COMMENTARY

The Curious Case of Nonrepetitive Centromeric DNA Sequences in *Candida auris* and Related Species

Alexander Lorenz,[a] Nicolas Papon[b]

[a]Institute of Medical Sciences (IMS), University of Aberdeen, Aberdeen, United Kingdom
[b]Univ Angers, Univ Brest, GEIHP, SFR ICAT, Angers, France

ABSTRACT 2009 saw the first description of *Candida auris*, a yeast pathogen of humans. *C. auris* has since grown into a global problem in intensive care settings, where it causes systemic infections in patients with underlying health issues. Recent whole-genome sequencing has discerned five *C. auris* clades with distinct phenotypic features which display genomic divergence on a DNA sequence and a chromosome structure level. In the absence of sexual reproduction in *C. auris*, the mechanism(s) behind the rapid genomic evolution of this emerging killer yeast has remained obscure. Yet, one important bit of information about chromosome organization was missing, the identification of the centromeres. In a recent study, Sanyal and coworkers (A. Narayanan, R. N. Vadnala, P. Ganguly, P. Selvakumar, et al., mBio 12:e00905-21, 2021, https://doi.org/10.1128/mBio.00905-21) filled this knowledge gap by mapping the centromeres in *C. auris* and its close relatives. This represents a major advance in the chromosome biology of the *Candida*/*Clavispora* clade.

KEYWORDS *Candida auris*, centromeres, chromosomes, karyotype evolution, centromere

Centromeres are essential features of chromosomes; they are sites where kinetochores are assembled which connect the chromosomes to the microtubular spindle to segregate them during mitosis and meiosis (1). Considering their conserved and fundamental role, a surprising diversity of centromere sequences and types is present in eukaryotes. In fungi, different centromere categories have been described, ranging from the tiny sequence-specific point centromeres of *Saccharomyces*, which represent the pinnacle of genome streamlining, to the large epigenetically defined regional centromeres flanked by repetitive DNA in *Schizosaccharomyces* (Fig. 1) (2). Common to centromeres in most organisms is a special histone H3 variant called CENP-A (centromeric protein A, usually called Cse4 in fungi), which replaces histone H3 in centromeric nucleosomes, where it forms the recruiting platform for kinetochores (1, 2). Thus, in recent studies centromeric DNA in fungi was identified through its propensity to be bound by CENP-A and other kinetochore components (3–7). Fungal centromeric DNAs are often found in large regions of low GC content, so-called GC-poor troughs (8). A notable exception to the latter is *Candida albicans*, which harbors small regional centromeres (~3 to 5 kb) defined by unique nonrepetitive sequences (Fig. 1) (3, 8). In *C. albicans*, the pericentromeric regions can contain repetitive DNA and have features reminiscent of heterochromatin as gene expression is repressed in the vicinity of centromeres and some fitting histone modifications are present (3, 9). *Clavispora* (*Candida*) *lusitaniae* has somewhat similarly sized (~4.0 to 4.5 kb) centromeres with unique sequences (Fig. 1), but these are located in GC-poor troughs, lack typical heterochromatin marks, and do not suppress gene expression in their vicinity (4).

Candida auris, *Candida haemulonii*, *Candida pseudohaemulonii*, and *Candida duobushaemulonii* make up the *Candida haemulonii* complex which, together with *C. lusitaniae*,

Citation Lorenz A, Papon N. 2021. The curious case of nonrepetitive centromeric DNA sequences in *Candida auris* and related species. mBio 12:e01476-21. https://doi.org/10.1128/mBio.01476-21.

Address correspondence to Alexander Lorenz, a.lorenz@abdn.ac.uk.

For the article discussed, see https://doi.org/10.1128/mBio.00905-21.

The views expressed in this article do not necessarily reflect the views of the journal or of ASM.

