Centromere mediated genome rearrangements and functional analysis of an outer kinetochore protein in human fungal pathogens

A thesis submitted for the degree of

Doctor of Philosophy

by

Sundar Ram Sankaranarayanan

To

Molecular Biology and Genetics Unit

Jawaharlal Nehru Center for Advanced Scientific Research

Jakkur, Bengaluru- 560064, India

May 2020

Dedicated to the loving memory of Ganesh

DECLARATION

I do hereby declare that this thesis entitled "**Centromere mediated genome rearrangements and functional analysis of an outer kinetochore protein in human fungal pathogens**" is an authentic record of research work carried out by me towards my Doctor of Philosophy under the guidance and supervision of Prof. Kaustuv Sanyal at the Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India and that this work has not been submitted elsewhere for the award of any other degree. In keeping with the norm of reporting scientific observations, due acknowledgments have been made whenever the work described was based on the findings of other investigators. Any omission, which might have occurred by oversight or misjudgment, is regretted.

Sundar Ram S

Place: Bangalore

Date: 15th May 2020

Jawaharlal Nehru Center for Advanced Scientific Research

Kaustuv Sanyal *PhD, FAAM (USA), FNA, FASc, FNASc*
Professor & Tata Innovation Fellow

Visiting Professor
Graduate School of Frontier Biosciences
Osaka University
Suita, Osaka 565 0871, Japan

15 May 2020

Certificate

This is to certify that the work described in this thesis "Centromere mediated genome rearrangements and functional analysis of an outer kinetochore protein in human fungal pathogens" is the result of the investigations carried out by **Mr. Sundar Ram Sankaranarayanan** towards his Doctor of Philosophy degree as part of the Integrated-PhD program at the Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore, India under my guidance. The results presented in this thesis have not previously formed the basis for the award of any other diploma, degree, or fellowship.

Kaustuv Sanyal, Ph.D
Professor
Molecular Mycology Laboratory
Molecular Biology and Genetics Unit
Jawaharlal Nehru Centre for
Advanced Scientific Research
Jakkur Post, Bangalore - 560 064, India.
Ph : +91-80-22082878, Email : sanyal@jncasr.ac.in

Place: Bangalore

Acknowledgments

This thesis is the outcome of constant support and encouragement of several people over the years of my stay at JNCASR. I take this opportunity to express my sincere gratitude to all of them and thank them for their unconditional support.

First and foremost, I would like to thank Prof. Kaustuv Sanyal, who has been an excellent mentor and helped me cross the bridge to come to this stage of submitting my thesis. Right from the start, he has involved me actively in several projects, helping me understand and appreciate diverse aspects of chromosome segregation in fungi. He always pushed an inch further so that one could do the very best of their ability. I would like to thank him for showing what difference can be made by that little extra effort and hope to better it in the coming years. His untiring enthusiasm in framing hypothesis with each result, and willingness to collaborate on aspects that the lab doesn't specialize on, has greatly improved many projects, including mine. He has been very understanding and supportive of many of the problems I faced, without which it would have been impossible to come to this stage. I really appreciate it and thank him for all his patience and guidance.

Before this course, I did my undergraduate studies in an engineering college where we hardly got an opportunity to work in anything close to a laboratory. The transition from such a background to working in a research lab would not have been possible without the mentorship of Dr. G.R. Ramesh. I spent my first two semesters with him, during which he introduced us to the world of experimental science. The basics of routine molecular biology techniques, troubleshooting a failed experiment, and his mantra for research life- 'plan, execute and analyze' are some of the things that have immensely helped during Ph.D. work.

I would like to thank Prof. Ranga Uday Kumar and his lab for supporting me during the first lab rotation. His passion for work and systematic nature is something every young researcher needs to imbibe. I honed my cloning skills during my stay in his lab, and I would like to thank Prabhu, Mahesh, and Rebu for their constant motivation and support.

Our coursework introduced me to the amazingly diverse topics in biology. In this aspect, I would like to thank the MBGU faculty, Prof. Uday Kumar Ranga, Prof. Tapas Kumar Kundu, Prof. Maneesha Inamdar, Prof. Hemalatha Balaram, Prof. Anuranjan Anand, Prof. Kaustuv Sanyal, Prof. Jayanta Haldar, Prof. Ravi Manjithaya, and the faculties coordinating the Scientific communication course.

I was introduced to *Malassezia* during our bi-weekly Skype call with Prof. Joe Heitman's group. What started off as a side project matured to be a chapter of my thesis. I am thankful to Prof. Heitman and

Prof. Tom Dawson for their continuous support and encouragement for working with *Malassezia*. Working with *Malassezia* was never easy given we faced difficulties even in growing this organism in the lab. Over time I could learn to handle *Malassezia* due to constant interactions with Giuseppe and Sheng, and I thank them for their time and support. Marco's contribution to the visualization of our results is immense. His crystal-clear illustrations helped a great deal during publication of this part of work. I would also like to thank Christian and Prof. Tom Dawson for providing us with the genome and RNA sequence data. I thank Prof, Rahul Siddharthan, and his student Rakesh for performing motif scans across several *Malassezia* genomes. Their findings gave us confidence and support for our centromere predictions in these species. I would also like to thank Bhagya and Promit for helping me with synteny analysis, ChIP-seq, and RNA-seq data, and for teaching me the basics of genome analysis. I am very grateful to Hashim for his constant help, support, and encouragement, without which I would not have completed the *Malassezia* study. His meticulous and systematic approach towards addressing a problem is something I look up to.

The Dad2 part of the thesis is also a collaborative effort of many individuals. I would like to thank Arti for laying the foundation for this work. Many of her results, both positive and negative, helped me design the course of this study. I also acknowledge Prof. K.V.R. Chary's lab for their efforts on deducing the structure of Dad2 through NMR. Although not a part of this thesis work, insights from their work helped us shape the course of this study. I am thankful to Prof. Tapas Manna's Lab at IISER-Tvm and his students Renjith and Sujaya for their help with the microtubule cosedimentation assay which helped us establish that Dad2 can bind to microtubules.

I would like to thank Prof. Wolfe Heyer, Prof. Scott Keeney, and Prof. Jeyaprakash Arulanandam for their constructive criticism during discussions about this work during the Chromosome stability meeting. I also take this opportunity to thank Prof. Ganesh Nagaraju and Prof. Deepak Saini for their suggestions during comprehensive exam presentation.

I thank the JNCASR administration for the fellowship and hostel facilities. I would like to thank the technical staff- Dr. Prakash at the animal house, the Clevergene team at the sequencing facility, and Ms. Suma and Mr. Sunil at the microscopy facility for helping with specific experiments.

I thank all the members of MML for making me feel home right from the beginning. The senior trio when I joined the lab- Sreyoshi, Laxmi, and Gautam were fantastic teachers. I especially thank Gautam for all his efforts to teach me the basics of microbiology, working with *Candida*, and the countless discussions on music. I thank the past members of the lab Asif, Abhishek, Tanmoy, Uttara, Lakshmi, Vikas, Jigyasa, Parijat, Bornika, and Aditi for the cheerful time in the lab. I am equally thankful to the present members Hashim, Shweta, Aswathy, Krishnendu, Shreyas, both the Priyas, Satyadev,

Kuldeep, Rashi, and Padma for all the madness in the lab and keeping it lively! Neha and Rima will always remain special to me for all the wonderful time we spent over these years discussing anything under the sun, and for their constant help, support, and encouragement. Also, I am extremely thankful to Mr. Nagaraj, the backbone of MML, for his excellent management of lab resources.

Sports, music, and coffee have been my way of rejuvenating from stress. I thank the table tennis group – Bharath, Sanjeev, Lakshmeesha, Asutosh, Aditya, and Arun for introducing me to this beautiful game. I am especially thankful to Mr. Sharanappa for that perfect cup of coffee at any time of the day.

I take this moment to my teachers Dr. Antony Caesar and Dr. Naina Baburaj, who influenced and encouraged me to pursue this career. Their teachings, encouragement, and blessings have been the backbone of my academic dreams.

Lastly, I would like to thank my family for their unconditional support during my tenure here. No words can match the sacrifices they have made to make me realize my dreams. I am grateful to my in-laws and my guardian angels- Amruthaa, Nandi, and Divya, for their unwavering support.

~Sundar Ram S

Table of contents

List of figures

List of tables

Abbreviations

CR	- Chromosome rearrangement
WGD	- Whole genome duplication
TARE	- Telomere associated repeat elements
FISH	- Fluorescence *in situ* hybridization
PFGE	- Pulsed field gel electrophoresis
RITS	- RNAi induced transcriptional silencing
CLRC	- Clr4 methyltransferase complex
HDAC	- Histone deacetylase
CENP	- Centromere protein
CCAN	- Constitutive centromere associated network
KMN	- Knl1C-Mis12C-Ndc80C
SAC	- Spindle assembly checkpoint
MCC	- Mitotic checkpoint complex
APC/C	- Anaphase promoting complex/ cyclosome
PV	- Pityriasis versicolor
SD/D	- Seborrheic dermatitis/ Dandruff
AD	- Atopic dermatitis
HTG	- Horizontally transferred gene
SMRT	- Single molecule real-time (Pacific Biosciences)
MRS	- Major repeat sequences
MTL	- Mating type like
PWM	- Position weight matrix
DSB	- Double-strand break
DSS	- Dad2 signature sequence
BLAST	- Basic local alignment tool
ARS	- Autonomously replicating sequence
GFP	- Green fluorescent protein
DMSO	- Dimethyl sulfoxide
DNA	- Deoxyribonucleic acid
RNA	- Ribonucleic acid
MYA	- Million years ago

NAT	- Nourseothricin
ORF	- Open reading frame
PAGE	- Polyacrylamide gel electrophoresis
PBS	- Phosphate buffer saline
RPM	- Revolutions per minute
SPB	- Spindle pole body
ChIP	- Chromatin immunoprecipitation
K_d	- Dissociation constant
PCR	- Polymerase chain reaction
qPCR	- Quantitative PCR
SD	- Standard deviation
SEM	- Standard error of mean
ANOVA	- Analysis of variance
ºC	- Degree Celsius
µg	- microgram
mg	- milligram
g	- gram
µL	- microlitre
mL	- millilitre
L	- litre
µM	- micromolar
mM	- millimolar
M	- molar
N	- normal
µm	- micrometer
bp	- base pair
kb	- kilobase
Mb	- megabase
s	- second
min	- minute
h	- hour

Chapter 1

Introduction

Genome evolution and adaptation

Comparison of genomes of closely related species or different isolates within a species show variations at the DNA sequence level or in their structural organization in a chromosome (BARRACLOUGH *et al.* 2003). The generation of genetic diversity is imperative for an organism to adapt and survive in diverse environmental conditions (EICHLER AND SANKOFF 2003). The outcome of these rearrangements could be beneficial or detrimental to the cell. Based on the nature of this outcome, they get selected for, or lost in the population by natural selection. Meiotic recombination is one of the well-known drivers of genome evolution mediated by recombining genetic material from both parents (HOLLIDAY 1984). However, several asexual/ mitotic mechanisms can also result in such rearrangements. Signs of polymorphisms or rearrangements in the genomes of species with rare or no observed sexual cycle suggests that alternate mechanisms mediating chromosomal rearrangements (CRs) exist (SEIDL AND THOMMA 2014). Generating genetic diversity through such modes is considered to assist asexual organisms in adapting and growing in different niches.

Modes of genome evolution

Duplications/ gene amplification

Over 50 years ago, Ohno proposed that the duplication of genetic material as an easy and effective way of generating large scale genetic diversity rather than creating genes *de novo* (OHNO 1970). Duplication of genetic material, the most explored drivers of genome evolution (EICHLER 2001), is classified into two types based on the scale of duplication- whole-genome duplication (WGD) and segmental duplication.

WGD has been reported independently in several eukaryotic lineages, from fungi up to vertebrates (WOLFE AND SHIELDS 1997; MCLYSAGHT *et al.* 2002). As the name implies, the process involves duplication of the entire genetic material, after which further rearrangements and loss of redundant gene copies stabilize the genome. In this process, the genes or gene families that provide any fitness advantage enabling the extant species to adapt to specific environmental conditions are retained (OTTO AND WHITTON 2000). In some cases, the duplicated genes can also diverge to acquire novel function such that they can provide an adaptive edge to the organism- a process termed as neofunctionalization (LYNCH AND CONERY 2000). Among fungi, ancient WGD events are proposed to have shaped the genomes of the basal fungi like *Rhizopus*, *Mucor*, and *Phycomyces* species (MA *et al.* 2009; CORROCHANO *et al.* 2016). The expanded gene families that were retained include those involved in signal transduction, cell wall biogenesis, secreted proteases. Evidence for the specialization of genes related to phototropism was also identified in these species. Among the Ascomycota, WGD was

evidenced in the Saccharomycetes (WOLFE AND SHIELDS 1997; KELLIS *et al.* 2004). Analysis of the duplicated gene families suggested that the increase in copy number of glycolytic genes conferred a selective advantage for growth in glucose-rich environments to all the post-WGD species through rapid glucose fermentation (CONANT AND WOLFE 2007). At least a single round of WGD in early chordate evolution resulting in diversification/ amplification of signaling pathways has been proposed in the vertebrates (ABI-RACHED *et al.* 2002; GU *et al.* 2002; MCLYSAGHT *et al.* 2002). The extant *Arabidopsis* genome is also a consequence of a WGD event during the evolution from *Brassicaceae* species (VISION *et al.* 2000; BOWERS *et al.* 2003).

Segmental duplications involve duplications of shorter segments of the genome that are either juxtaposed such that they form tandem repeats or get interspersed in the genome by subsequent non-allelic recombination (Figure 1-1). Rearrangements between such duplicated segments resulting in rapidly evolving stretches have been identified in human genomes, wherein 5% of the genome consists of such duplications (BAILEY *et al.* 2002). Closely related *Caenorhabditis* species *C. elegans* and *C. briggsae* are separated by 14 segmental duplications since their divergence from a common ancestor (COGHLAN AND WOLFE 2002). Insecticide resistance in *Anopheles gambiae* was ascribed to duplication of the loci harboring genes that include cytochrome P-450, glutathione transferases, and carboxylesterases (RANSON et al. 2002). Another form of segmental duplication observed in pathogenic fungus *Candida albicans* is the formation of isochromosome in response to azoles. Isochromosome formation involving the left arm of Chr5, resulting in an increased copy number of *ERG11* and *TAC1*, was observed upon treatment with azoles (SELMECKI *et al.* 2006; SELMECKI *et al.* 2008). Similarly, the trisomy of Chr4 was also implicated in resistance to fluconazole (ANDERSON *et al.* 2017). In another human fungal pathogen, *Cryptococcus neoformans*, the disomy of Chr1 is associated with fluconazole resistance (SIONOV *et al.* 2010).

Amplification of gene families is another mechanism that helps species adapt to an ecological niche. For instance, the amplification of gene families such as glycosyl hydrolases, peptidases, and transporters are the signatures of pathogenic lifestyle in many plant-pathogenic fungal and oomycete genomes (TYLER *et al.* 2006; SPANU *et al.* 2010; KLOSTERMAN *et al.* 2011; SEIDL *et al.* 2012). Members of the *Malassezia* species complex, a lipophilic group of yeasts show amplification of genes essential for assimilating lipids and fatty acids, enabling their survival on animal skin (XU *et al.* 2007; WU *et al.* 2015). Gene amplification as an adaptive response to external stress conditions is also known in fungi. The genes responsible for sugar assimilation such as *HXT6*, *HXT7* are amplified upon nutrient starvation in *Saccharomyces cerevisiae* (BROWN et al. 1998; GRESHAM et al. 2008). Similar adaptive means for assimilation of copper and arsenite have been reported in *S. cerevisiae* and *C. neoformans* respectively (CHOW *et al.* 2012; HULL *et al.* 2017; ZHAO *et al.* 2017).

Recombination through repetitive sequences- deletions, inversions, and translocations

Most eukaryotic genomes are characterized by the presence of various forms of repetitive elements that are substrates for recombination events. The outcome of these recombination events are diverse and includes deletions, inversions, translocations based on the orientation of repeat units (reviewed in (FIERRO AND MARTIN 1999; GUSA AND JINKS-ROBERTSON 2019)). Commonly observed repeat elements in eukaryotic genomes include long-interspersed nucleotide elements (LINEs), short-interspersed nucleotide elements (SINEs), DNA- and retro-transposons, rDNA repeats, sub-telomeric repeats, and other species-specific repeat elements (BISCOTTI *et al.* 2015).

