Open Access

Citation: Yadav V, Sreekumar L, Guin K, Sanyal K (2018) Five pillars of centromeric chromatin in fungal pathogens. PLoS Pathog 14(8): e1007150. https://doi.org/10.1371/journal.ppat.1007150

Editor: Donald C Sheppard, McGill University, CANADA

Published: August 23, 2018

Copyright: © 2018 Yadav et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The Senior Research Fellowship from the Council of Scientific and Industrial Research (CSIR), Government of India (grant number 09/733 (0179)/2012-EMR-I) was received by VY. The Senior Research Fellowship from the Council of Scientific and Industrial Research (CSIR), Government of India (grant number 09/733(0178)/2012-EMR-I) was received by LSK. The Shyama Prasad Mukherjee Fellowship from the Council of Scientific and Industrial Research (CSIR), Government of India (grant number 07/733(0181)/2013-EMR-I) was received by KG. The Tata innovation fellowship (grant number BT/HRT/35/01/03/2017) was received by KS. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

DNA sequence and organization of DNA sequence elements

DNA sequence features provide the necessary template to act as a binding platform for kinetochore proteins. The genus Candida, which harbors several pathogenic species, presents a diverse array of centromere types. Candida glabrata carries point genetic centromeres, much like the 125-bp DNA sequence that serves as a fully functional point centromere in the budding yeast Saccharomyces cerevisiae [3–5]. Typically, genetic centromeres have specific and conserved DNA sequence motifs and confer mitotic stability to otherwise unstable plasmids carrying an autonomous replicating sequence (ARS) during cell division. Despite high-structural homology in DNA sequence elements, the point centromeres of C. glabrata are not fully functional in S. cerevisiae, suggesting that centromere function is species-specific [5, 6]. Short regional genetic centromeres of Candida tropicalis comprise a central core flanked by inverted repeats, similar to those of the fission yeast Schizosaccharomyces pombe [3, 7]. The sequence and orientation of these repeats are important for centromere function. Due to the presence of inverted repeats, the centromeres in C. tropicalis can acquire a hairpin loop-like secondary structure that might be crucial for kinetochore assembly. Candida albicans and Candida dubliniensis, on the other hand, possess unique and different centromere DNA sequences on each of their chromosomes [8, 9]. While C. tropicalis centromeres can stabilize an ARS plasmid,
indicating a role of DNA sequence in centromere identity, the same does not hold true for *C. albicans*. *Cryptococcus neoformans*, a basidiomycetous pathogen that diverged from *C. albicans* more than 900 million years ago, harbors large regional centromeres that are rich in centromere-specific retroelements [10–12]. While the presence of such retroelements hints toward the functional dependence on DNA sequence in centromere function, more studies are needed to explore such links. Some fungal centromeres possess specific DNA sequence features. For example, *Candida lusitaniae* (teleomorph *Clavispora lusitaniae*) and *Malassezia sympodialis* centromeres are present in highly AT-rich regions of the genome but lack any easily detectable conserved sequence motifs or repeats [13, 14]. Whether any AT-rich DNA sequence can act as a centromere in these organisms, similar to what is observed in diatoms [15], remains an open question.

**Centrochromatin—CENP-A and chromatin modifications**

The conserved centromere-specific histone H3 variant CENP-A/Cse4 is specifically present at all fungal centromeres identified to date but is largely excluded from other regions of the genome. *Mucor circinelloides* and *Phycomyces blakesleeanus* are notable exceptions in this regard, as they have no obvious CENP-A homologs, even though their centromeres are not yet physically mapped [16]. CENP-A is considered as the epigenetic determinant of centromere identity, as these molecules can seed the formation of a functional centromere in most organisms. This is supported by the fact that ectopic CENP-A incorporation can result in neocentromere formation, which is activated when an endogenous centromere becomes nonfunctional [17]. Although it is not well understood how CENP-A acts as the epigenetic determinant of the centromere, structural properties like a longer alpha N-terminal (αN) helix and the L1 loop region and biophysical features of the CENP-A nucleosome array, such as higher condensation properties, might be crucial for this role [18]. The process of CENP-A incorporation has been studied in *C. albicans*. Like other species, the CENP-A chaperone Holliday junction recognition protein/suppressor of chromosome mis-segregation 3 (HJURP/Scm3) is found to be crucial for CENP-A loading in *C. albicans* (our unpublished results). Regional centromeres harbor canonical histone H3 along with CENP-A. Post-translational modifications of histone H3 are crucial in forming a functional kinetochore [19]. Biochemical studies revealed the presence of heterochromatin histone marks such as dimethylation of histone H3 lysine 9 (H3K9diMe) across the centromeres of *C. neoformans* [10]. Apart from histone marks, DNA methylation has also been observed at the centromeres in *C. neoformans*, but its functional significance is unclear yet [10].