Published 3 August 2021

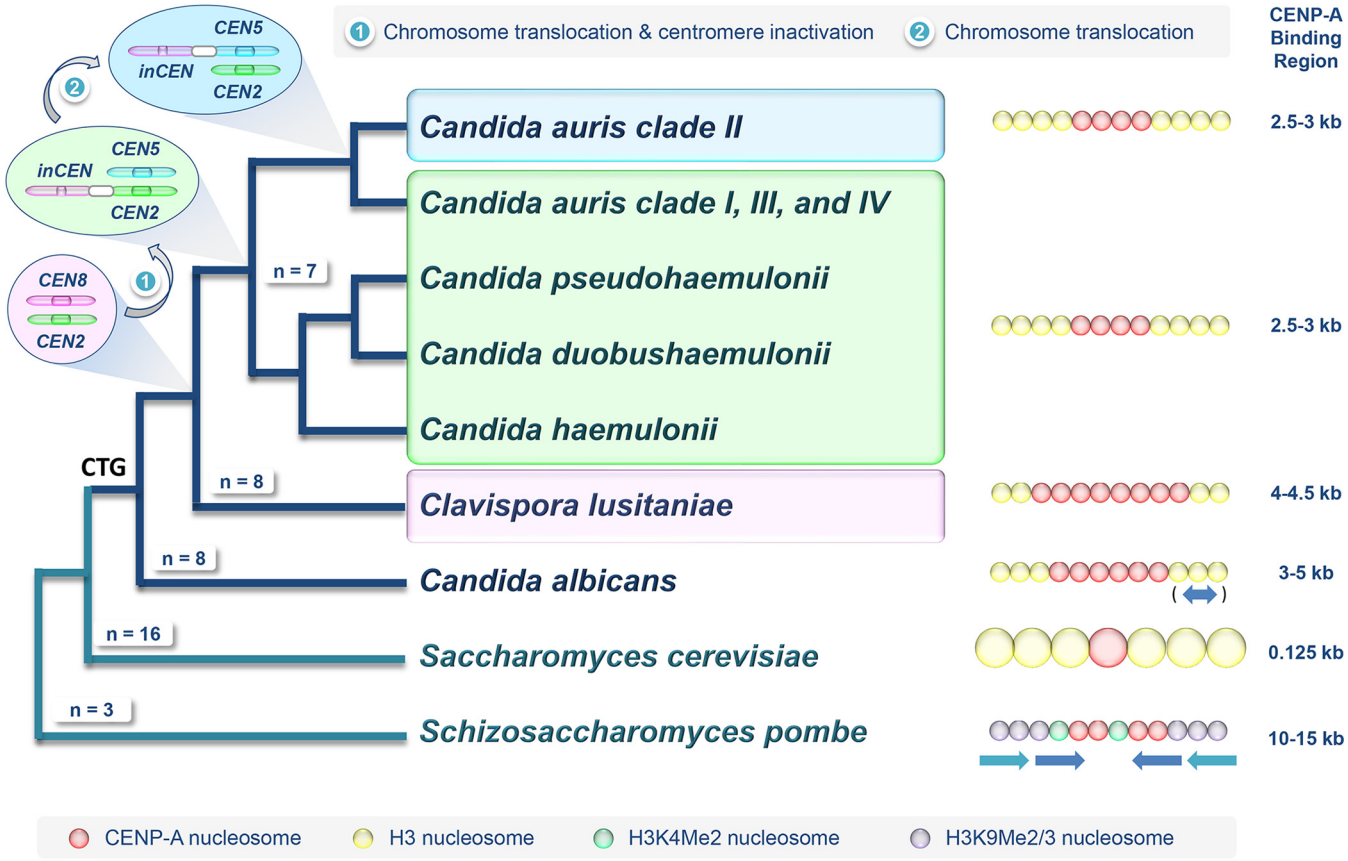

FIG 1 Tracing the evolution of centromeres in *C. auris* and other ascomycetous species. Branch lengths of the phylogenetic tree are arbitrary, but topologies resemble the current understanding of the relationships of the Metschnikowiaceae (including *C. auris*, *C. haemulonii*, *C. pseudohaemulonii*, and *C. duobushaemulonii*, and *C. lusitaniae*) (16). The model yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, as well as *C. albicans*, are also indicated. A first event of chromosome translocation and centromere inactivation in an ancestor of Metschnikowiaceae likely resulted in chromosome number reduction in species belonging to the *C. haemulonii* complex (including *C. haemulonii*, *C. pseudohaemulonii*, *C. duobushaemulonii*, and *C. auris*). Another chromosome translocation further repositions the inactive centromere in *C. auris* clade II. Centromeres are numbered using *C. lusitaniae* as the reference. On the right side a schematic representation of centromere types (not to scale) and lengths of the CENP-A-positive core centromeres for each species/clade are given. Arrows for *C. albicans* and *Schizosaccharomyces pombe* indicate DNA repeats in pericentromeres; in the former, repeats do not occur in every pericentromere.

belongs to the *Clavispora* clade of the Metschnikowiaceae (10). This group of *Clavispora* yeasts is part of the *Candida* CTG clade which also contains *C. albicans* and its close relatives (11). Since its first identification in 2009 (12), *C. auris* has become known as a nosocomial yeast pathogen (13). The species has been subdivided into 5 geographical clades; the genomes of the different clades show substantial differences between them on a DNA sequence and a chromosome structure level (14–18). The karyotype (chromosome complement) of *C. auris* is thus fairly plastic and can change quickly upon stress exposure (17). This karyotype plasticity could at least in part