The presence of direct repeats is one of the primary sources of deletions in the genome. When present in the same chromosome, recombination between direct repeats can result in the excision of the intervening DNA sequence (Figure 1-1). When present in nonhomologous chromosomes, such repeats can facilitate reciprocal translocation of sequences through these repeats (Figure 1-1). Such translocations have been observed in various organisms studied to date. Breakpoints of conserved gene order (also known as gene synteny) between human and mouse chromosomes mapped to L1 repeats and LTR retrotransposon elements indicative of translocation at these loci (DEHAL *et al.* 2001). The translocations observed between closely related worm species *C. elegans* and *C. briggsae* coincided with the dispersed repeats present in these genomes (COGHLAN AND WOLFE 2002). Similar rearrangements have been reported in fungi as well. For instance, different types of repeats have been implicated in the rearrangements observed between *S. cerevisiae* and its related species *Saccharomyces paradoxus*, *Saccharomyces mikatae*, and *Saccharomyces bayanus* (KELLIS et al. 2004). Reciprocal translocations were mapped to the Ty class transposons and the highly similar rRNA genes present in the genome. Rearrangements mediated by transposable elements in the *Magnaporthe* genome accounted for the loss of conserved gene synteny observed during the evolution in Sordariomycetes that includes filamentous fungi such as *Magnaporthe, Neurospora, and Fusarium* species (THON *et al.* 2006). The genome of *C. albicans* contains unique repeat elements called as the major repeat sequences (MRS) that are present in almost all the chromosomes. Rearrangements mediated by the MRS account for the majority of karyotype diversity observed among *C. albicans* isolates and between *Candida dubliniensis* (CHIBANA *et al.* 1994; CHINDAMPORN *et al.* 1998; LEPHART AND MAGEE 2006).

Recombination between repeat blocks that are present in inverted orientation can result in inversion of the intervening sequences (Figure 1-1). Besides the translocations described above between *S. cerevisiae* and related species, several inversions were detected upon the comparison of these genomes. These inversion breakpoints were mapped to highly similar tRNA genes present in an inverted orientation in the genome (KELLIS *et al.* 2004).

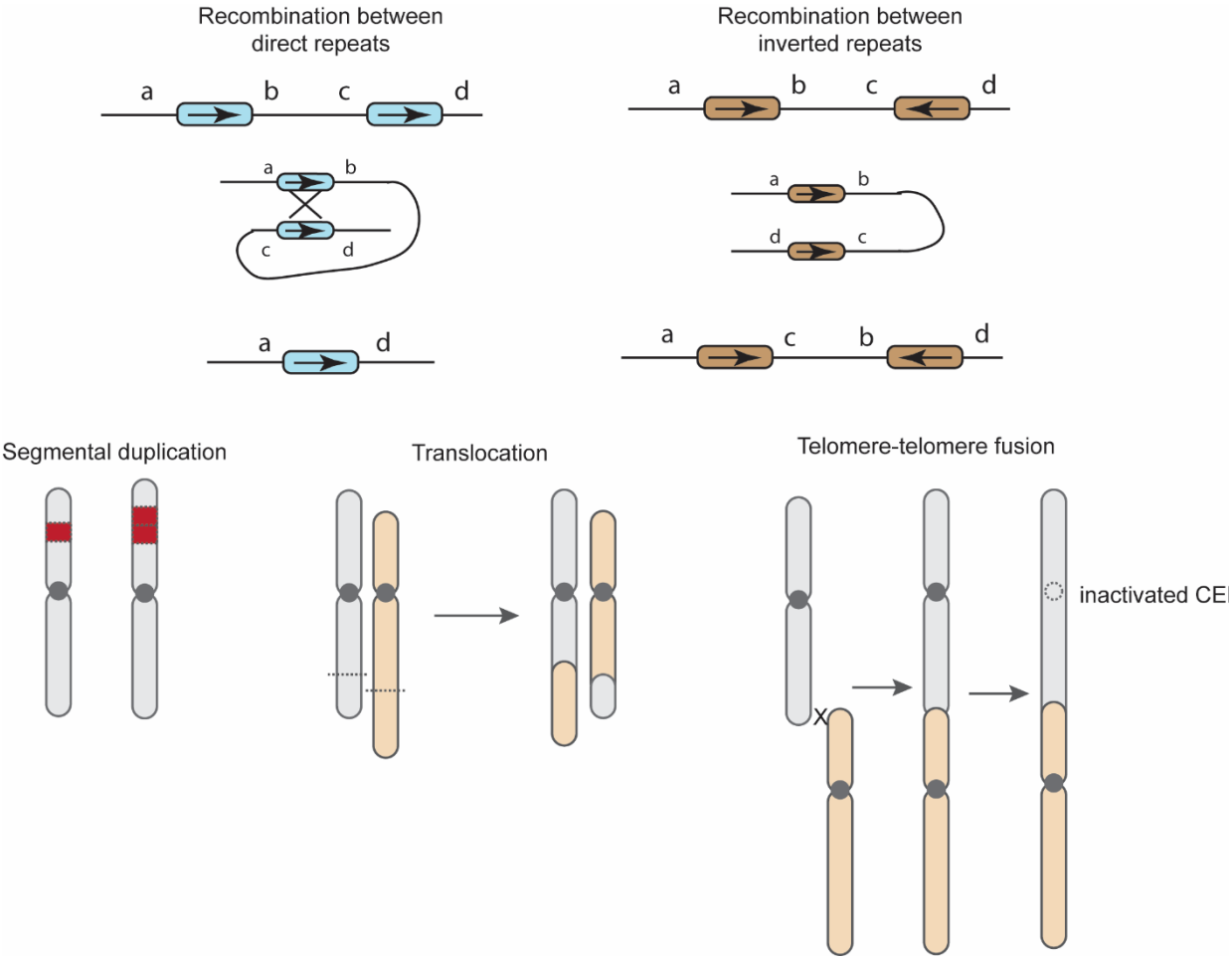


Figure 1-1. Different modes of eukaryotic genome evolution.
Schematic representation of the recombination between direct and inverted repeats in the genome resulting in the deletion and inversion of the intervening DNA sequences respectively is present in the top panel. Segmental duplication in a chromosome, interchromosomal recombination resulting in translocation of chromosome fragments, and chromosome number reduction by telomere-telomere fusion of chromosomes are schematically depicted in the bottom panel.

The comparison of gene order between *S. cerevisiae* and *C. albicans* revealed over 1100 single gene inversions between them since their divergence (SEOIGHE *et al.* 2000). Similarly, over 70 inversions were detected when the genomes of *A. gambiae* and *Anopheles funestus*, two closely related malaria vectors that diverge about five million years ago, were compared (SHARAKHOV *et al.* 2002). While these inversions are fixed in the population, the fitness advantage and the cause of such inversions remain elusive.

The subtelomeric region is another known recombination hotspot that facilitates translocations between nonhomologous chromosomes through the repeats present in these regions. In *S. cerevisiae*, the recombination between subtelomeric regions is proposed as the mechanism leading to the amplification of the *SUC* gene family that is related to the assimilation of sucrose (LOUIS *et al.* 1994; NESS AND AIGLE 1995). Comparative analysis of *Plasmodium* genomes revealed significant homogenization at the sub-telomeric regions indicative of the exchange of genetic material at these loci. These telomere-associated repeat elements (TAREs), along with rRNA repeats, are the primary drivers of inter-chromosomal exchanges in these species (CARLTON *et al.* 2002; GARDNER *et al.* 2002). Recombination between subtelomeric repeats and a homologous sequence present in the chromosome arm was shown to result in terminal chromosome translocations and terminal deletions in *Aspergillus nidulans* and *Neurospora crassa* respectively (NEWMEYER AND GALEAZZI 1977; SEXTON AND ROPER 1984).

Variation in chromosome number

Apart from the rearrangements mentioned above, closely related species from several eukaryotic lineages have also been found to contain different numbers of chromosomes. Indeed, observations on differences in chromosome numbers were one of the earliest data available to support genetic variation between species. The well-known mechanisms that were shown to result in chromosome number variations in different organisms are discussed below.

Telomere-telomere fusion is a common mode of chromosome number reduction that has been observed in several systems (Figure 1-1). Examples of this type include the formation of extant human chromosome 2 by fusion of two chromosomes in a common ancestor shared with the great apes (JDO *et al.* 1991), karyotype reduction identified in the post-WGD lineage of Saccharomycotina that includes *Candida glabrata, Vanderwaltozyma polyspora, Kluyveromyces lactis,* and *Zygosaccharomyces rouxii* (GORDON et al. 2011), Robertsonian translocation/fusion of acrocentric chromosomes resulting in the formation of a metacentric chromosome in *Arabidopsis* and related *Brassicaceae* species, and in several mouse races (HAUFFE AND PIÁLEK 1997; LYSAK *et al.* 2006).

Nested chromosome insertions (NCIs), wherein an entire donor chromosome is ‘inserted’ into or near the centromere of a nonhomologous recipient chromosome, is another mode of chromosome number reduction most commonly observed in grasses (LUO *et al.* 2009; MURAT *et al.* 2010) and sporadically in other lineages like cucumber and other *Brassicaceae* genera (LYSAK *et al.* 2006; YANG *et al.* 2014). Similar to telomere fusions, NCIs also result in the formation of dicentric chromosomes when the insertion occurs outside the centromere of the recipient chromosome.

Polyploidy, a consequence of WGD, is another common mechanism of chromosome number evolution in plants. Independent WGD events- one in the common ancestor of all seed plants and another

in the common ancestor of angiosperms were proposed to have driven the evolution of these species (JIAO *et al.* 2011). Many extant plant genomes like those of cereals and grapevines are polyploidy in their genetic content. In animals, incidences of inter-species hybridization resulting in duplicated gene content, also known as allopolyploidy, exist in diverse eukaryotic lineages. Known examples of allopolyploids include the *Xenopus laevis* (amphibians; (SESSION *et al.* 2016)), *Carassius auratus* (fishes; (CHEN *et al.* 2019)), and *Ambystoma* salamanders (reptiles; (MCELROY *et al.* 2017)).

In the case of holocentric chromosomes, many chromosomal regions can associate with the microtubules, resulting in faithful inheritance during cell division (DERNBURG 2001). Consistent with this property, runaway chromosome number expansions by chromosome fission are predominantly observed in organisms that have holocentric chromosomes. Most extreme examples of such chromosome number variations are observed in insects. In the case of *Apiomorpha* (Hemiptera), haploid chromosomal complement ranging from n=2 to 96 was reported (COOK 2000). Among lepidopterans, examples of such variations include the species of *Godyris* (n=13-120; (BROWN *et al.* 2004)), *Leptidea* (n=28-103; (SICHOVA *et al.* 2015)), and *Lysandra* (n=24-93; (TALAVERA *et al.* 2013)). The *Agrodiaetus* butterflies show karyotype variations ranging from n=10-134 (VERSHININA AND LUKHTANOV 2017). Outside of the insect lineages, similar expansion in the karyotype variation were reported in Scorpions (MATTOS *et al.* 2018; UBINSKI *et al.* 2018), and in angiospermous plant genus *Carex* (Sedges, (HIPP 2007))

Studying chromosome number variations

In the 1970s, karyotypic variation in diverse organisms such as those among species of bats, muntjacs, new-world monkeys, and fishes was well established (EGOZCUE 1969; FREDGA 1977; BAKER AND BICKHAM 1980). Observations on variations in the number and structure of chromosomes were one of the earliest known markers for genetic variation. Such karyotypic changes and other forms of CRs discussed earlier are potential drivers of speciation. Such changes can obstruct the pairing of homologous sequences during meiosis, which upon recombination can result in unbalanced progeny with reduced fitness, thereby preventing gene flow between them (WHITE 1969; WHITE 1978). CRs can also create regions of suppressed recombination during meiosis, accumulating specialization genes that eventually drive populations to speciation (NOOR *et al.* 2001; RIESEBERG 2001). In organisms lacking a conventional sexual cycle, CRs can assist in overcoming the limitations of lack of sex in generating diversity to adapt to a given niche or stress (EICHLER AND SANKOFF 2003; SEIDL AND THOMMA 2014; GUSA AND JINKS-ROBERTSON 2019).

Differences in chromosome number and structure were primarily estimated by G-, C-banding techniques, fluorescence in situ hybridization (FISH), and by pulsed-field gel electrophoresis (PFGE) based karyotyping in case of smaller genomes like those of fungi. While these methods reveal the

differences in chromosome number and size, additional resources like centromere identity, genome sequences are required for an understanding of the mechanisms that resulted in the karyotypic change.

The centromere is an essential genetic and structural component of every eukaryotic chromosome. Any evolutionary change in chromosome number should be accompanied by a concomitant gain/ loss of centromeres (DARLINGTON 1937). Also, chromosome number changes usually occur when recombination results in a fusion product with two centromeres, of which one centromere gets inactivated, as in the case of telomere-fusions or NCIs described earlier. Thus, knowledge on centromere identity of an organism and the factors governing centromere function is critical for understanding karyotype evolution. When used in conjunction with a well-assembled genome sequence information, one can identify and trace the synteny blocks between rearranged genomes that resulted in the karyotypic variation. Another important aspect of studying karyotype variation is to understand the path of evolution, the ancestral state from which the variation ensued. Estimates of chromosome numbers from several closely related organisms along with their phylogenetic relationship can help one identify the ancestral state, critical to study the direction of evolution.

Centromeres

Centromeres, identified as the primary constriction on chromosome spreads, are unique locus on each chromosome that physically associates with the microtubules to drive chromosome segregation (FLEMMING 1882). The distinct feature of this locus, when compared to the bulk chromatin, is the presence of a variant histone H3 called CENP-A that is incorporated into the nucleosomes instead of canonical histone H3 (PALMER *et al.* 1991; SULLIVAN *et al.* 1994). Indeed, CENP-A is called as the epigenetic mark of centromeres. A remarkable diversity in both sequence and the organization of these sequence elements has been observed at the centromeres to perform an essential, conserved function like chromosome segregation (MCKINLEY AND CHEESEMAN 2016).

Diversity in centromere structure

Chromosomes are either monocentric or holocentric based on the attachment of microtubules to a single locus on it or across the length of the chromosome (STEINER AND HENIKOFF 2015). Among the monocentric chromosomes, different configurations have been identified based on the size and the nature of sequences present in the centromere (Figure 1-2).

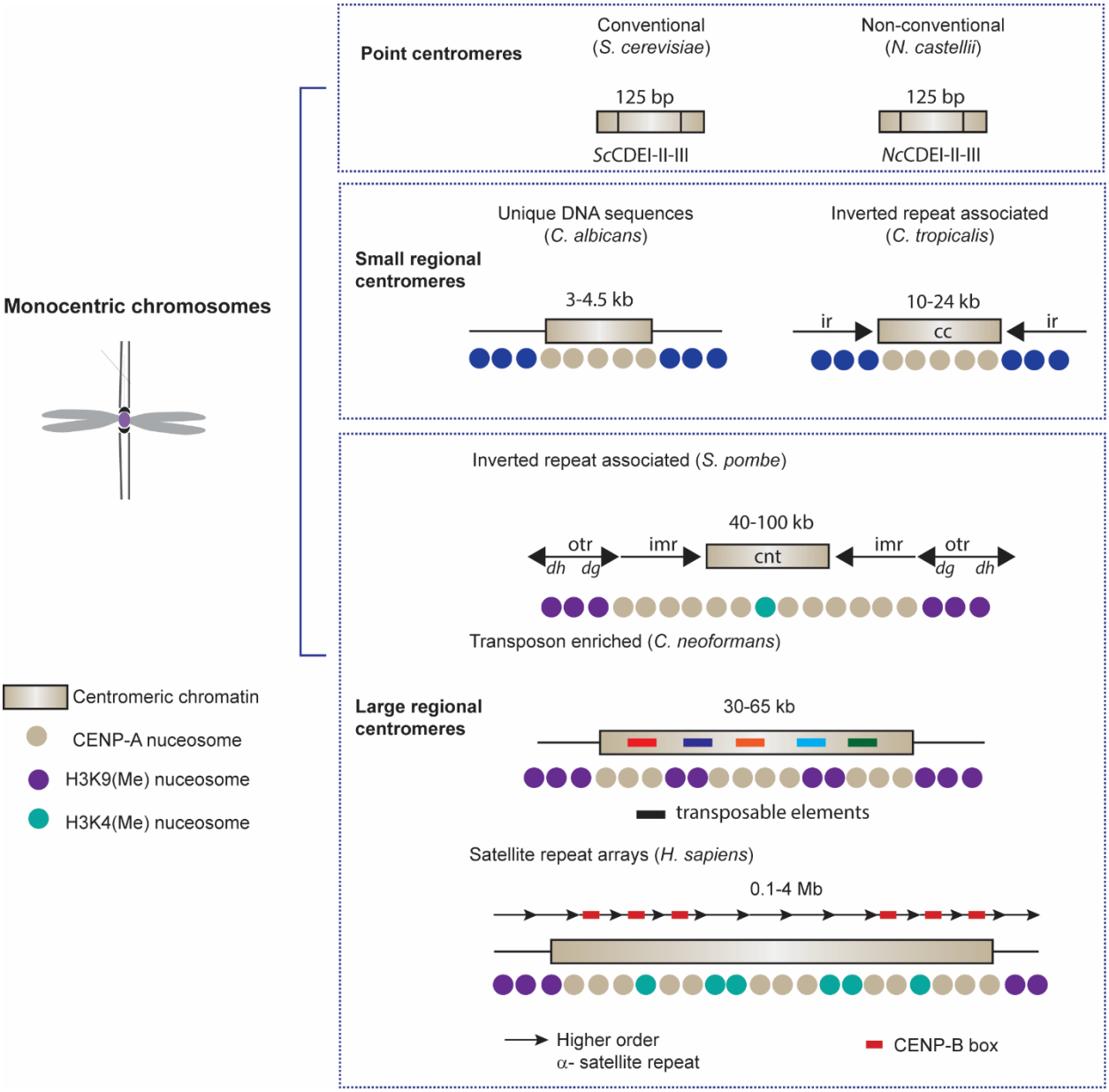


Figure 1-2. Structural diversity observed in centromeres from diverse organisms.
Monocentric chromosomes are classified into point, small regional, and large region centromeres based on length of centromere DNA. A representative of each of these centromere types is schematically represented with their size, constituent nucleosomes and other sequence features.