**Transcription and RNAi**

For a very long time, the centromere locus was considered heterochromatic and transcriptionally inert. While centromere regions are generally transcription poor, landmark studies in several yeast species revealed that small interfering RNAs (siRNAs) derived from pericentromeric regions are necessary for centromere function [20, 21]. These studies indicated that centromere transcription is permissible and has a functional significance. A pan–fungal analysis of RNA interference (RNAi) proteins revealed that a few species, including *C. glabrata* and *Ustilago maydis*, have lost all of the proteins required for functional RNAi during the course of evolution, whereas species including *C. albicans* and *C. tropicalis* harbor a cryptic RNAi machinery [22]. A recent study in the pathogenic *Cryptococcus* species complex correlated the loss of RNAi with the length of centromeres, thereby proposing that RNAi helps in maintaining long repetitive, transposon-rich centromeres [10]. Whether the RNAi machinery has a functional significance in the centromere biology of this species complex and other fungal...
apathogens remains unexplored. Apart from RNAi, centromere transcription can also play a functional role through long noncoding RNA (ncRNA). Indeed, pervasive levels of transcription have been documented in various fungal species [23, 24]. However, the transcripts generated from the centromeres are significantly low in number compared to the rest of the genome, as shown in the Cryptococcus and Ustilago species complex [10].

Replication and repair
Unlike metazoans, most fungal species have early replicating centromeres [25–27]. This temporally distinct replication timing not only allows better tolerance toward replication stress but also ensures proper kinetochore assembly at the centromeres. In C. albicans, centromeres replicate early in every synthesis phase (S-phase) and are associated with an early firing replicating origin [26]. Additionally, the formation of a neocentromere advances the replication time of the flanking region by activating an early replicating origin. This proximity effect was explained by a replication-coupled repair mechanism in a kinetochore-dependent manner [28]. Centromere-proximal origins stall randomly at the centromere, leading to accumulation of single-stranded DNA, which then recruits the homologous recombination proteins such as Rad51 and Rad52. These proteins physically interact with CENP-A in C. albicans and load it at the site of stalled replication forks; that is, centromeres. How this process is regulated to occur only during S-phase remains unknown, with possible implications for the CENP-A chaperone Scm3. Based on studies in many other nonpathogenic fungal species, the physical proximity of a partitioning locus (centromere) and an initiator site (replication origin) is relevant when one dissects the functional aspects of genome maintenance. However, evidence toward this connection is just beginning to emerge.