explain how *C. auris* generates genetic diversity in the absence of sexual reproduction and meiosis (19). What role centromeres play in the karyotype diversification of *C. auris* and other *Clavispora* yeasts could not be answered until recently, when Sanyal and coworkers mapped the centromeric DNA by chromatin immunoprecipitation approaches of CENP-A (7).

To identify the centromeric DNA sequences of *C. auris*, Narayanan et al. (7) employed ChIP-seq (chromatin immunoprecipitation followed by next-generation sequencing) of the centromere-specific histone H3 variant Cse4 (also known as CENP-A). *C. auris CSE4* was identified by its homology to *C. albicans CSE4*. A C-terminal protein A tag was then introduced at the endogenous locus of *CSE4* in a representative *C. auris* clade I strain (B8441). To check the correct localization of Cse4 and the functionality of the tagged version, the authors performed indirect immunofluorescence in mitotic *C. auris* cells carrying

the Cse4-protein A construct. As expected, Cse4-protein A formed one discrete focus per nucleus in nondividing cells and a focus each at the leading edge of dividing daughter nuclei (7); this is indicative of centromere clustering (20). Sequencing of the DNA associated with Cse4-protein A revealed a single peak each on 7 different contigs of the published B8441 genome sequence. The peaks of Cse4-binding (i.e., centromeric) DNA were between ~2.5 and ~2.9 kb in size (Fig. 1) and comprised unique DNA sequences which show no homologies to each other or to any sequences within the genome. Centromeric DNAs were found in GC-poor troughs in all instances; the only repeat sequences detected were 40-bp poly(A) and poly(T) stretches in or near all centromeric DNA regions. Genes in the vicinity of centromeres did not show suppressed expression, which according to the authors suggests that pericentromeric heterochromatin is absent in *C. auris* (7). This situation is thus very similar to *C. lusitaniae* (see above) (4).