In fungi

Much of the structural diversity in centromeres are known among the ascomycetes of fungi (ROY AND SANYAL 2011). The centromeres of budding yeast *S. cerevisiae* was the first centromere to be identified and cloned. These centromeres spanning <200 bp in length are called as point centromeres and

have been identified in related budding yeasts of the family *Saccharomycetaceae* (CLARKE AND CARBON 1980; GORDON *et al.* 2011; KOBAYASHI *et al.* 2015). These centromeres are organized into conserved DNA elements I, II, and III (CDEI- III) that span 125 bp in length and are recognized by a cognate kinetochore protein complex called the CBF3 complex, making them genetically defined centromeres. Small regional centromeres, identified in several *Candida* species, form the second category (SANYAL *et al.* 2004; PADMANABHAN *et al.* 2008; KAPOOR *et al.* 2015; CHATTERJEE *et al.* 2016) and have a 2 to 5 kb region enriched by kinetochore proteins. These centromeres can either have unique DNA sequences as observed in *C. albicans*, *C. dubliniensis* and *Candida lusitaniae* or have a homogenized core that is flanked by inverted repeats as reported in *Candida tropicalis* and *Komagataella phaffii*. The third type of centromere structure is the large regional centromere, often repetitive in sequence and spanning longer than 15 kb. Large regional centromeres can be transposon enriched as in *N. crassa*, and *Magnaporthe oryzae* or organized into repeat structures around a central core as in *Schizosaccharomyces pombe* and related species (CHIKASHIGE *et al.* 1989; CLARKE AND BAUM 1990; SUN *et al.* 2017; YADAV *et al.* 2018b).

In the Basidiomycota, the centromeres of *Cryptococcus* species are of the large regional type enriched with transposon sequences reminiscent of those observed in *M. oryzae* and *N. crassa* (SUN *et al.* 2017; YADAV *et al.* 2018b). Centromeres of similar features were predicted in the *Ustilago* species complex that is closely related to *Cryptococcus* species (YADAV *et al.* 2018b). From the whole genome assembly, clusters of retrotransposons were proposed to be the centromeres of *Coprinopsis cinerea* (STAJICH *et al.* 2010).

The centromeres of an early divergent, basal fungi *Mucor circinelloides* of phylum Mucoromycota, were found to harbor a short kinetochore bound region flanked by regions enriched with retroelements (NAVARRO-MENDOZA *et al.* 2019). This organization was described as a ‘mosaic’ centromere bearing features of both point and regional centromeres.

In animals and plants

Other than fungi, the centromeres from species representative of other kingdoms such as fruit fly, chicken, primates, humans, and different phyla of plants are also known. In all of the cases, the centromeres are large-regional type, consisting of species-specific repetitive elements or transposons (CHOO 1997; PLOHL *et al.* 2014). Examples include the primate centromeres built over arrays of 171 bp monomer alpha-satellite DNA (RUDD *et al.* 2003; ALKAN *et al.* 2007), *Mus musculus* centromeres with a homogenized core containing 120 bp minor satellite repeats flanked by major satellite repeats (JOSEPH *et al.* 1989), and the centromeres in *Arabidopsis* and rice with 178 bp and 155 bp repeat monomers (NAGAKI *et al.* 2003; LEE *et al.* 2005).

Holocentromeres

Holocentricity has been proposed to have independently evolved in various taxa at least 13 times (Melters et al. 2012). Of the known holocentromeres, *C. elegans* is relatively well studied, and the centromeres are considered as 'dispersed point centromeres' given the occurrence of CenH3 nucleosomes at accessible regions such as transcription factor hotspots (Steiner and Henikoff 2014). In specific insect lineages, the transition to holocentricity is associated with loss of CENP-A and associated kinetochore proteins that specify the site for kinetochore assembly (Drinnenberg et al. 2014). Notable exceptions to this include the case in Trypanosomes (Akiyoshi and Gull 2014; Garcia-Silva et al. 2017) and *M. circinelloides* (Navarro-Mendoza et al. 2019), which lost CENP-A but retained monocentric chromosomes.

Determinants of centromere identity and function

DNA-sequence elements

Point centromeres of the budding yeast is a classical case where DNA sequence governs centromere function (CLARKE AND CARBON 1980). The contribution of DNA sequence elements in centromere function was elegantly demonstrated by constructing stable minichromosomes in which the centromeres were cloned. A replicative plasmid that contains sequences for initiation of replication (ARS) were not stably inherited in a population of cells. However, the addition of a centromere sequence conferred mitotic stability to these plasmids. This suggested that the cloned centromere sequence enabled functional kinetochores assembly that facilitated the proper segregation of these plasmids. This also served as a powerful tool to dissect the critical domains within the ~125 bp centromere, such as the CCG motif present in CDEIII. A single base substitution in this motif to CTG abolished centromere activity in the plasmids or when introduced in the chromosome (MCGREW *et al.* 1986; NG AND CARBON 1987). The CBF3 complex was identified as a centromere associated protein complex that binds specifically through CDEI and CDEIII (LECHNER AND CARBON 1991). This provided a direct link between the centromere sequence and its function (Figure 1-3A) (CAI AND DAVIS 1989; CAI AND DAVIS 1990; JEHN *et al.* 1991; LECHNER AND CARBON 1991). The proteins Ndc10, Cep3, and Ctf13 of the CBF3 complex were found restricted to organisms containing point centromeres (MERALDI *et al.* 2006). It should be noted that the deposition of CENP-A^{Cse4} to these centromeres is achieved with the assistance of Ndc10, forming the genetic basis of centromere function (CAMAHORT *et al.* 2007; STOLER *et al.* 2007).

Similar minichromosome based assays were used to dissect the functionally critical domains within the *S. pombe* centromeres that are of the large regional type (HAHNENBERGER *et al.* 1989; CLARKE AND BAUM 1990; MATSUMOTO *et al.* 1990; BAUM *et al.* 1994). As mentioned earlier, the centromeres in

this species contain a central core region (*cnt*) that is flanked by inverted repeat structures called the innermost repeats (*imr*, also called B repeats). This central region (*cnt* and *imr* repeats) is flanked by outer repeats (*otr*) that contain *dg* and *dh* elements (also called K and L repeats) (CHIKASHIGE *et al.* 1989; CLARKE AND BAUM 1990). Testing the requirement of each of these elements using a *CEN1* or *CEN2* containing minichromosome suggested that the central region with the K- type repeat was sufficient for mitotic stability (HAHNENBERGER *et al.* 1989; CLARKE AND BAUM 1990; BAUM *et al.* 1994). However, the otr corresponding to the pericentric heterochromatin was essential for proper segregation in meiosis I (HAHNENBERGER *et al.* 1989). Unlike the case in *S. cerevisiae*, no cognate kinetochore complex that recruits CENP-A to the centromeres are present in *S. pombe*. However, the DNA sequences facilitate centromere function by alternate mechanisms described in the next section.

The short regional centromeres in *C. albicans* were one of the first fungal centromeres that were found to have unique DNA sequences in each of its chromosomes (SANYAL *et al.* 2004). Unlike the cases described above, cloned centromere DNA sequences failed to recruit CENP-A to minichromosomes when exogenously introduced into *C. albicans*. Further, a chromosome was truncated *in vivo* to generate an 85 kb mitotically stable chromosome fragment in this diploid organism. This fragment with a functional centromere, when isolated and reintroduced into *C. albicans* failed to recruit CENP-A, rendering the chromosome fragment unstable (BAUM *et al.* 2006). This suggested that the centromeres are specified by an unknown epigenetic means in this organism. Interestingly, the centromeres of another closely related *Candida* species *C. tropicalis* could impart stability to an ARS plasmid (CHATTERJEE *et al.* 2016). The implications of this divergent behavior of centromere sequences is discussed later in the context of centromere evolution.

DNA sequence directed binding of centromere proteins have been shown in repetitive centromeres of humans as well. An example of such sequence-specific binding is CENP-B that recognizes a specific repeat element called the CENP-B box, a 17-mer sequence present on the higher-order repeats of alpha-satellite DNA in primate and rodent centromeres (MASUMOTO *et al.* 1989; MURO *et al.* 1992; HAAF *et al.* 1995; KIPLING *et al.* 1995). The alpha satellite DNA sequence was shown to recruit CENP-A *de novo* in a CENP-B dependent manner, enabling stable inheritance of human and yeast artificial minichromosomes in human cell lines (HARRINGTON *et al.* 1997; IKENO *et al.* 1998; OHZEKI *et al.* 2002).

Epigenetic factors

While the genetic basis of centromere function represents a straightforward mechanism to perform this conserved function, several lines of evidence described below highlight the DNA sequence-independent epigenetic regulation of centromere function (MCKINLEY AND CHEESEMAN 2016).

Experimental or spontaneous deletion of a centromere has been shown to form centromere on another locus (neocentromere) in the chromosome with sequences unrelated to the deleted centromere (VOULLAIRE *et al.* 1993; SHANG *et al.* 2013; THAKUR AND SANYAL 2013). In cases like Chr17 in humans, two alpha satellite arrays, each capable of stabilizing a human artificial chromosome are present in tandem. However, CENP-A was found associated with only one of the two blocks, establishing an epiallele like behavior (MALONEY *et al.* 2012). Studies on primate centromeres reveal that a centromere can reposition itself in short evolutionary time scales to a new locus with less/no sequence similarity, forming evolutionarily new centromeres (TOLOMEO *et al.* 2017; SCHUBERT 2018). These observations allude towards the fact that the centromere DNA sequence is neither sufficient nor essential for centromere function.

However, the unifying feature of all functional centromeres is the presence of CENP-A (EARNSHAW AND MIGEON 1985; STOLER *et al.* 1995; WARBURTON *et al.* 1997; THAKUR AND SANYAL 2013). Across taxa, CENP-A is essential for a functional kinetochore (HOWMAN *et al.* 2000; MOORE AND ROTH 2001; OEGEMA *et al.* 2001; LIU *et al.* 2006; THAKUR AND SANYAL 2012; FACHINETTI *et al.* 2013; LOGSDON *et al.* 2015), and tethering of CENP-A or its assembly factors to non-centromeric loci has been shown to impart centromere activity in different systems (VAN HOOSER *et al.* 2001; BARNHART *et al.* 2011; MENDIBURO *et al.* 2011). In the following section, the epigenetic factors that regulate CENP-A dynamics at the centromeres and their effects on centromere propagation and centromere function are discussed.

Transcription and RNAi

Studies over the last two decades on several model organisms have uncovered a significant role of transcription in regulating the identity and function of centromeres, conventionally thought to be a heterochromatic region in the genome. A classic case highlighting the role of transcription on centromere function is from the pericentric repeats of *S. pombe* (Figure 1-3B) (HALL *et al.* 2002; VOLPE *et al.* 2002; VOLPE *et al.* 2003; FOLCO *et al.* 2008). The RNAPII generated transcripts from the *dg* and *dh* repeats of these pericentric regions are processed by dicer into siRNAs, which then attracts the argonaute containing RITS complex into these pericentric regions (REINHART AND BARTEL 2002; MOTAMEDI *et al.* 2004; VERDEL *et al.* 2004). The RITS complex then recruits the CLRC complex that contains Clr4 (the methyltransferase for H3K9) and other chromodomain containing proteins like Swi6/HP1, the SHREK complex, and the HDAC Clr3, which facilitate the formation of repressive chromatin (LACHNER *et al.* 2001; ZHANG *et al.* 2008). The interaction between Dcr1 and RDRC amplifies this signaling cascade by

further production of siRNAs from the pericentric regions (SUGIYAMA *et al.* 2005).

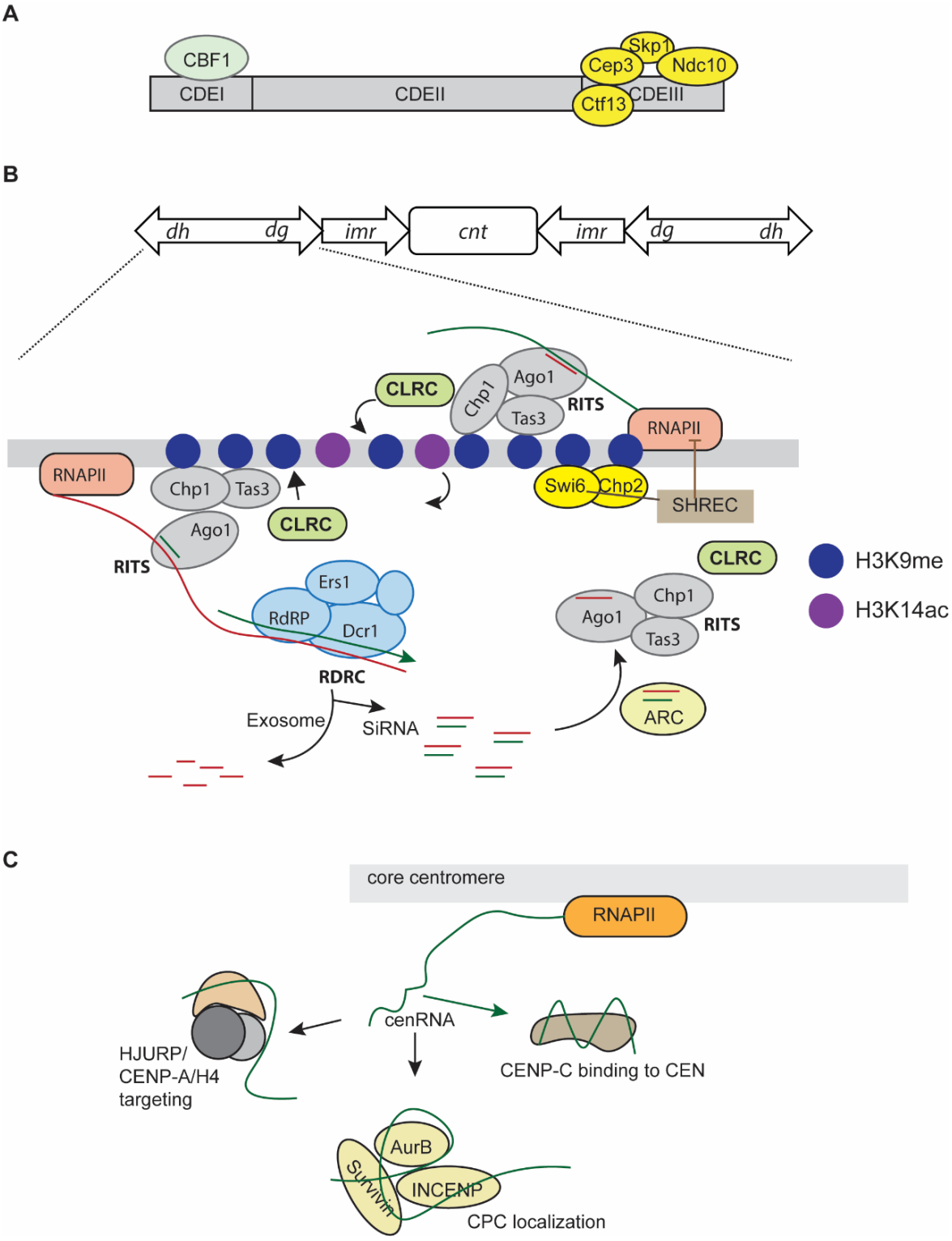


Figure 1-3. Genetic and epigenetic determinants of centromere identity.
(A) Centromeres of *S. cerevisiae* schematically depicted as a representative of genetically defined centromeres. The association of centromere DNA sequence in *S. cerevisiae* by a cognate CBF3

complex facilitates exclusive loading CENP-A^{Cse4} at the centromeres [adapted from (BIGGINS 2013)]. (B-C) Known epigenetic factors that were reported to govern centromere function in diverse systems are schematically depicted. These include the role of pericentric transcription and RNAi in *S. pombe* centromere function (B), and other known roles of transcription at core centromeres in CENP-A deposition, kinetochore assembly, and Aurora B kinase function (C).