Spatial location
Most fungal centromeres are clustered near spindle pole bodies (SPBs). This association may result in folding back of chromosomes and positioning them such that telomeres are juxtaposed in the interphase nucleus, giving rise to the Rabl conformation [29]. This phenomenon has been shown to occur in both animal and plant pathogens including C. albicans, C. tropica lis, and Fusarium graminearum. Using chromosome conformation capture (3C) experiments, clustered centromere DNA regions were shown to be present in close spatial proximity, leading to physical interactions between different centromeres [30–32]. It has been proposed that the clustering of centromeres aids in determining the site of centromere formation in these organisms. According to this hypothesis, a part of the nucleus is enriched with a pool of CENP-A proteins to form a CENP-A–rich zone or CENP-A cloud [33, 34]. It was proposed that the region of a chromosome that is near this CENP-A cloud would attract a higher level of CENP-A and thus serves as a preferred site for centromere formation. In S. cerevisiae, for example, a locally enriched population of accessory CENP-A molecules at pericentric chromatin has been shown to serve as a reservoir for rapid incorporation of CENP-A in the event of premature eviction from centromeres [35]. Further evidence supporting the CENP-A cloud hypothesis stems from studies in C. albicans in which neocentromeres were formed close to the native centromere in most cases [34]. In addition, neocentromeres in C. albicans change the spatial location to be a part of the centromere cluster by 3C experiments [31]. Interestingly, centromeres were found to be unclustered in premitotic C. neoformans cells that eventually cluster at the onset of mitosis [36]. Whether this centromere clustering also arises as a result of physical interactions among centromeres is not yet known.

Overall, here we summarize five key determinants among many that are needed for centromere identity in fungal pathogens (Fig 1). These factors may work sequentially or in parallel to
Fig 1. **Five key determinants of centromere identity in pathogenic fungi.** (A) The length and type of centromeres in various pathogenic fungi have been depicted in a cladogram. *Candida glabrata* and *Malassezia sympodialis* have point centromeres (blue); *C. albicans*, *C. tropicalis*, and *Cryptococcus deuterogattii* have small regional centromeres (green); and *C. neoformans* and *C. deneformans* harbor large regional repetitive centromeres (pink). (B) Centromeric heterochromatin or centrochromatin refers to the various chemical modifications associated with the histones (both canonical and variant) as well as DNA at the centromeric locus. For example, centromere DNA is methylated at cytosines, and CENP-A nucleosomes are interspersed with H3K9-dimethylated nucleosomes in *C. neoformans*. (C) Transcription and a functional RNAi machinery at the centromere are required to preserve centromere identity in some organisms. The centromeres in the *Ustilago* and *Cryptococcus* species complexes are poorly transcribed. In *C. neoformans*, centromeric transcription is regulated by the RNAi machinery that silences centromeric retrotransposons to stabilize centromere structure. Unprocessed long noncoding RNA, whose function is unknown, is also produced from centromeres in this species. (D) DNA replication and repair proteins are known to play a key role in *C. albicans* centromere stability. The centromere proximal origins help in maintaining centromere function in this organism. Replication forks converging toward the centromere stall randomly, which leads to accumulation of single-stranded DNA (ssDNA) that recruits repair proteins like Rad51 and Rad52, along with new CENP-A molecules in *C. albicans*. (E) The spatial location of centromeres within the nucleus determines its activity and interaction with other nuclear subcompartments. All centromeres are clustered toward the nuclear periphery to form a CENP-A–rich zone in *C. albicans* and many other pathogenic fungi. This preferential spatial distribution helps to determine the site of centromere assembly in every cell cycle. CENP-A, centromere protein A; H3K9, histone H3 lysine 9; Rad, radiation sensitive; RNAi, RNA interference; ssDNA, single-stranded DNA.

https://doi.org/10.1371/journal.ppat.1007150.g001
ensure that centromere identity is maintained at all times in every cell cycle. For example, the repeats present at the centromere can form secondary structures, which may lead to double-stranded DNA breaks during replication and attract more CENP-A to these regions through DNA-repair pathways. The nuclear subdomain near SPBs is usually heterochromatic in nature, and the presence of the centromere cluster at a nuclear peripheral region contributes to poor transcription. Thus, a combination of factors can lead to more efficient CENP-A incorporation at the centromere. The CENP-A incorporation affects assembly of kinetochore proteins; for example, in C. neoformans, CENP-A assembly initiates the loading of other proteins, giving rise to a sequential kinetochore assembly in this organism [36]. On the contrary, C. albicans shows an interdependent kinetochore formation in which probably all kinetochore proteins assemble as a single complex [37]. However, our knowledge of the underlying mechanisms of CENP-A loading is still at its infancy, given the number of fungal species known to exist. The lack of knowledge is primarily due to difficulties associated with genetic manipulation of most pathogenic fungal species. Probing into the molecular mechanisms of centromere identity using diverse fungal species will yield significant insight into the structure-function-evolution of centromeres during speciation. Studies in Candida and Cryptococcus species complexes revealed that centromere-mediated recombination might have contributed to variations in the centromere structure and sequence [7, 10, 38]. Additionally, centromere sequences can also help in identification and classification of two or more closely related species. For example, C. albicans and C. dubliniensis centromere sequences show a high level of divergence while the rest of the genome sequences are highly homologous [9]. Discovery of highly efficient gene-targeting technologies such as the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR-Cas9) system for various Candida and Cryptococcus species as well as the application of efficient and improved DNA sequencing technologies for obtaining better genome assemblies are some of the significant recent developments that will advance pathogenic fungal molecular genetic research.