The centromere sizes and positions were corroborated by ChIP-qPCR (ChIP followed by quantitative PCR) in the clade I strain (B8441) harboring the protein A-tagged Cse4 (7). This strategy was then employed to map centromeres in *C. auris* clade II, III, and IV strains as well as in strains from the other species of the *C. haemulonii* complex to confirm centromere predictions from bioinformatic analyses based on ORF (open reading frame) content, gene synteny, and GC content. Within *C. auris*, clade II differs most from the other three clades due to major chromosomal rearrangements (7, 18). Interestingly, the related clade II isolates B11220 and CBS10913T contain a large tandem duplication which encompasses the centromeric DNA (7, 18); technically, it is not possible to discern which of those two centromere candidates is the active one. Other clade II strains do not harbor this duplication (18). Intriguingly, centromeric DNA sequences evolve more quickly than intergenic regions as suggested by a high incidence of substitution mutations observed between the different clades (7). The karyotypes of clades I, III, and IV are quite similar, with only 2 major chromosome translocation events detected between clades III and IV. Clade II has undergone multiple chromosome rearrangements, and several of those map to the centromeric regions and break the gene synteny surrounding them (7, 18). This suggests that centromeres can be sources of karyotype diversification in *C. auris*.

Strains of the other *C. haemulonii* complex species displayed the same centromere properties as *C. auris* and *C. lusitaniae* (4, 7). These data together with published genome assemblies allowed the authors to present a convincing hypothesis on chromosome/karyotype evolution of *Clavispora* yeasts. *C. duobushaemulonii* is the only other species with a chromosome-level genome assembly, and its karyotype is very similar to karyotypes of *C. auris* clades I, III, and IV. This supports the hypothesis that the *C. auris* clade II karyotype is a derived state, i.e., the *C. auris* clade likely split off the remaining *C. haemulonii* complex fairly recently. A key difference between *C. lusitaniae* and the species of the *C. haemulonii* complex is that the former has 8 chromosomes whereas the studied representatives of the latter have 7 chromosomes. Analysis indicated that *C. lusitaniae* chromosome 8 has been translocated as 3 fragments onto other chromosomes in the last common ancestor of the *C. haemulonii* complex species. The *C. lusitaniae* centromere 8 can still be detected in the genomes of *C. haemulonii* complex species (Fig. 1) but is inactive in *C. haemulonii* complex species as it has undergone substantial DNA sequence attrition and fails to recruit Cse4 (7). Finally, the authors explore the published genomes of additional species belonging to the *Clavispora* clade, all of which contain 8 potential centromeric regions defined by gene synteny in GC-poor troughs. In *Candida heveicola*, one of the 8 centromeres showed DNA sequence attrition similar to *C. haemulonii* complex species, whereas the other species studied (*Candida blattae*, *Candida intermedia*, *Candida oregonensis*) had 8 full-length centromeres like *C. lusitaniae*.

The nonrepetitive nature of centromeric DNA in *Clavispora* raises several intriguing questions about their chromosome biology. It is generally accepted that centromeric DNA repeats are a source of genome instability via ectopic recombination, but it has also

been suggested that recombination between centromeric repeats is important for their maintenance (21). When centromeric DNA sequences are unique and nonrepetitive, as in *Clavispora* species, how do they contribute to karyotype diversification? Sanyal and coworkers suggest that due to the proximity of centromeric DNAs in the centromere clusters (7, 20), chromosome rearrangements could be a consequence of replication fork perturbations near centromeres. Maybe the poly(A) and poly(T) stretches play a role as sites for ectopic recombination? In the absence of common DNA sequence signatures, how is centromere identity defined? In *C. albicans*, centromeres seem to be specified epigenetically as unchromatinized centromeric DNA sequence does not establish a functional centromere (22). But then *C. albicans* centromeres also display identifying, partially heterochromatic, histone modifications (9); mechanistically imposing such modifications *de novo* presumably enables the formation of neocentromeres (23). Histones at *C. lusitaniae* centromeres also have specific modifications, albeit none which are typical for heterochromatin (4). Do centromeres of other *Clavispora* yeasts, including *C. auris*, harbor the same histone modifications as *C. lusitaniae*? Are these histone modifications indeed defining centromere identity? And finally, if the underlying DNA sequence is not important, why is centromeric DNA attrition observed after chromosome fusions, as is the case with the centromere of *C. lusitaniae* chromosome 8 in *C. haemulonii* complex species (7)? Answering these questions will not only give us insight into the fascinating chromosome biology of these opportunistic pathogens but will also provide clues to how genetic and genomic diversity is generated to evolve virulence and antimicrobial resistance traits.