Any defects in this cascade of events impaired *de novo* deposition of CENP-A^{Cnp1} onto minichromosomes. The studies on the impact of pericentric transcription in other systems were based on mutants of proteins that process these transcripts, like Dicer. Ablation of Dicer in Chicken-human hybrid cells led to mitotic defects, including precocious sister chromatid segregation as a consequence of loss of HP1 binding (FUKAGAWA *et al.* 2004). Similar to the case in *S. pombe*, dicer deficiency in the vertebrate cells mentioned above and in mouse embryonic stem cells resulted in the accumulation of pericentric transcripts highlighting a role in maintaining the heterochromatic state (FUKAGAWA *et al.* 2004; KANELLOPOULOU *et al.* 2005). While the pericentric transcription primarily contributes to the formation of heterochromatin, transcription at the core centromere is associated with kinetochore assembly and function (Figure 1-3C).

Transcripts from the core centromeres have been detected in several systems like the budding and fission yeasts (CHOI *et al.* 2011; OHKUNI AND KITAGAWA 2011), *Plasmodium falciparum* (LI *et al.* 2008), flies (ROSIC *et al.* 2014; BOBKOV *et al.* 2018), mouse (COHEN *et al.* 1973), and in human centromeres and neocentromeres (WONG *et al.* 2007; CHUEH *et al.* 2009). Indeed chromatin remodelers and RNAPII interacting proteins have been shown to facilitate the deposition of CENP-A in *Drosophila* (CHEN *et al.* 2015) and *S. pombe* (CARLSTEN et al. 2012) highlighting a role for both transcription and the resulting transcripts in centromere homeostasis. In the genetically defined point centromeres of *S. cerevisiae*, RNAPII derived cenRNA was shown to be critical for proper kinetochore assembly (OHKUNI AND KITAGAWA 2012). The transcription factor Cbf1 bound to centromere DNA (HEMMERICH *et al.* 2000) and was shown to be essential for producing CEN transcripts. Loss of function of Cbf1 or factors that regulate Cbf1 activity such as Ste12 or Dig1 resulted in chromosomal instability (OHKUNI AND KITAGAWA 2011). Similar to the above case, a GATA-binding factor Ams2 was reported to regulate the levels of CENP-A^{Cnp1} in *S. pombe* by its association with the CEN core sequences (CHEN *et al.* 2003). In humans, the physical association of a lncRNA from centromeres with pre-nucleosomal CENP-A/HJURP complex was shown to be critical for CENP-A deposition (CHAN AND WONG 2012; QUENET AND DALAL 2014). Another kinetochore protein CENP-C also binds to cenRNA and is critical for its localization and for proper kinetochore function in different organisms (WONG *et al.* 2007; DU *et al.* 2010; CHAN *et al.* 2012a; ROSIC *et al.* 2014). While the above examples highlight the importance of transcription at the

centromere core, over-expression of centromere transcripts was found to be counter-productive, inhibiting kinetochore assembly in both *S. cerevisiae* and in humans (HILL AND BLOOM 1987; BERGMANN *et al.* 2012). Taken together, these results suggest that a balanced and regulated level of transcription is critical for centromere function.

Besides kinetochore assembly, centromere transcription is also implicated in the recruitment of the chromosome passenger complex (CPC) to the inner centromeres and activation of Aurora B kinase that phosphorylates several kinetochore proteins essential for regulating microtubule attachments (IDEUE *et al.* 2014; BLOWER 2016; HINDRIKSEN *et al.* 2017). Recently a novel mechanism of Aurora B kinase activation, mediated by the formation of R-loops at the centromeres during mitosis was proposed (KABECHE *et al.* 2018). This finding adds another function to centromere transcription as a sensor for chromosomal attachments with the microtubules (PEREA-RESA AND BLOWER 2018).

Chromatin state

The distinct roles of transcription at the core centromere and the pericentromere is also evident by distinct chromatin signatures in the form of post-translational marks added to the histone H3 nucleosomes in these regions (see Figure 1-2) (SULLIVAN AND KARPEN 2004). The pericentric regions are often found to be heterochromatic, containing histone marks such as H3K9 hypermethylation (EYMERY *et al.* 2009; BERGMANN *et al.* 2012). As described in the previous section, the pericentric regions in *S. pombe* are enriched with H3K9me2 marks. Abrogation of the pericentric H3K9me2 marks has been shown to compromise CENP-A deposition and impaired mini-chromosome stabilization (VOLPE *et al.* 2002; VOLPE *et al.* 2003; FOLCO *et al.* 2008; ZHANG *et al.* 2008). On the other hand, signatures of transcriptionally active regions such as H3K4me2 and H3K36me2 have been detected at the core centromeres in humans and flies, and H3K4me2 at the core centromeres in *S. pombe*, suggesting a role of fine-tuned centromere transcription for the stable inheritance of centromeres (BLOWER *et al.* 2002; NAKANO *et al.* 2008; RIBEIRO *et al.* 2010; CHOI *et al.* 2011; OHZEKI *et al.* 2012). However, studies on several fungi identified repressive chromatin marks on repetitive centromeres (FRIEDMAN AND FREITAG 2017). In the case of *N. crassa* and *Fusarium fujikuroi*, the centromeric chromatin was found to be enriched in H3K9me3 mark (SMITH *et al.* 2011; SMITH *et al.* 2012; WIEMANN *et al.* 2013). While the loss of this mark was shown to result in shortening of CENP-A enriched regions in *N. crassa*, several developmental defects were observed in the case of *F. fujikuroi* (SMITH *et al.* 2011; WIEMANN *et al.* 2013). The H3K9me2 mark was found enriched at the centromere, pericentromere, and in sub-telomeric regions in *C. neoformans* (DUMESIC *et al.* 2015; YADAV *et al.* 2018b). The repetitive centromeres in chicken were also found to be enriched with H3K9me3 (HORI *et al.* 2014). The non-repetitive

centromeres in this organism are also heterochromatic, but alternative mechanisms were proposed for heterochromatin formation (FUKAGAWA 2017).

Besides histone H3, post-translational modification of other histones at the centromeres have also been reported. For instance, H4K5ac and H4K12ac are correlated with maintenance of CENP-A levels in chicken centromeres (HORI *et al.* 2014; SHANG *et al.* 2016; FUKAGAWA 2017). These marks were found predominantly in the pre-nucleosomal CENP-A complexed with its deposition machinery. Similarly, H4K20 mono-methylation, a transcription activation mark found commonly in CENP-A nucleosomes has been implicated in kinetochore assembly in both humans and chickens (SULLIVAN AND KARPEN 2004; WANG *et al.* 2008b; HORI *et al.* 2014). In *S. pombe*, mono-ubiquitination of H2A at K119 was shown essential for transcription at the centromere core (SADEGHI *et al.* 2014). Several modifications to CENP-A in humans have also been reported. These include ubiquitination and acetylation of K124, tri-methylation at G1, phosphorylation of S7, S16, and S18 (ZEITLIN *et al.* 2001; BAILEY *et al.* 2013; NIIKURA *et al.* 2015). However, the mechanistic details of their contribution to centromere function remain unclear.

Centromeres: rapidly evolving sites performing a conserved function

Analyses of centromere DNA sequences of closely related groups of species revealed that the centromeres show an enhanced rate of sequence divergence as compared to any other non-coding regions in the genome (MERALDI *et al.* 2006; BENSASSON *et al.* 2008; PADMANABHAN *et al.* 2008). This rapidly evolving nature of centromere sequences is commonly known as the 'centromere paradox' (HENIKOFF *et al.* 2001). In the point centromeres of *S. cerevisiae*, the DNA sequences corresponding to CDEII showed accelerated evolution as compared to other sequences (BENSASSON *et al.* 2008). Similar observations were derived by the comparison of centromere sequences from *C. albicans* and *C. dubliniensis* (PADMANABHAN *et al.* 2008). Both organisms have unique DNA sequences at the centromeres, even among its chromosomes. Strikingly, the centromeres of a related *Candida* species *C. tropicalis* that diverged 50 MYA from a common ancestor were found to contain homogenized core regions flanked by inverted repeats, marking a major transition in both sequence and structure at the centromere (CHATTERJEE *et al.* 2016). Among the fission yeast *Schizosaccharomyces* species, *Schizosaccharomyces japonicus* harbor transposon enriched centromeres while two related species *S. pombe* and *Schizosaccharomyces octosporus* lost transposons and contain repeats at the centromeres (RHIND *et al.* 2011).

This phenomenon is not restricted to fungal centromeres. One of the early reports on the rapid evolution of centromeres were based on the comparison of satellite repeat sequences between homologous centromeres of humans and chimpanzee (HAAF AND WILLARD 1997). Large scale analysis of

about 282 plant and animal centromere repeat elements identified variations in the centromeric tandem repeat monomer composition and length (MELTERS *et al.* 2013). The cultivated potato *Solanum tuberosum* contains both repetitive and repeatless centromeres in its chromosomes of which the satellite DNA sequences were found to have diverged from that present in the ancestral wild species *Solanum verrucosum* (GONG *et al.* 2012; ZHANG *et al.* 2014). Similarly, closely related rice species *Oryza sativa* and *Oryza brachyantha* harbor 155 bp and 154 bp long centromere repeat monomers respectively that do not share any sequence homology (LEE *et al.* 2005; YI *et al.* 2013).

Centromeres and species boundary

Much similar to the evolving nature of centromere DNA sequences, the kinetochore proteins that bound them were also found to show signatures of sequence evolution (MALIK AND HENIKOFF 2001; TALBERT *et al.* 2004; BAKER AND ROGERS 2006; MERALDI *et al.* 2006; SCHUELER *et al.* 2010). This adaptive evolution of centromeres and cognate kinetochore proteins have been proposed to result in reproductive isolation/ hybrid sterility in crosses between independently evolving populations (HENIKOFF *et al.* 2001; MALIK AND HENIKOFF 2009). In light of such an effect, speciation has been proposed to be an inevitable consequence of centromere evolution. Centromeres were indeed found to be species-specific, centromeres from one species are non-functional even in related species (ROACH *et al.* 2012). Functional incompatibilities in the centromere-kinetochore association have been shown to result in uniparental genome elimination in several interspecies plant hybrids (RAVI AND CHAN 2010; SANEI *et al.* 2011; WANG *et al.* 2014a) and proposed as an alternative method to produce haploid lines (COMAI 2014; WANG *et al.* 2019). In the case of viable hybrids of maize in the genetic background of oat, it was observed that maize centromeres expanded in size to match that of the other so that it can effectively associate with the mitotic spindle for proper segregation (WANG *et al.* 2014a).

Centromere dysfunction

Defects in pathways that regulating centromere assembly or function can lead to structural or numerical changes in the genome (MCKINLEY AND CHEESEMAN 2016). Structural changes typically include the rearrangements involving breaks at or adjacent to the centromeres. Indeed, an analysis of over 8000 cancer genomes revealed whole arm translocations, commonly derived from breakage and recombination at the centromere loci, suggestive of a compromised centromere structure (KIM *et al.* 2013). Centromere integrity can be compromised due to features like complex topological forms, secondary structure formation, and unfinished DNA replication (reviewed in (BARRA AND FACHINETTI 2018). Also, breakage at the centromeres results when imbalanced microtubule attachments are formed, leading to unequal forces from opposite poles during chromosome segregation (GUERRERO *et al.* 2010a;

GUERRERO *et al.* 2010b). Kinetochores mediate the attachments between centromeres and microtubules by physically linking them. Defects in the centromere/kinetochore assembly and function can compromises microtubule association and lead to massive chromosome segregation defects (SANTAGUIDA AND MUSACCHIO 2009). In unicellular organisms like fungi, eliciting chromosome segregation defects can be lethal and are attractive targets for therapeutic intervention. In the following section, the composition of the kinetochores complex, their association with the dynamic microtubules, and its significance in chromosome segregation are discussed.

The kinetochore complex

The kinetochore is an ensemble of several protein sub-complexes that assembles on a defined locus on a chromosome and anchors the mitotic spindle in a spatiotemporally regulated manner to facilitate segregation of chromosomes (CHEESEMAN 2014). Early understanding of the kinetochore was derived from the attempts to understand this locus that specifically attaches to the mitotic spindle, and dates back to the early 19th century. The kinetochore, implying *movement place*, was defined as 'the specialized region in the chromosome where the property of active mobility on the spindle is manifested, irrespective of number or extensions of these regions in the chromosome body and other properties that can be associated with the region of active mobility' (LIMA-DE-FARIA 1949). The terms centromere and kinetochore were used synonymously until the 1980s, owing to a lack of clear understanding of the kinetochore structure (BATTAGLIA 2003). This ambiguity was subsequently resolved to consider the kinetochore as the ultrastructural element that precisely attached to the spindle microtubules and centromeres, less precisely as the region on the chromosome the kinetochore is associated with (RIEDER 1982; BATTAGLIA 2003).

Initial insights into the kinetochore structure were primarily based on electron micrographs that could identify an inner and outer plate in the kinetochore separated by a translucent layer (RIEDER 1982). Studies over the past two decades have identified over 100 kinetochore proteins that are classified similarly: the chromatin-associated inner kinetochore and the microtubule-associated outer kinetochore (PESENTI *et al.* 2016; MUSACCHIO AND DESAI 2017) (Figure 1-4). Besides these structural components, the kinetochore ensemble also includes a regulatory module consisting of kinases and phosphatases that act as checkpoints to orchestrate error-free chromosome segregation in each cell cycle (FOLEY AND KAPOOR 2013; LONDON AND BIGGINS 2014).