**Acknowledgments**

We thank members of the Molecular Mycology Laboratory, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India and the Heitman lab, Duke University, United States of America for useful discussion, critical reading, and comments on the manuscript.

**References**

1. Flemming W. Zellsbestand, Kern und Zelltheilung: F. C. W. Vogel, Leipzig: 1882.
2. Henikoff S, Ahmad K, Malik HS. The centromere paradox: Stable inheritance with rapidly evolving DNA. Science. 2001; 293(5532):1098–102. Epub 2001/08/11. https://doi.org/10.1126/science.1062939 PMID: 11498581.
3. Roy B, Sanyal K. Diversity in requirement of genetic and epigenetic factors for centromere function in fungi. Eukaryot Cell. 2011; 10(11):1384–95. Epub 2011/09/13. https://doi.org/10.1128/EC.05165-11 PMID: 21908596.
4. Clarke L, Carbon J. Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature. 1980; 287(5782):504–9. Epub 1980/10/09. PMID: 6999364.
5. Kitada K, Yamaguchi E, Arisawa M. Isolation of a Candida glabrata centromere and its use in construction of plasmid vectors. Gene. 1996; 175(1–2):105–8. PMID: 8917084.
6. Kitada K, Yamaguchi E, Hamada K, Arisawa M. Structural analysis of a Candida glabrata centromere and its functional homology to the Saccharomyces cerevisiae centromere. Current genetics. 1997; 31 (2):122–7. PMID: 9021128.
7. Chatterjee G, Sankaranarayanan SR, Guin K, Thattikutty Y, Padmanabhan S, Siddharthan R, et al. Repeat-associated fission yeast-like regional centromeres in the ascomycete budding yeast Candida tropicalis. PLoS Genet. 2016; 12(2):e1005839. https://doi.org/10.1371/journal.pgen.1005839 PMID: 26845548.
8. Sanyal K, Baum M, Carbon J. Centromeric DNA sequences in the pathogenic yeast Candida albicans are all different and unique. Proc Natl Acad Sci U S A. 2004; 101(31):11374–9. https://doi.org/10.1073/pnas.0404318101 PMID: 15272074

9. Padmanabhan S, Thakur J, Siddharthan R, Sanyal K. Rapid evolution of Cse4p-rich centromeric DNA sequences in closely related pathogenic yeasts, Candida albicans and Candida dubliniensis. Proc Natl Acad Sci U S A. 2008; 105(50):19797–802. Epub 2008/12/09. https://doi.org/10.1073/pnas.0809770105 PMID: 19060206.

10. Yadav V, Sun S, Billmyre RB, Thimmappa BC, Shea T, Lintner R, et al. RNAi is a critical determinant of centromere evolution in closely related fungi. Proc Natl Acad Sci U S A. 2018; 115(12):3108–13. https://doi.org/10.1073/pnas.1713725115 PMID: 29507212.

11. Janbon G, Ormerod KL, Paulet D, Byrnes EJ 3rd, Yadav V, Chatterjee G, et al. Analysis of the genome and transcriptome of Cryptococcus neoformans var. grubii reveals complex RNA expression and micro-evolution leading to virulence attenuation. PLoS Genet. 2014; 10(4):e1004261. Epub 2014/04/02. https://doi.org/10.1371/journal.pgen.1004261 PMID: 24743168.