ACKNOWLEDGMENT

We thank Adele L. Marston (University of Edinburgh, United Kingdom) for critical reading of the manuscript.

REFERENCES

1. Talbert PB, Henikoff S. 2020. What makes a centromere? Exp Cell Res 389:111895. https://doi.org/10.1016/j.yexcr.2020.111895.
2. Guin K, Sreekumar L, Sanyal K. 2020. Implications of the evolutionary trajectory of centromeres in the fungal kingdom. Annu Rev Microbiol 74:835–853. https://doi.org/10.1146/annurev-micro-011720-122512.
3. Sanyal K, Baum M, Carbon J. 2004. Centromeric DNA sequences in the pathogenic yeast *Candida albicans* are all different and unique. Proc Natl Acad Sci U S A 101:11374–11379. https://doi.org/10.1073/pnas.0404318101.
4. Kapoor S, Zhu L, Froyd C, Liu T, Rusche LN. 2015. Regional centromeres in the yeast *Candida lusitaniae* lack pericentromeric heterochromatin. Proc Natl Acad Sci U S A 112:12139–12144. https://doi.org/10.1073/pnas.1508749112.
5. Coughlan AY, Hanson SJ, Byrne KP, Wolfe KH. 2016. Centromeres of the yeast *Komagataella phaffii* (*Pichia pastoris*) have a simple inverted-repeat structure. Genome Biol Evol 8:2482–2492. https://doi.org/10.1093/gbe/evw178.
6. Chatterjee G, Sankaranarayanan SR, Guin K, Thattikota Y, Padmanabhan S, Siddharthan R, Sanyal K. 2016. Repeat-associated fission yeast-like regional centromeres in the ascomycetous budding east *Candida tropicalis*. PLoS Genet 12:e1005839. https://doi.org/10.1371/journal.pgen.1005839.
7. Narayanan A, Vadnala RN, Ganguly P, Selvakumar P, Rudramurthy SM, Prasad R, Chakrabarti A, Siddharthan R, Sanyal K. 2021. Functional and comparative analysis of centromeres reveals clade-specific genome rearrangements in *Candida auris* and a chromosome number change in related species. mBio 12:e00905-21. https://doi.org/10.1128/mBio.00905-21.
8. Lynch DB, Logue ME, Butler G, Wolfe KH. 2010. Chromosomal G + C content evolution in yeasts: systematic interspecies differences, and GC-poor troughs at centromeres. Genome Biol Evol 2:572–583. https://doi.org/10.1093/gbe/evq042.
9. Freire-Benéitez V, Price RJ, Buscaino A. 2016. The chromatin of *Candida albicans* pericentromeres bears features of both euchromatin and heterochromatin. Front Microbiol 7:759. https://doi.org/10.3389/fmicb.2016.00759.
10. Daniel H-M, Lachance M-A, Kurtzman CP. 2014. On the reclassification of species assigned to *Candida* and other anamorphic ascomycetous yeast genera based on phylogenetic circumscription. Antonie Van Leeuwenhoek 106:67–84. https://doi.org/10.1007/s10482-014-0170-z.
11. Turner SA, Butler G. 2014. The *Candida* pathogenic species complex. Cold Spring Harb Perspect Med 4:a019778. https://doi.org/10.1101/cshperspect.a019778.
12. Satoh K, Makimura K, Hasumi Y, Nishiyama Y, Uchida K, Yamaguchi H. 2009. *Candida auris* sp. nov., a novel ascomycetous yeast isolated from the external ear canal of an inpatient in a Japanese hospital. Microbiol Immunol 53:41–44. https://doi.org/10.1111/j.1348-0421.2008.00083.x.