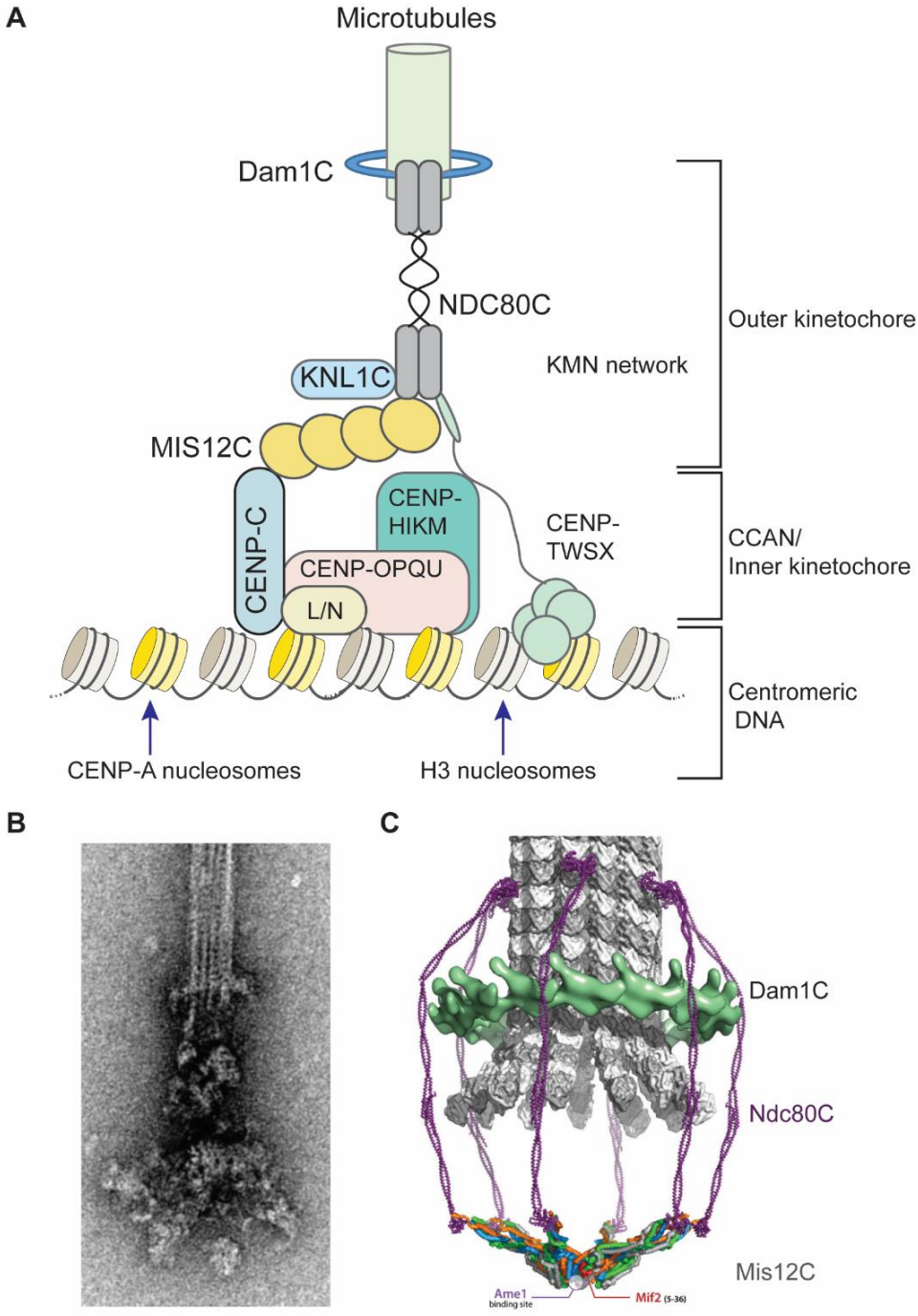


Figure 1-4. The kinetochore architecture in yeasts.
(A) Schematic representation of the yeast kinetochore ensemble with the constituent subcomplexes categorized into the inner and outer kinetochore. (B) Electron micrograph of *in vitro* purified *S. cerevisiae* kinetochore particles in association with microtubules (GONEN *et al.* 2012). (C) A model of yeast kinetochore associated with microtubules based on the structures known for the Mis12C, Ndc80C, and the Dam1C (DIMITROVA *et al.* 2016).

Structural organization of the kinetochore

The foundation to our present understanding of the composition and structure of the kinetochore was laid by a serendipitous discovery employing the use of serum from individuals with auto-immune syndrome CREST (**C**alcinosis, **R**eynaud's syndrome, **E**sophageal dysmotility, **S**clerodactyly, **T**elangiectasia) that localized to the kinetochores (MOROI *et al.* 1980; MOROI *et al.* 1981). Seminal studies by Bill Earnshaw's lab and Don Cleveland's Lab using these **a**nti-**c**entromere **a**ntibodies (ACA) identified three antigens that were detected by these antibodies- CENP-A, B, and C (**cen**tromere **p**roteins), marking the first identification of a centromere/kinetochore proteins (GULDNER *et al.* 1984; EARNSHAW AND MIGEON 1985; EARNSHAW AND ROTHFIELD 1985; PALMER *et al.* 1987; SAITOH *et al.* 1992; SULLIVAN *et al.* 1994; EARNSHAW 2015).

The Inner kinetochore layer

CENP-A - laying the foundation for the kinetochore

Coincidentally, the protein identified and named CENP-A was later found to be the hallmark of centromeric chromatin, earmarking centromeres as the sites for the assembly of kinetochore complex. Analysis of CENP-A identified it to be a histone H3 variant that is conserved from yeasts to humans (PALMER *et al.* 1991; SULLIVAN *et al.* 1994; STOLER *et al.* 1995).

At the sequence level, CENP-A shares a conserved histone fold domain with Histone H3 but contains additional features that distinguish it from the canonical one. CENP-A sequences contain a 22 amino acid-long insertion within the histone fold domain called **C**ENP-**A** **t**argeting **d**omain (CATD), comprising the loop 1 and alpha-helix 2 of CENP-A (Figure 1-5). This region was found to impart centromere specificity to CENP-A, replacing canonical H3 nucleosomes with CENP-A nucleosomes at the centromeres (BLACK *et al.* 2004; BLACK *et al.* 2007). Strikingly, the canonical H3, when expressed as a chimeric Histone H3 that included CATD, could be localized to the centromeres as well (BLACK *et al.* 2007). Another distinguishing feature in the CENP-A sequence is the presence of a highly variable N-terminus and a hydrophobic C-terminus as compared to the canonical H3 (MUSACCHIO AND DESAI 2017). In all of the systems studied till date, CENP-A localization has been a prerequisite for localization of any other kinetochore protein, placing it on the very top of the localization hierarchy (HOWMAN *et al.* 2000; MOORE AND ROTH 2001; OEGEMA *et al.* 2001; LIU *et al.* 2006; THAKUR AND SANYAL 2012; FACHINETTI *et al.* 2013; LOGSDON *et al.* 2015).

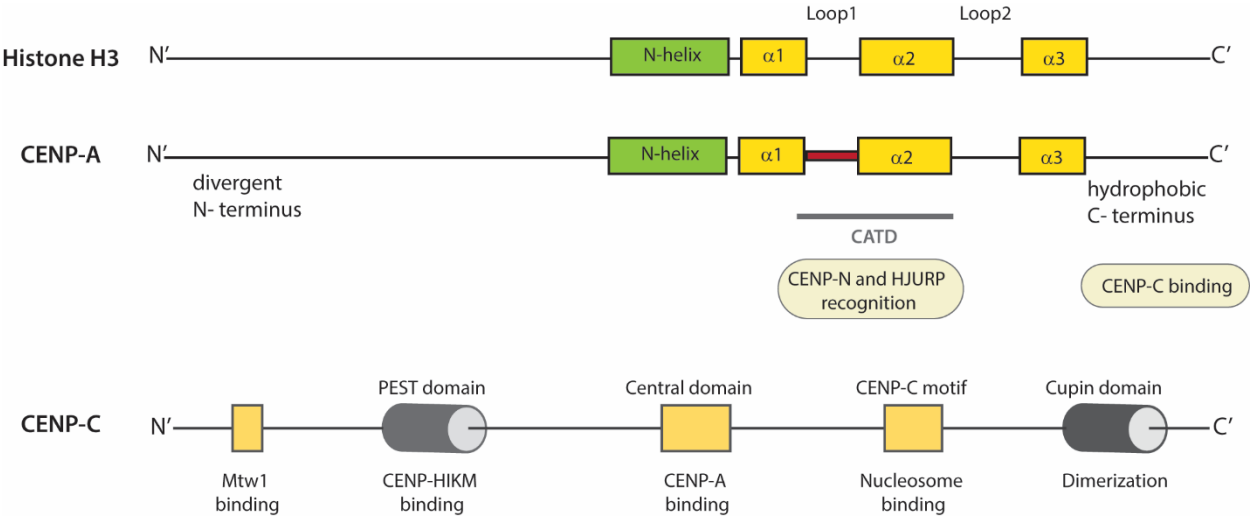


Figure 1-5. The domain architecture of CENP-A and CENP-C.
The sequence variations in CENP-A are depicted in comparison with histone H3. The loop1 and the alpha-helix2 constitute the CENP-A targeting domain (CATD) in CENP-A. The proteins known to interact with different regions of CENP-A are indicated below the line diagram. The motifs present in CENP-C (Mif2) along with the function of each of these regions is schematically represented.

This makes it critical to understand the factors and mechanisms involved in loading CENP-A (MELLONE *et al.* 2009; STELLFOX *et al.* 2013). This essential function is performed by the CENP-A chaperone HJURP along with the Mis16 and Mis18 complex in several different systems (CAMAHORT *et al.* 2007; STOLER *et al.* 2007; DUNLEAVY *et al.* 2009; FOLTZ *et al.* 2009; PIDOUX *et al.* 2009; WILLIAMS *et al.* 2009; BERNAD *et al.* 2011). The mode of action of this deposition machinery is dependent on the nature of centromere. For instance, HJURP interacts with the CBF3 complex proteins to load CENP-A^{Cse4} at the genetically defined point centromeres in the budding yeast (CAMAHORT *et al.* 2007). In the case of fission yeast, Scm3 recognizes and complexes with the Mis16, Mis18, and Mis6 to localize and deposit CENP-A^{Cnp1} at the kinetochores (PIDOUX *et al.* 2009; WILLIAMS *et al.* 2009). The protein Cal1 was found to be the homolog of Scm3 in flies wherein it was shown to be sufficient to recruit CENP-A^{CID1} upon tethering to an ectopic locus (CHEN *et al.* 2014). The deposition pathway in humans is proposed to be through a di-nucleosome model in which CENP-C, another inner kinetochore protein, plays a critical role in attracting the deposition machinery to the CENP-A nucleosomes. It is proposed that upon dilution of CENP-A during S- phase, H3.3 acts as a placeholder. A di-nucleosome of CENP-A/H3.3, along with the associated CENP-C, is recognized by the deposition machinery containing HJURP-Mis18 complex that chaperones prenucleosomal CENP-A into these regions (MUSACCHIO AND DESAI 2017).

Across different kingdoms of life, sporadic loss of CENP-A has been reported (DRINNENBERG *et al.* 2014; VAN HOOFF *et al.* 2017a; NAVARRO-MENDOZA *et al.* 2019). Despite this loss, organisms like trypanosomes and *Mucor circinelloides* retained monocentricity. Studies on such non-conventional model systems will shed light on the variations naturally seen in a conserved process such as chromosome segregation (AKIYOSHI AND GULL 2014; NAVARRO-MENDOZA *et al.* 2019).

The constitute centromere associated network (CCAN)

The identification of CENP-A, B, and C in the early 1990s was followed by the identification of two other proteins CENP-H and CENP-I, the latter being the homolog of fission yeast protein Mis6 (SAITOH *et al.* 1997; SUGATA *et al.* 1999; NISHIHASHI *et al.* 2002). Subsequent studies to identify the interactome of CENP-A, -H, and -I in vertebrate cells uncovered a group of about 20 proteins (CENP-K, CENP-L, CENP-M, CENP-N, CENP-O, CENP-P, CENP-Q, CENP-R, CENP-S, CENP-T, CENP-U, CENP-W, and CENP-X) that localized closely to the centromeric chromatin (FOLTZ *et al.* 2006; IZUTA *et al.* 2006; HORI *et al.* 2008). Based on biochemical reconstitutions and localization dependencies, these proteins were grouped into smaller subcomplexes- CENP-LN complex (POT *et al.* 2003; HINSHAW AND HARRISON 2013; MCKINLEY *et al.* 2015), CENP-HIKM complex (MEASDAY *et al.* 2002; DE WULF *et al.* 2003; OKADA *et al.* 2009; KLARE *et al.* 2015), CENP-OPQU complex (DE WULF *et al.* 2003; HORI *et al.* 2008), and CENP-TWSX complex (HORI *et al.* 2008; AMANO *et al.* 2009; NISHINO *et al.* 2012; SCHLEIFFER *et al.* 2012), which together constitute the CCAN (Figure 1-6) (PERPELESCU AND FUKAGAWA 2011). The homologs of most of these proteins are found in both budding yeast (Ctf19 complex and Cnn1 complex) and fission yeasts (Sim4, Mal2, and Fta (**F**our **t**wo **a**ssociated) proteins) (FLEIG *et al.* 1996; DE WULF *et al.* 2003; PIDOUX *et al.* 2003; LIU *et al.* 2005; BIGGINS 2013).

Of the known kinetochore proteins, only CENP-C and CENP-N can exclusively associate with CENP-A nucleosomes over H3 nucleosomes (CARROLL *et al.* 2009; CARROLL *et al.* 2010; KATO *et al.* 2013). These specific associations have functional significance in terms of recruitment of CENP-A to the kinetochores and linking the centromeric chromatin to the microtubule-binding module: the KMN network. The budding yeast homolog of CENP-N is Chl4, a protein known to be essential for *de novo* loading of CENP-A, but dispensable for propagation of previously established centromeric chromatin (MYTHREYE AND BLOOM 2003; POT *et al.* 2003).

CENP-C was identified along with CENP-A with the use of ACA as described earlier and was shown to localize to the kinetochore by immune electron microscopy (SAITOH *et al.* 1992). The gene was found to be conserved across eukaryotes, known as Mif2 in the budding yeast (MELUH AND KOSHLAND 1995). CENP-C/Mif2 is known to have an N-terminal domain that interacts with and recruits the Mis12/Mtw1 complex that subsequently forms the outer kinetochore and the CENP-HIKM complex that

recruits other members of CCAN (Figure 1-5) (GUSE *et al.* 2011; SCREPANTI *et al.* 2011; KLARE *et al.* 2015). The C- terminal half of CENP-C contains a central domain that directly interacts with the hydrophobic tail of CENP-A and a CENP-C motif that can bind to nucleosomes (GUSE *et al.* 2011; FACHINETTI *et al.* 2013; KATO *et al.* 2013; LOGSDON *et al.* 2015). According to the di-nucleosome model for CENP-A loading, these two domains of CENP-C recognize a CENP-A: H3.3 di-nucleosome and labels these sites for deposition of CENP-A (MUSACCHIO AND DESAI 2017). HJURP and Mis18 complex along with prenucleosomal CENP-A binds to the cupin domain of one such CENP-C, facilitating CENP-A loading at these marked sites. CENP-C is also known to interact with the acidic patches in H2A and H2B, helping stabilize the CENP-A nucleosomes (COHEN *et al.* 2008). In yeasts, Mif2 is one of the only three essential components of the Ctf19 complex (the yeast equivalent of the CCAN). Mutants of Mif2 show compromised levels of CENP-A and elevated levels of chromosome segregation defects (MELUH AND KOSHLAND 1995; MELUH AND KOSHLAND 1997; WESTERMANN *et al.* 2003). Mif2 forms one pathway of Mtw1 recruitment in the budding yeast kinetochore through an N-terminal Mtw1 binding motif (HORNUNG *et al.* 2011; HORNUNG *et al.* 2014; KILLINGER *et al.* 2020). Further, this association between Mif2 and Mtw1 is facilitated only when Mif2 is associated with centromeric nucleosomes, highlighting an autoinhibitory regulation of this interaction (KILLINGER *et al.* 2020).

The cooperative binding of CENP-C and CENP-LN, along with the CENP-HIKM complex, has been proposed to specifically recognize CENP-A nucleosomes and bridge it to the KMN network, forming the structural base for centromere function (WEIR *et al.* 2016). Knockdown of CENP-HIKM complex proteins showed defects in the localization of other CCAN proteins suggesting that these proteins function as an interdependent network (MCKINLEY *et al.* 2015). These proteins collectively bridge adjacent CENP-A nucleosomes, thereby imparting modularity in the kinetochore-microtubule interactions in higher eukaryotes (NAGPAL AND FUKAGAWA 2016; WEIR *et al.* 2016).

The CENP-OPQU complex, also known as the Ctf19 or COMA complex in yeasts, requires other CCAN proteins like CENP-C and CENP-HIKM for its kinetochore localization (ORTIZ *et al.* 1999; FOLTZ *et al.* 2006; HORNUNG *et al.* 2014; SAMEJIMA *et al.* 2015). Diverse functions have been ascribed for these proteins in different systems studied. In humans, CENP-Q and -U have been implicated in chromosome congression and recruitment of Plk1 that monitor kinetochore-microtubule attachments (AMARO *et al.* 2010; BANCROFT *et al.* 2015). In the budding yeast, Ctf19 and Mcm22 of the Ctf19 complex are not essential for viability, but mutants of these genes show defects in chromosome segregation and pericentric cohesion (FERNIUS AND MARSTON 2009). Additionally, mutants of Ctf19 result in spores with very poor viability indicative of an essential role in meiosis (AGARWAL *et al.* 2015). The two other subunits Okp1 and Ame1, are essential for viability in mitosis and are known to interact with Mis12/Mtw1, providing a second link between the inner and outer kinetochore apart from the link

provided by Mif2– Mtw1 interaction (Figure 1-6) (HORNUNG *et al.* 2014). Recently, it was shown that the region of Mtw1 recognized by Mif2 and Okp1/Ame1 were identical, but they do not compete for the binding site. This was suggestive of independent/parallel pathways to assemble Mtw1 and other outer kinetochore subunits (KILLINGER *et al.* 2020).