12. Loftus BJ, Fung E, Roncaglia P, Rowley D, Amedeo P, Bruno D, et al. The genome of the basidiomycete yeast and human pathogen Cryptococcus neoformans. Science. 2005; 307(5713):1321–4. https://doi.org/10.1126/science.1103773 PMID: 15653466.

13. Kapoor S, Zhu L, Froyd C, Liu T, Rusche LN. Regional centromeres in the yeast Candida lusitaniae lack pericentromeric heterochromatin. Proc Natl Acad Sci U S A. 2015; 112(39):12139–44. https://doi.org/10.1073/pnas.1508749112 PMID: 26373987.

14. Zhu Y, Engstrom PG, Tellgren-Roth C, Baudo CD, Kennell JC, Sun S, et al. Proteogenomics produces comprehensive and highly accurate protein-coding gene annotation in a complete genome assembly of Malassezia sympodialis. Nucleic Acids Res. 2017; 45(5):2629–43. https://doi.org/10.1093/nar/gkx006 PMID: 28100699.

15. Diner RE, Noddings CM, Lian NC, Kang AK, McQuaid JB, Jablanovic J, et al. Diatom centromeres suggest a mechanism for nuclear DNA acquisition. Proc Natl Acad Sci U S A. 2017; 114(29):E6015–E24. https://doi.org/10.1073/pnas.1700764114 PMID: 28673887.

16. van Hooff JJ, Tromer E, van Wijk LM, Snel B, Kops GJ. Evolutionary dynamics of the kinetochore network in eukaryotes as revealed by comparative genomics. EMBO Rep. 2017; 18(9):1559–71. https://doi.org/10.15252/embr.201744102 PMID: 28642229.

17. Steiner FA, Henikoff S. Diversity in the organization of centromeric chromatin. Curr Opin Genet Dev. 2015; 31:28–35. https://doi.org/10.1016/j.gde.2015.03.010 PMID: 25956076.

18. Bui M, Walkiewicz MP, Dimitriadis EK, Dalal Y. The CENP-A nucleosome: a battle between Dr Jekyll and Mr Hyde. Nucleus. 2013; 4(1):37–42. https://doi.org/10.4161/nuc1.23588 PMID: 23324462.

19. Fukagawa T. Critical histone post-translational modifications for centromere function and propagation. Cell Cycle. 2017;1–7. https://doi.org/10.1080/15384101.2017.1325044 PMID: 28598241.

20. Pidoux AL, Allshire RC. The role of heterochromatin in centromere function. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2005; 360(1555):569–79. https://doi.org/10.1098/rstb.2004.1611 PMID: 15905142.

21. Volpe T, Schramke V, Hamilton GL, White SA, Teng G, Martienssen RA, et al. RNA interference is required for normal centromere function in fission yeast. Chromosome Res. 2003; 11(2):137–46. PMID: 12733640.

22. Nakayashiki H, Kadotani N, Mayama S. Evolution and diversification of RNA silencing proteins in fungi. J Mol Evol. 2006; 63(1):127–35. https://doi.org/10.1007/s00239-005-0257-2 PMID: 16786437.

23. Choi ES, Straflors A, Castillo AG, Durand-Dubief M, Ekwall K, Allshire RC. Identification of noncoding transcripts from within CENP-A chromatin at fission yeast centromeres. J Biol Chem. 2011; 286(26):23600–7. https://doi.org/10.1074/jbc.M111.228510 PMID: 21531710.

24. Ohkuni K, Kitagawa K. Endogenous transcription at the centromere facilitates centromere activity in budding yeast. Curr Biol. 2011; 21(20):1695–703. https://doi.org/10.1016/j.cub.2011.08.056 PMID: 22000103.

25. Kim SM, Dubey DD, Huberman JA. Early-replicating heterochromatin. Genes Dev. 2003; 17(3):330–5. https://doi.org/10.1101/gad.1046203 PMID: 12569122.