13. Vila T, Sultan AS, Montelongo-Jauregui D, Jabra-Rizk MA. 2020. *Candida auris*: a fungus with identity crisis. Pathog Dis 78:ftaa034. https://doi.org/10.1093/femspd/ftaa034.
14. Lockhart SR, Etienne KA, Vallabhaneni S, Farooqi J, Chowdhary A, Govender NP, Colombo AL, Calvo B, Cuomo CA, Desjardins CA, Berkow EL, Castanheira M, Magobo RE, Jabeen K, Asghar RJ, Meis JF, Jackson B, Chiller T, Litvintseva AP. 2017. Simultaneous emergence of multidrug-resistant *Candida auris* on 3 continents confirmed by whole-genome sequencing and epidemiological analyses. Clin Infect Dis 64:134–140. https://doi.org/10.1093/cid/ciw691.
15. Chow NA, Muñoz JF, Gade L, Berkow EL, Li X, Welsh RMM, Forsberg K, Lockhart SR, Adam R, Alanio A, Alastruey-Izquierdo A, Althawadi S, Araúz ABB, Ben-Ami R, Bharat A, Calvo B, Desnos-Ollivier M, Escandón P, Gardam D, Gunturu R, Heath CH, Kurzai O, Martin R, Litvintseva AP, Cuomo CA. 2020. Tracing the evolutionary history and global expansion of *Candida auris* using population genomic analyses. mBio 11:e03364-19. https://doi.org/10.1128/mBio.03364-19.
16. Muñoz JF, Gade L, Chow NA, Loparev VN, Juieng P, Berkow EL, Farrer RA, Litvintseva AP, Cuomo CA. 2018. Genomic insights into multidrug-resistance, mating and virulence in *Candida auris* and related emerging species. Nat Commun 9:5346. https://doi.org/10.1038/s41467-018-07779-6.
17. Bravo Ruiz G, Ross ZK, Holmes E, Schelenz S, Gow NAR, Lorenz A. 2019. Rapid and extensive karyotype diversification in haploid clinical *Candida auris* isolates. Curr Genet 65:1217–1228. https://doi.org/10.1007/s00294-019-00976-w.
18. Muñoz JF, Welsh RM, Shea T, Batra D, Gade L, Howard D, Rowe LA, Meis JF, Litvintseva AP, Cuomo CA. 2021. Clade-specific chromosomal rearrangements and loss of subtelomeric adhesins in *Candida auris*. Genetics 218:iyab029. https://doi.org/10.1093/genetics/iyab029.

19. Ross ZK, Lorenz A. 2020. Is *Candida auris* sexual? PLoS Pathog 16:e1009094. https://doi.org/10.1371/journal.ppat.1009094.
20. Jin QW, Fuchs J, Loidl J. 2000. Centromere clustering is a major determinant of yeast interphase nuclear organization. J Cell Sci 113:1903–1912. https://doi.org/10.1242/jcs.113.11.1903.
21. McFarlane RJ, Humphrey TC. 2010. A role for recombination in centromere function. Trends Genet 26:209–213. https://doi.org/10.1016/j.tig.2010.02.005.
22. Baum M, Sanyal K, Mishra PK, Thaler N, Carbon J. 2006. Formation of functional centromeric chromatin is specified epigenetically in *Candida albicans*. Proc Natl Acad Sci U S A 103:14877–14882. https://doi.org/10.1073/pnas.0606958103.
23. Ketel C, Wang HSW, McClellan M, Bouchonville K, Selmecki A, Lahav T, Gerami-Nejad M, Berman J. 2009. Neocentromeres form efficiently at multiple possible loci in *Candida albicans*. PLoS Genet 5:e1000400. https://doi.org/10.1371/journal.pgen.1000400.