A third connection between the inner and outer kinetochore exists, in the form of CENP-TWSX complex, also known as the Cnn1 complex in budding yeasts (Figure 1-6) (HORI *et al.* 2008; AMANO *et al.* 2009; NISHINO *et al.* 2012). All four subunits of this tetrameric complex contain histone fold like domains that can form nucleosome like structure upon oligomerization (NISHINO *et al.* 2012). Indeed this complex was shown to bind inter-nucleosomal DNA upon incubation with dinucleosomes of CENP-A or H3 (TAKEUCHI *et al.* 2014; THAKUR AND HENIKOFF 2016). Localization of the CENP-T complex in humans is dependent on other CCAN proteins like CENP-C and CENP-HIKM, in addition to DNA binding (MCKINLEY *et al.* 2015; SAMEJIMA *et al.* 2015). Also, the ablation of CENP-TWSX resulted in defects in the outer kinetochore structure due to compromised KMN protein levels (HORI *et al.* 2008; SAMEJIMA *et al.* 2015). Indeed CENP-T was shown to recruit Ndc80 directly and indirectly through recruitment of additional Mis12 in line with the previous observations, forming the third linker between inner and outer kinetochore (RAGO *et al.* 2015; HUIS IN 'T VELD *et al.* 2016).

In the case of yeast kinetochore, *de novo* kinetochore assembly studies revealed that Cnn1 was dependent on all the CCAN subunits tested (includes Cse4, Mif2, Mcm22, Okp1, and Chl4) (LANG *et al.* 2018). This suggested that Cnn1 was the most distal component among the CCAN subunits. The N-terminus of Cnn1 has been shown to compete with the Mtw1 complex for anchoring the Ndc80 complex (MALVEZZI *et al.* 2013; PEKGOZ ALTUNKAYA *et al.* 2016). The transition from metaphase to anaphase is also marked by a switch from an Mtw1-anchored Ndc80 complex to a Cnn1-anchored Ndc80 complex. The anaphase enrichment gains relevance when the phosphorylation of Dsn1 is removed as cells progress through mitosis (LANG *et al.* 2018). However, the non-essential nature of Cnn1 suggests that the Mis12 and Okp1/Ame1 pathway recruit the majority of the Ndc80 complex in yeasts (HORNUNG *et al.* 2014).

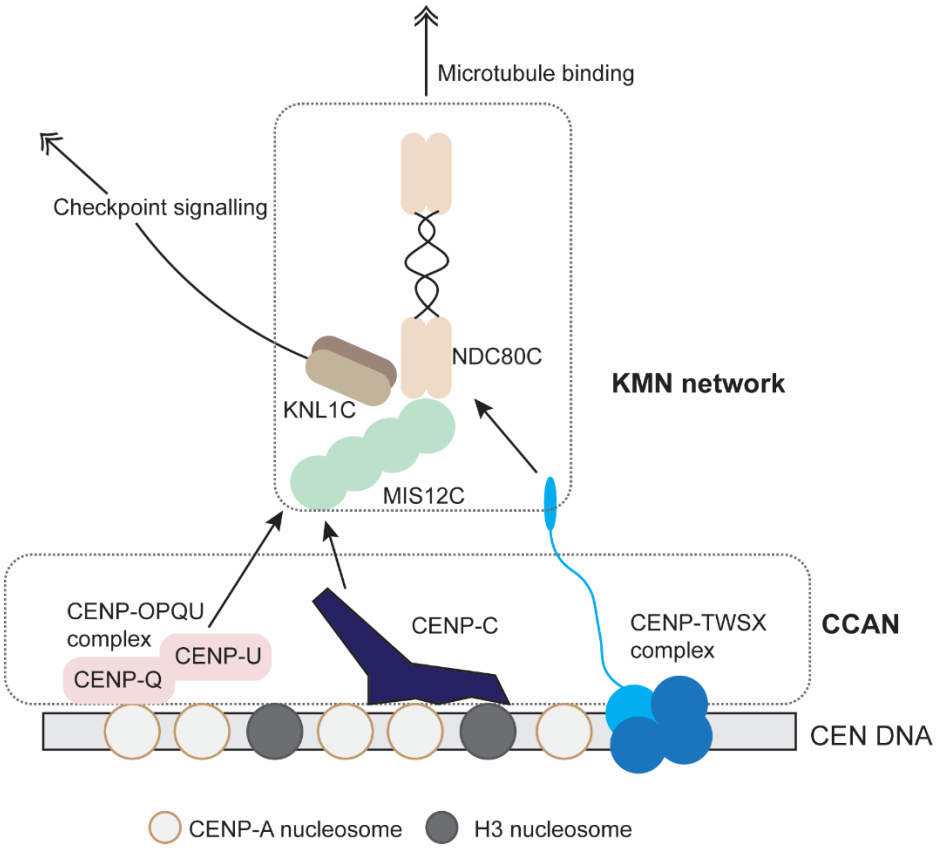


Figure 1-6. Multivalent attachments to the KMN network by CCAN.
One of the CCAN functions in forming multiple linker pathways between the KMN and inner kinetochore proteins is schematically represented. The three linker pathways are formed by CENP-C, CENP-O and CENP-Q of the CENP-OPQU complex, and the CENP-TWSX complex respectively.

The outer kinetochore layer

The KMN network

The KMN network is an ensemble of three complexes that include the Knl1C, Mis12C, and the Ndc80C (CHEESEMAN AND DESAI 2008; MUSACCHIO AND DESAI 2017). The KMN network forms the principle microtubule-binding module of the kinetochore with Ndc80C as the primary anchor point for the microtubules (CHEESEMAN *et al.* 2006). In line with this function, mutants of the KMN network fail to make microtubule attachments and exhibit mitotic arrest/delay phenotype (NEKRASOV *et al.* 2003; WEI *et al.* 2005; PINSKY *et al.* 2006; PAGLIUCA *et al.* 2009).

The Mis12C is a 4-protein complex comprised of Mis12/Dsn1/Nsl1/Nnf1 that does not bind microtubules but anchors the other two complexes that directly binds to microtubules (DE WULF *et al.* 2003; PINSKY *et al.* 2003; WESTERMANN *et al.* 2003; CHEESEMAN *et al.* 2006; MASKELL *et al.* 2010;

HORNUNG *et al.* 2011). The structure of the yeast and human Mis12 complexes were identified recently and was found to be an elongated structure with two globular heads (MASKELL *et al.* 2010; HORNUNG *et al.* 2011; DIMITROVA *et al.* 2016; PETROVIC *et al.* 2016). The Mis12/Mtw1 shows direct interaction with CENP-C and CCAN complex proteins, facilitating the kinetochore localization of the Mis12 complex. The C-termini of Dsn1 and Nsl1 was shown to anchor Ndc80C and Knl1C through interactions with Spc24-Spc25 and C-terminus of Knl1 respectively (WEI *et al.* 2005; MASKELL *et al.* 2010). The binding of Mis12C to the kinetochores is dependent on the phosphorylation of an autoinhibitory motif in Dsn1 by Aurora B kinase in both yeasts and humans (DIMITROVA *et al.* 2016; LADURNER AND STRAIGHT 2016).

The Knl1C, comprising of Knl1 and Kre28, forms one arm of kinetochore-microtubule interactions. The Knl1C was less studied *in vitro* owing to its poor stability until it was found to be stabilized upon co-expression with Mis12C (CHEESEMAN *et al.* 2006; PAGLIUCA *et al.* 2009). Knl1 and its yeast homolog Spc105 show a weak microtubule-binding activity, primarily mediated by the N-terminus of Knl1. The C-terminus is required for interactions with the Mis12C as mentioned earlier (MASKELL *et al.* 2010). Rather than force transduction, the primary role of Knl1C is to host the members of the spindle assembly checkpoint that monitors and regulates the kinetochore-microtubule interactions (KIYOMITSU *et al.* 2007; LIU *et al.* 2010; KIYOMITSU *et al.* 2011; LONDON *et al.* 2012; YAMAGISHI *et al.* 2012). A series of MELT repeats in Knl1 are phosphorylated by the mitotic kinases to regulate microtubule attachments (CALDAS AND DELUCA 2014).

The Ndc80C forms the other arm of kinetochore-microtubule interaction (CHEESEMAN *et al.* 2006; WEI *et al.* 2007; CIFERRI *et al.* 2008; ALUSHIN *et al.* 2010). The complex consists of two heterodimers, Spc24-Spc25 and Ndc80-Nuf2, that tetramerize to form a dumb-bell shaped structure *in vitro* (CIFERRI *et al.* 2005; WEI *et al.* 2005; WEI *et al.* 2006; WEI *et al.* 2007; CIFERRI *et al.* 2008). The contacts with inner kinetochore proteins and Mis12C are mediated by the Spc24-Spc25 end, and the interactions with microtubules are mediated by the Ndc80-Nuf2 end through the calponin homology domains in these subunits. Besides these two defined structures, Ndc80 contains two additional domains/motifs that influence its function. First, the N-terminus of Ndc80 contains a tail/ extension that has been shown to improve microtubule-binding efficiency (GUIMARAES *et al.* 2008; MILLER *et al.* 2008). The second structural feature is the internal Ndc80 loop region that interrupts the coiled-coil regions of Ndc80. While the function of the Ndc80 loop is not explicitly clear, it is shown to be critical for interactions with the Dam1/Ska complex proteins (WANG *et al.* 2008a; MAURE *et al.* 2011; VARMA *et al.* 2012).

The couplers for kinetochore-microtubule interactions: Dam1/Ska complex

While the KMN network constitutes the core microtubule-binding module of kinetochores, its efficiency is augmented by accessory complexes such as the Dam1 complex and the Ska complex in different organisms (Figure 1-7A) (LAMPERT *et al.* 2010; TIEN *et al.* 2010; HELGESON *et al.* 2018).

The Dam1 complex

The Dam1 complex is a heterodecameric complex consisting of Dad1, Dad2, Dad3, Dad4, Dam1, Duo1, Ask1, Spc19, spc24, and Hsk3 that has been shown to form a ring-like structure *in vitro* and *in vivo* (HOFMANN *et al.* 1998; JONES *et al.* 1999; CHEESEMAN *et al.* 2001a; CHEESEMAN *et al.* 2001b; ENQUIST-NEWMAN *et al.* 2001; JANKE *et al.* 2002; MIRANDA *et al.* 2005; WESTERMANN *et al.* 2005; NG *et al.* 2019). The Dam1 complex subunits are essential for viability in most of the organisms where it has been characterized, with mutants showing mitotic arrest and defects in maintaining proper microtubule attachments (THAKUR AND SANYAL 2011; CHATTERJEE *et al.* 2016; SRIDHAR *et al.* 2019). These proteins were found non-essential in *M. oryzae* and *S. pombe*, but the mutants showed elevated levels of missegregation (LIU *et al.* 2005; SHAH *et al.* 2019).

The ability to co-express and purify all the ten subunits of the Dam1 complex paved the way for the identification of a ring-like structure and its contribution towards microtubule association (Figure 1-7B) (MIRANDA *et al.* 2005). Each monomer of the Dam1 ring is 160 Å long rod that is made of two coiled-coil arms that merge in the middle from which a protrusion emerges perpendicular to the arms (Figure 1-7C) (JENNI AND HARRISON 2018). The Arms I and II are composed of Dad2-Dad4-Ask1-Spc19-Hsk3 and Dad1-Dad3-Dam1-Duo1-Spc34 respectively. The protrusion is formed by the extended C-terminus of Spc19 and Spc34. The junction where the two arms merge and the protrusion emerge forms the central domain that contains buried hydrophobic amino acids of the subunits. Truncations or perturbation of the subunits that can expose the hydrophobic residues in the central domain can result in disassembly or degradation of the Dam1 ring structure. Of the ten subunits, interactions with tubulin monomers have been detected only for Dam1 and Duo1. Interactions with Ndc80 complex subunits have been detected at the Ask1 C-terminus and the Spc19-Spc34 protrusions.

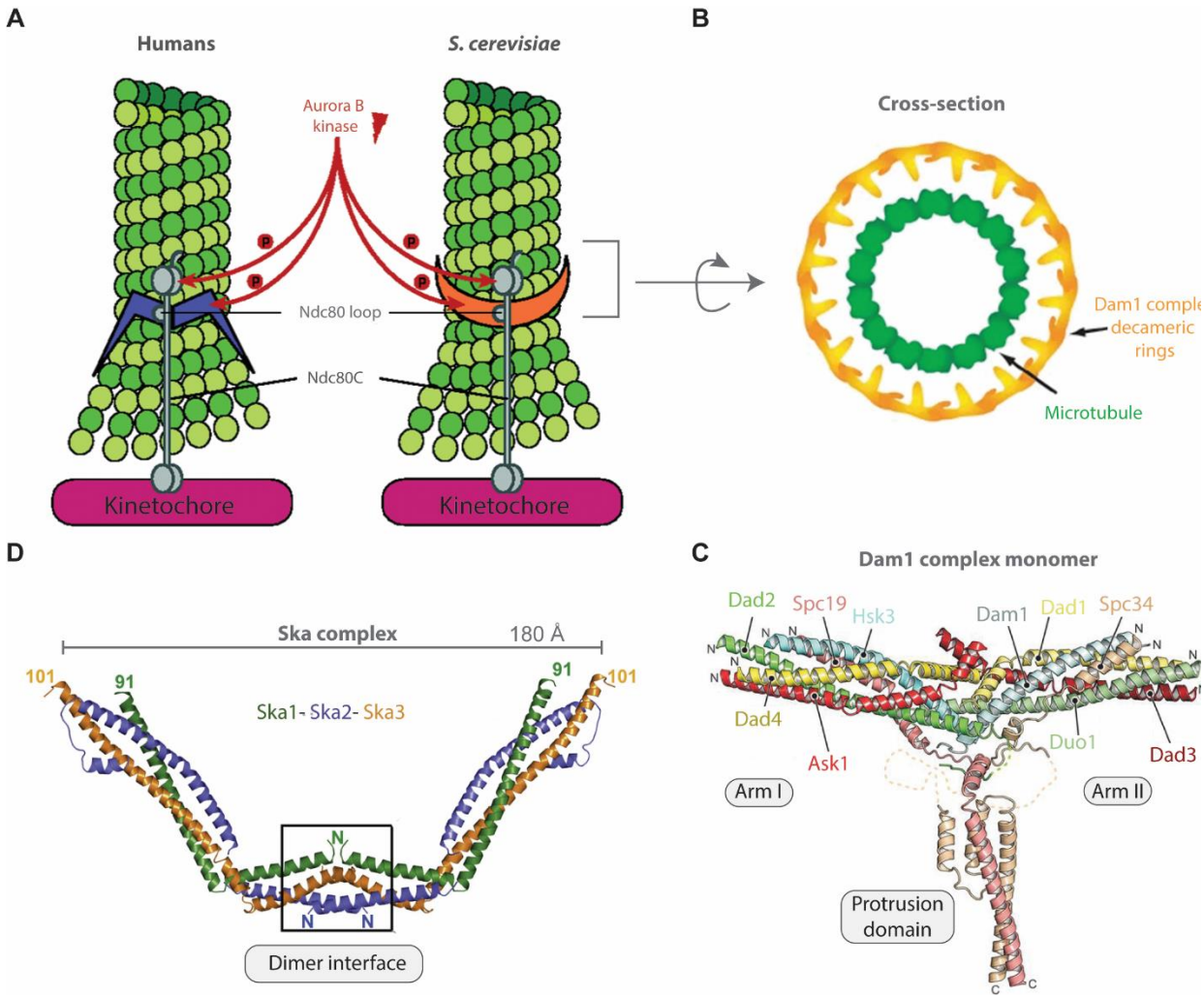


Figure 1-7. Structural organization of the outer kinetochore Dam1 and Ska complex.
(A) Illustration of functional similarities between Ska1 and Dam1 complex at the kinetochore-microtubule interface. (B) A cross-section view of the Dam1 complex rings encircling microtubules. (C) The subunit organization of the Dam1 complex monomer indicating the two coiled-coil arms and the protrusion domain. (D) The structure of the Ska complex dimer forming a wedge-like structure with the open ends of the dimer interacting with microtubules. [Figures adapted from (SANDALL AND DESAI 2007; JEYAPRAKASH *et al.* 2012; VAN HOOFF *et al.* 2017b; JENNI AND HARRISON 2018)]

Studies on the Dam1 complex rings *in vivo* suggested that the association with microtubules was critical for Dam1 ring formation (NG *et al.* 2019). The experimental ablation of microtubule association resulted in the accumulation of partial rings *in vivo*. The microtubule-binding by Dam1 and Duo1 forms an interaction bridge between the Dam1 complex and the kinetochores (LEGAL *et al.* 2016; JENNI AND HARRISON 2018; NG *et al.* 2019). When the underlying microtubule protofilament disassembles, the interaction bridge is broken, enabling the Dam1 complex to diffuse away from the depolymerizing ends. Despite depolymerization, the stable association of the Dam1 complex is ensured by curved protofilament

ends that trap the Dam1 complex on one side along with the formation of transient interaction bridge between the complex and the curved microtubule. Once the diffused Dam1 complex finds a linear protofilament, the interaction bridge is reestablished. Cycles of dissociation and reassociation of the interaction bridge between the Dam1 complex and microtubules are proposed to enable stable attachments with the microtubule plus ends (NG *et al.* 2019).