26. Koren A, Tsai HJ, Tirosh I, Burrell LS, Barkai N, Berman J. Epigenetically-inherited centromere and neocentromere DNA replicates earliest in S-phase. PLoS Genet. 2010; 6(8):e1001068. https://doi.org/10.1371/journal.pgen.1001068 PMID: 20808889.

27. Raghuraman MK, Winzeler EA, Collingwood D, Hunt S, Wodicka L, Conway A, et al. Replication dynamics of the yeast genome. Science. 2001; 294(5540):115–21. https://doi.org/10.1126/science.294.5540.115 PMID: 11588253.
28. Mitra S, Gomez-Raja J, Larriba G, Dubey DD, Sanyal K. Rad51-R ad52 mediated maintenance of centromeric chromatin in Candida albicans. PLoS Genet. 2014; 10(4):e1004344. https://doi.org/10.1371/journal.pgen.1004344 PMID: 24762765.

29. Rabl C. Ueber Zelltheilung. Morphologisches Jahrbuch. 1885; 10:214–300.

30. Duan Z, Andronescu M, Schultz K, Mollwain S, Kim YJ, Lee C, et al. A three-dimensional model of the yeast genome. Nature. 2010; 465(7296):363–7. Epub 2010/05/04. https://doi.org/10.1038/nature08973 PMID: 20436457.

31. Burrack LS, Hutton HF, Matter KJ, Clancey SA, Liachko I, Plemons AE, et al. Neocentromeres provide chromosome segregation accuracy and centromere clustering to multiple loci along a Candida albicans chromosome. PLoS Genet. 2016; 12(9):e1006317. https://doi.org/10.1371/journal.pgen.1006317 PMID: 27662467.

32. Anderson M, Haase J, Yeh E, Bloom K. Function and assembly of DNA looping, clustering, and microtubule attachment complexes within a eukaryotic kinetochore. Mol Biol Cell. 2009; 20(19):4131–9. https://doi.org/10.1091/mbc.E09-05-0359 PMID: 19656849.

33. Fukagawa T, Earnshaw WC. The centromere: chromatin foundation for the kinetochore machinery. Dev Cell. 2014; 30(5):496–508. https://doi.org/10.1016/j.devcel.2014.08.016 PMID: 25203206.

34. Thakur J, Sanyal K. Efficient neocentromere formation is suppressed by gene conversion to maintain centromere function at native physical chromosomal loci in Candida albicans. Genome Res. 2013; 23(4):638–52. Epub 2013/02/27. https://doi.org/10.1101/gr.141614.112 PMID: 23439889.

35. Haase J, Mishra PK, Stephens A, Haggerty R, Quammen C, Taylor RM 2nd, et al. A 3D map of the yeast kinetochore reveals the presence of core and accessory centromere-specific histone. Curr Biol. 2013; 23(19):1939–44. Epub 2013/10/01. https://doi.org/10.1016/j.cub.2013.07.083 PMID: 24076245.

36. Kozubowski L, Yadav V, Chatterjee G, Sridhar S, Yamaguchi M, Kawamoto S, et al. Ordered kinetochore assembly in the human-pathogenic basidiomycetous yeast Cryptococcus neoformans. mBio. 2013; 4(5):e00614–13. Epub 2013/10/03. https://doi.org/10.1128/mBio.00614-13 PMID: 24085781.

37. Thakur J, Sanyal K. A coordinated interdependent protein circuitry stabilizes the kinetochore ensemble to protect CENP-A in the human pathogenic yeast Candida albicans. PLoS Genet. 2012; 8(4): e1002661. Epub 2012/04/27. https://doi.org/10.1371/journal.pgen.1002661 PMID: 22536162.

38. Sun S, Yadav V, Billmyre RB, Cuomo CA, Nowrousian M, Wang L, et al. Fungal genome and mating system transitions facilitated by chromosomal translocations involving intercentromeric recombination. PLoS Biol. 2017; 15(8):e2002527. https://doi.org/10.1371/journal.pbio.2002527 PMID: 28800596.