The Ska complex

The function of Dam1 complex in higher eukaryotes is taken over by a three subunit Ska complex (WELBURN *et al.* 2009; ABAD *et al.* 2014; ABAD *et al.* 2016; VAN HOOFF *et al.* 2017b). Unlike the Dam1 complex that forms rings, analysis of Ska complex structure revealed it to be W- shaped with Ska1 and Ska3 exhibiting microtubule-binding activity (Figure 1-7D) (JEYAPRAKASH *et al.* 2012; ABAD *et al.* 2014). While the Ndc80 complex can associate only with linear protofilaments, the Ska complex can remain stably associated with both linear and curved protofilaments. This is suggestive of a role for the Ska complex in retaining microtubule attachments at the depolymerizing ends (ABAD *et al.* 2014).

The Dam1 and Ska complexes are dependent on the Ndc80 complex for kinetochore localization. Both these complexes can bind microtubules *in vitro*, track depolymerizing ends of microtubules, remain stably associated with both linear and curved protofilaments, and can transport cargo with this property. The microtubule-binding of both these protein complexes are electrostatic and is under the surveillance of Aurora B kinase. Based on these lines of evidence, the Ska complex was proposed to be functionally analogous to the Dam1 complex (WELBURN *et al.* 2009; ABAD *et al.* 2014).

Kinetochore function

The kinetochore ensemble, in conjunction with the microtubules, must perform each of the following functions to segregate chromosomes between cells faithfully. These functions can be broadly categorized as (i) to facilitate the attachment of chromosomes to microtubules and couple the microtubule dynamics to the movement of chromosomes, and (ii) to monitor the kinetochore-microtubule interactions to detect and reset any aberrant associations (HINSHAW AND HARRISON 2018).

Kinetochore-microtubule attachments

Prior metaphase, one sister kinetochore of each chromosome makes a lateral attachment with the mitotic spindle such that it binds to the microtubule side/wall. The lateral attachment is one of the earliest attachments that occur between the chromosome/kinetochore and microtubules with which chromosome congression occurs, enabling the chromosomes to align at the metaphase plate (TANAKA *et al.* 2005a;

KAPOOR *et al.* 2006; MAGIDSON *et al.* 2011). This process is facilitated by additional proteins like CENP-E and dynein (MCEWEN *et al.* 2001; LI *et al.* 2007; YANG *et al.* 2007).

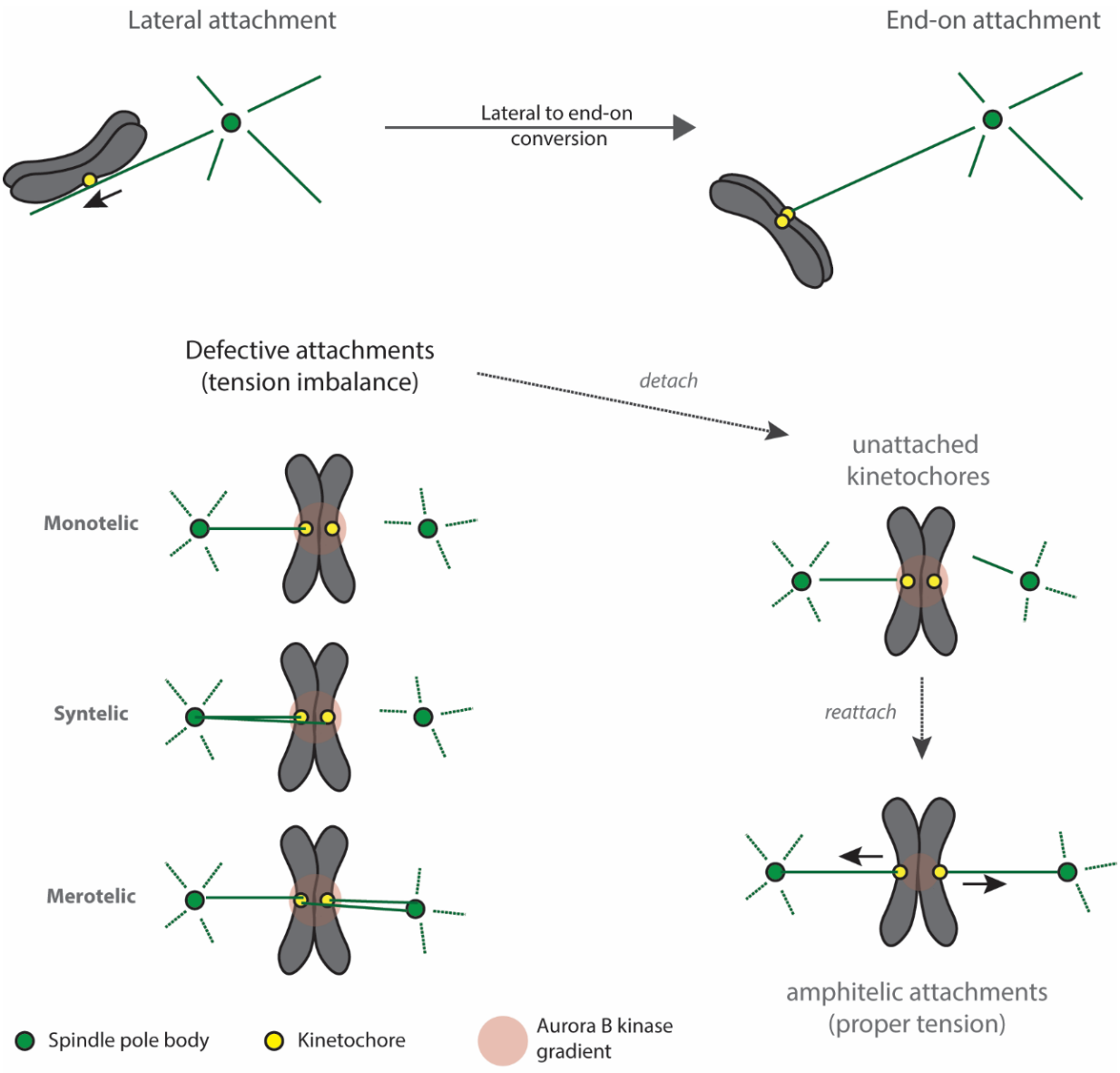


Figure 1-8. Types of kinetochore-microtubule attachment and their regulation.
Lateral association with the kinetochores is converted to end-on attachments by end-on conversion. Defective attachments in the form of monotelic, syntelic, or merotelic attachments, which do not generate tension, place the kinetochore cluster within the Aurora B kinase gradient (red halo). This enables the phosphorylation of substrates like Ndc80 and Dam1 by Aurora B kinase resulting in unattached kinetochores that are allowed to form proper bivalent attachments/ amphitelic attachments.

By a process called end-on conversion (SHRESTHA AND DRAVIAM 2013; SHRESTHA *et al.* 2017), the chromosomes reorient such that the unbound sister kinetochore takes a geometrically favorable position to encounter and bind microtubule from the opposing spindle pole (Figure 1-8). When such an attachment occurs (also called amphitelic attachment), the chromosome is said to be bi-oriented and is poised for the transition to anaphase (TANAKA *et al.* 2005b). However, cases of no attachment or less favored modes of attachment can also occur at this stage. These include (i) monotelic attachment in which only one sister kinetochore is attached to a microtubule while the other one remains unattached (ii) syntelic attachment wherein both the sister kinetochores capture microtubules emanating from the same spindle pole, and (iii) merotelic attachment in which one of the sister kinetochores makes an attachment with microtubules from both the spindle poles (Figure 1-8) (TANAKA 2010; GREGAN *et al.* 2011).

Whenever an erroneous attachment occurs, the cell senses and destabilizes these attachments providing another chance to make amphitelic attachment. One form of correction is based on the stability of kinetochore-microtubule attachments that is consequential of a balanced tension across the sister kinetochores (LI AND NICKLAS 1995; GRISHCHUK *et al.* 2005; FRANCK *et al.* 2007; AKIYOSHI *et al.* 2010). Any of the above-mentioned defective attachments does not create tension, rendering these attachments unstable and resulting in dissociation. However, additional sensors in the form of mitotic kinases such as Mps1 and Aurora B/Ipl1 identify such tension-defective attachments (BIGGINS AND MURRAY 2001; TANAKA *et al.* 2002; TANAKA 2010).

Tension sensors and the spindle assembly checkpoint

The Aurora B kinase is a tension sensor that localizes to the inner centromere as a part of the **c**hromosomal **p**assenger **c**omplex (CPC) (CARMENA *et al.* 2012; KRENN AND MUSACCHIO 2015). The microtubule-binding proteins in the kinetochore, including KMN network and Dam1/Ska complex proteins, harbor Aurora B phosphorylation sites (CHEESEMAN *et al.* 2002; CHEESEMAN *et al.* 2006; TIEN *et al.* 2010; WELBURN *et al.* 2010; DELUCA *et al.* 2011; ABAD *et al.* 2014). The spatial separation model was proposed to explain the tension sensing roles of Aurora B kinase at the kinetochores (LAMPSON AND CHEESEMAN 2011). In case of a bi-oriented state, the tension generated by the opposing spindles stretch the kinetochore proteins such that the outer kinetochore proteins fall outside the zone where they can be phosphorylated by Aurora B. In case of a defective attachment that fails to create tension, the force imbalance retains outer kinetochore closer to the centromere, facilitating phosphorylation of target sites. Given the electrostatic nature of KMN/Dam1/Ska complex interactions with the microtubules, phosphorylation results in detachment from microtubules that sends a “stop anaphase” signal by the activation of **s**pindle **a**ssembly **c**heckpoint (SAC).

Initial insights on the molecular composition of SAC complex and their function were derived from genetic screens in yeasts that identified *budding uninhibited by benzimidazole* (Bub mutants: BUB1, BUB2, BUB3) proteins and *mitotic arrest deficient* (Mad mutants: MAD1, MAD2, MAD3) proteins. (HOYT *et al.* 1991; LI AND MURRAY 1991). These proteins are recruited to the kinetochores in the event of unattached kinetochores that result in phosphorylation of MELT repeats of Knl1 by Mps1 kinase (LONDON *et al.* 2012; YAMAGISHI *et al.* 2012). Upon kinetochore localization and association with Mad1, Mad2 undergoes a structural transition from an open state to a closed state. Subsequently, these proteins form a **m**itotic **c**heckpoint **c**omplex (MCC) that sequesters Cdc20 that is required for the function of the **a**naphase-**p**romoting **c**omplex (APC/cyclosome) (Figure 1-9). The resulting anaphase delay forms the time window by which proper kinetochore-microtubule attachments are allowed to occur (reviewed in (LONDON AND BIGGINS 2014) and the references therein).

Once bi-orientation is achieved, the checkpoint is 'silenced' by various means such as dynein-mediated removal of checkpoint subunits (HOWELL *et al.* 2001), sequestering Mad2 by p31 (XIA *et al.* 2004), and the phosphatase activity of PP1/Glc7 and Cdc14 counteract phospho-regulation (PINSKY *et al.* 2009; MIRCHENKO AND UHLMANN 2010). At this juncture, the checkpoint is said to be satisfied, resulting in the release of Cdc20 followed by degradation of Securin/Pds1, activation of Separase, and anaphase onset.

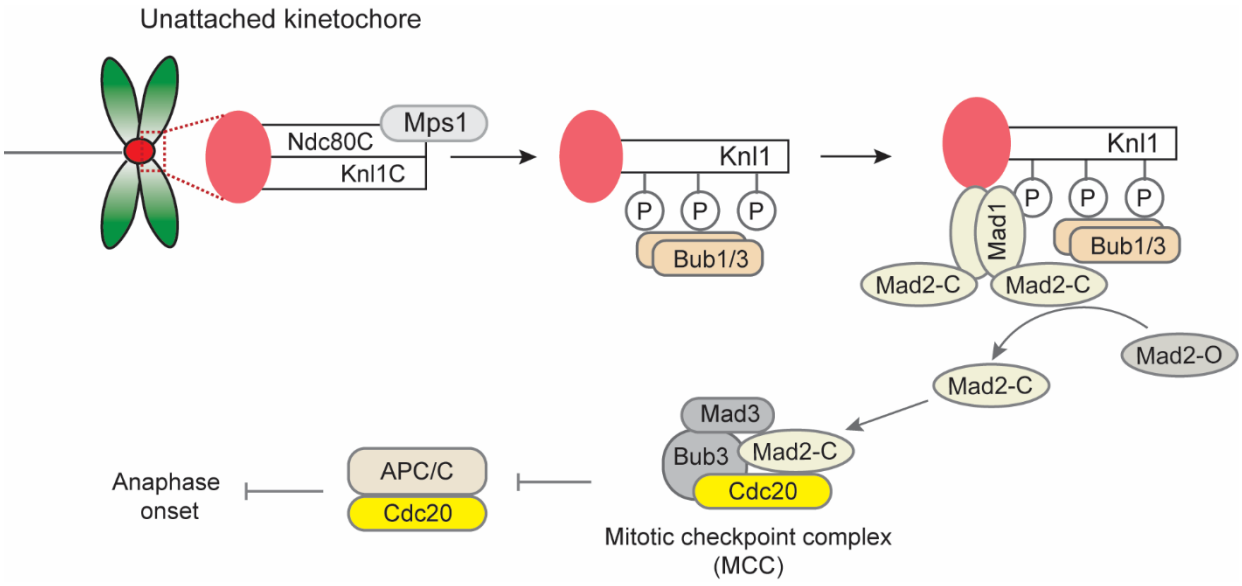


Figure 1-9. Activation of the spindle assembly checkpoint.
The stepwise assembly of a functional checkpoint-signaling complex initiated by the phosphorylation of Knl1 by Mps1 at the unattached kinetochores is schematically represented. Formation of the mitotic checkpoint complex sequesters Cdc20 and inhibits APC/C function, resulting in mitotic arrest.

Introduction to the model organism: The *Malassezia* species complex

Phylum	: Basidiomycota
Subphylum	: Ustilagomycotina
Class	: Malasseziomycetes
Order	: Malasseziales
Family	: *Malasseziaceae*
Genus	: *Malassezia*

Etiology

The *Malassezia* yeasts have been known to humans for over 150 years as a part of human cutaneous microflora that was thought to be the etiological agents of certain skin disorders (EICHSTEDT 1846). One of the first reports of identification of *Malassezia* came as early as in 1856 when the fungal nature of **p**ityriasis **v**ersicolor (PV), as observed by the presence of yeast and filamentous cells in the lesions was recognized (EICHSTEDT 1846). Presently, the *Malassezia* species complex consists of 18 species that belong to the Malasseziomycetes of the Basidiomycota (Figure 1-10) (THEELEN *et al.* 2018). However, the journey to arrive at this taxonomic classification was chaotic, filled with contradicting reports mainly due to their fastidious nature for *in vitro* culturing.

The species was first named in 1853 as *Microsporon furfur* by Robin, which was later changed by Baillon in 1889 as *Malassezia furfur* when he created a new genus *Malassezia* (reviewed in (ASHBEE AND EVANS 2002; BOEKHOUT *et al.* 2010)). Later in 1902, Sabouraud placed the yeast forms- oval and spherical cells, into one genus *Pityrosporum* (SABOURAUD 1902). While he observed similarities between the yeast cells associated with PV and **s**eborrheic **d**ermatitis/**d**andruff (SD/D), the presence of filaments in the lesions associated with PV and not on the skin associated with SD/D prevented him unify *Malassezia furfur* and *Pityrosporum* species under a single genus. Subsequent studies by Panja in 1927, wherein preliminary methods to culture these species were reported, identified similarities between the *M. furfur* and *Pityrosporum* species and suggested a single genus, *Malassezia*, for these yeasts (PANJA 1927).

However, this nomenclature was not widely accepted until the late 1970s when similarities between these species were shown by several means such as morphology, ultrastructure, immunological properties, and the ability of the *Pityrosporum* species to exhibit filamentous growth upon induction in culture (STERNBERG AND KEDDIE 1961; BARFATANI *et al.* 1964; PIERARD AND DOCKX 1972; DORN AND ROEHNERT 1977; PORRO *et al.* 1977; TANAKA AND IMAMURA 1979).

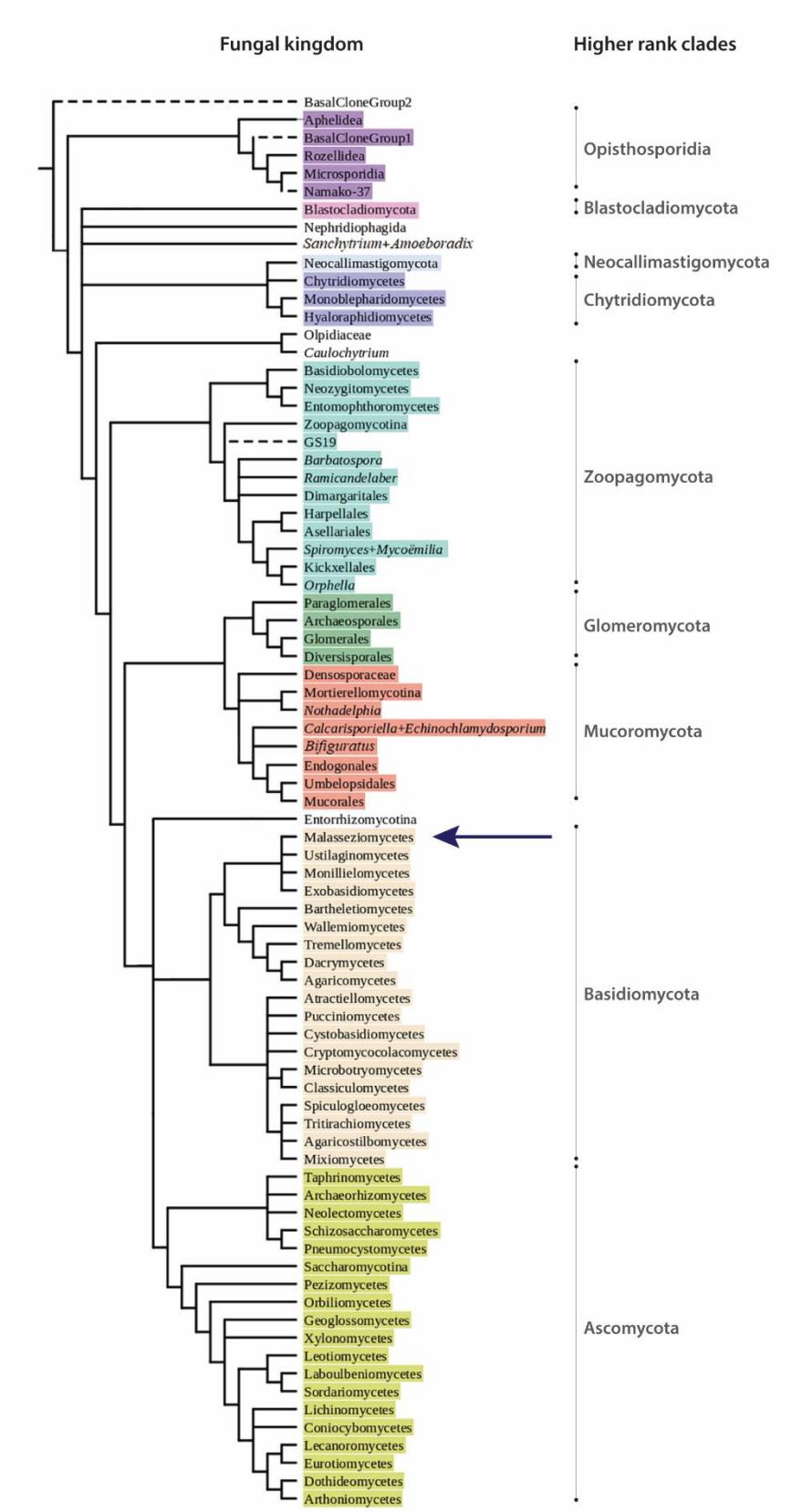

Fungal kingdom
Higher rank clades
BasalCloneGroup2
Aphelidea
BasalCloneGroup1
Rozellidea
Microsporidia
Namako-37
Opisthosporidia
Blastocladiomycota
Blastocladiomycota
Nephridiophagida
Sanchytrium+Amoeboradix
Neocallimastigomycota
Neocallimastigomycota
Chytridiomycetes
Monoblepharidomycetes
Hyaloraphidiomycetes
Chytridiomycota
Olpidiaceae
Caulochytrium
Basidiobolomycetes
Neozygitomycetes
Entomophthoromycetes
Zoopagomycotina
GS19
Barbatospora
Ramicandelaber
Dimargaritales
Harpellales
Asellariales
Spiromyces+Mycoëmilia
Kickxellales
Orphella
Zoopagomycota
Paraglomerales
Archaeosporales
Glomerales
Diversisporales
Glomeromycota
Densosporaceae
Mortierellomycotina
Nothadelphia
Calcarisporiella+Echinochlamydosporium
Bifiguratus
Endogonales
Umbelopsidales
Mucorales
Mucoromycota
Entorrhizomycotina
Malasseziomycetes
Ustilaginomycetes
Monillielomycetes
Exobasidiomycetes
Bartheletiomycetes
Wallemiomycetes
Tremellomycetes
Dacrymycetes
Agaricomycetes
Atractiellomycetes
Pucciniomycetes
Cystobasidiomycetes
Cryptomycocolacomycetes
Microbotryomycetes
Classiculomycetes
Spiculogloeomycetes
Tritirachiomycetes
Agaricostilbomycetes
Mixiomycetes
Basidiomycota
Taphrinomycetes
Archaeorhizomycetes
Neolectomycetes
Schizosaccharomycetes
Pneumocystomycetes
Saccharomycotina
Pezizomycetes
Orbiliomycetes
Geoglossomycetes
Xylonomycetes
Leotiomycetes
Laboulbeniomycetes
Sordariomycetes
Lichinomycetes
Coniocybomycetes
Lecanoromycetes
Eurotiomycetes
Dothideomycetes
Arthoniomycetes
Ascomycota

Figure 1-10. Cladogram based on multi-gene phylogenies depicting major classes of fungi. The position of *Malasseziomyctes* is indicated by the blue arrowhead. [Figure adapted from (NARANJO-ORTIZ AND GABALDON 2019)]

Identification of new *Malassezia* species

In a seminal work by Guillot and Gueho in 1995, about 100 species of *Malassezia* and *Pityrosporum* that included most of the species described till then, were compared using LSU rRNA sequences and nuclear DNA complementarity (GUILLOT AND GUEHO 1995). Their analysis identified eight species clusters that included *M. furfur*, *M. sympodialis*, and *M. pachydermatis*, with the rest being named as *Malassezia* species 1-5. In the following year, the same group described four new taxa that included *M. globosa*, *M. restricta*, *M. slooffiae*, and *M. obtusa* (GUEHO *et al.* 1996). The advent of pulsed-field gel electrophoresis (PFGE) based karyotyping, random amplification of polymorphic DNA (RAPD), restriction fragment length polymorphism (RFLP) analysis, and sequencing technology along with several culture-based assays paved the way for the classification of novel *Malassezia* species that were subsequently identified (BOEKHOUT AND BOSBOOM 1994; BOEKHOUT *et al.* 1998; SENCZEK *et al.* 1999).

By 2011, the number of species in the *Malassezia* genus rose to 14, including new species like *M. japonica* (SUGITA *et al.* 2003a), *M. dermatis* (SUGITA *et al.* 2002), *M. yamatoensis* (SUGITA *et al.* 2004), *M. nana* (HIRAI *et al.* 2004), *M. caprae* (CABANES *et al.* 2007), *M. equina* (CABANES *et al.* 2007), and *M. cuniculi* (CABANES *et al.* 2011). In the last decade or so, four more species, *M. psittaci* (CABANES *et al.* 2016), *M. brasiliensis* (CABANES *et al.* 2016), *M. vespertilionis* (LORCH *et al.* 2018), and *M. arunalokei* (HONNAVAR *et al.* 2016), identified from diverse sources were included in this genus making a total of 18 species in the *Malassezia* species complex at present.

Phylogenetic classification

The *Malassezia* species were placed under the polyphyletic class Exobasidiomycetes, which along with classes Ustilagomycetes and Entorrhizomycetes, formed the subphylum Ustilagomycotina of the Basidiomycota (Figure 1-10 and Figure 1-11). Multi-gene phylogenetic analysis identified these species to form a highly supported deep lineage with conserved features within the members and showed distinct genomic features and gene repertoire when compared to the closely related *Ustilago* species that are well-known plant pathogens (BEGEROW *et al.* 2006; WANG *et al.* 2014b). Further, the exclusion of *Malassezia* species from the phylogenetic analysis rendered the Exobasidiomycetes as a monophyletic class. Based on these analyses and phenotypic comparisons, a new class Malasseziomycetes was proposed within the Ustilagomycotina (WANG *et al.* 2014b)

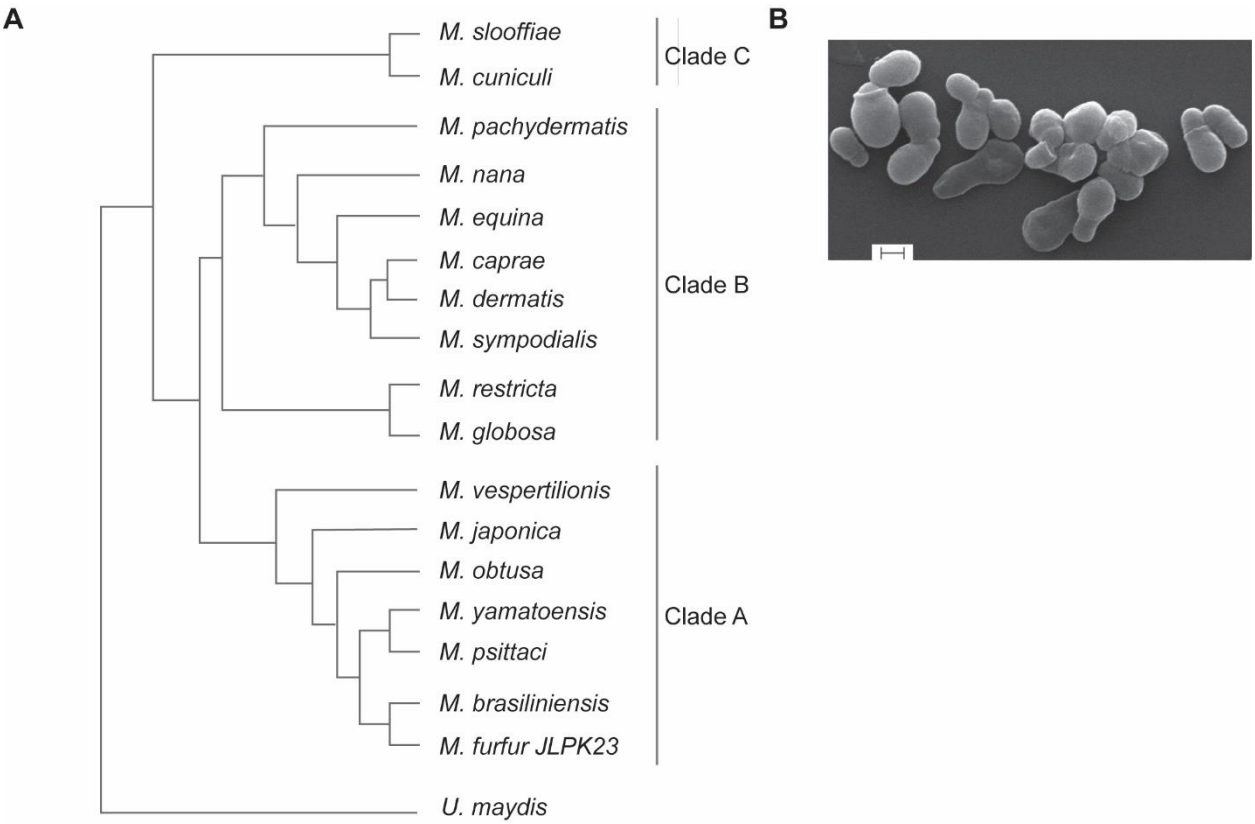


Figure 1-11. The *Malassezia* species complex.
(A) A cladogram depicting the phylogenetic relationship between 17 *Malassezia* species along with their classification into clades is represented [adapted from (THEELEN *et al.* 2018)]. (B) Scanning electron micrograph of *M. japonica* indicating cylindrical yeast-like cells along with buds (HONNAVAR *et al.* 2015).

Ecology and clinical relevance

Association with the human skin

Traditional culture bases studies of microbes associated with human hosts yielded an incomplete list owing to lack of appropriate culture conditions that ensure equal representation of all resident species including *Malassezia*. Analysis of the human skin microbiome using culture-independent, sequencing-based methods identified *Malassezia* species as the most abundant microbe detected in the human skin (Figure 1-12) (FINDLEY *et al.* 2013; FINDLEY AND GRICE 2014; OH *et al.* 2014). A total of 11 *Malassezia* species were identified in the study that constituted ~50-80% of the total skin mycobiome, with *M. restricta* and *M. globosa* being the most dominant ones followed by *M. sympodialis* to a lesser extent. Colonization of human skin with *Malassezia* is detected soon after birth albeit at low levels, but their rise

shows a strong association with the onset of puberty and the activation of sebaceous glands, the secretions of which provide a suitable habitat for these lipid dependent yeasts (NAGATA *et al.* 2012; JO *et al.* 2016). Besides healthy skin, these species are also detected in the lesions or regions affected by various skin disorders such as PV, SD/D, and atopic dermatitis, making the establishment of a causal relationship between *Malassezia* and these disorders complicated (FINDLEY *et al.* 2013; PROHIC *et al.* 2016; DE HOOG *et al.* 2017; THEELEN *et al.* 2018). Several independent reports suggest that *M. globosa*, *M. restricta*, *M. sympodialis* along with *M. obtusa* and *M. furfur* to a smaller extent are predominantly detected in skin/ lesions associated with the above disease conditions (DE HOOG *et al.* 2017).

Beyond the human skin

Besides their known contributions in skin disorders, *Malassezia* species such as *M. pachydermatis*, *M. furfur*, and *M. sympodialis* were also reported from systemic infections, especially in neonates receiving exogenous lipid supplementation through implants like catheters.

Recent studies on humans and mouse models suggest a role for *Malassezia* in the onset and progression of conditions such as Crohn's disease and pancreatic cancer. *M. restricta* and *M. globosa* were found to be the two predominant fungi present in the gut-associated with Crohn's disease and showed a strong association with individuals carrying a polymorphism S12N in the CARD9 gene (LIMON *et al.* 2019; SPARBER *et al.* 2019; WRIGHTON 2019). This provides a direct link between *Malassezia* colonization and inflammatory response in the gut. Also, *M. restricta* could singly exacerbate colitis in mouse models. More recently, pancreatic tissues associated with pancreatic ductal adenocarcinoma were found to have microbiota distinct from the healthy ones, dominated by *M. globosa* and *M. restricta* (AYKUT *et al.* 2019). These organisms activate the complement cascade and pro-inflammatory response, exacerbating pancreatic cancers.

Association with animals and other sources

Many of the 18 *Malassezia* species described to date have been isolated from animals both from healthy and disease states. Common examples include cats, dogs, horses, rabbits, pigs, cattle, goats, and others from which several *Malassezia* species have been isolated, without any host specificity (CABANES 2014; THEELEN *et al.* 2018). *M. psittaci* and *M. brasiliensis* were isolated from parrots, and a new thermo-tolerant species *M. vespertilionis* was isolated from migrating bats (CABANES *et al.* 2016; LORCH *et al.* 2018). This suggests a specific association of *Malassezia* with mammalian hosts.

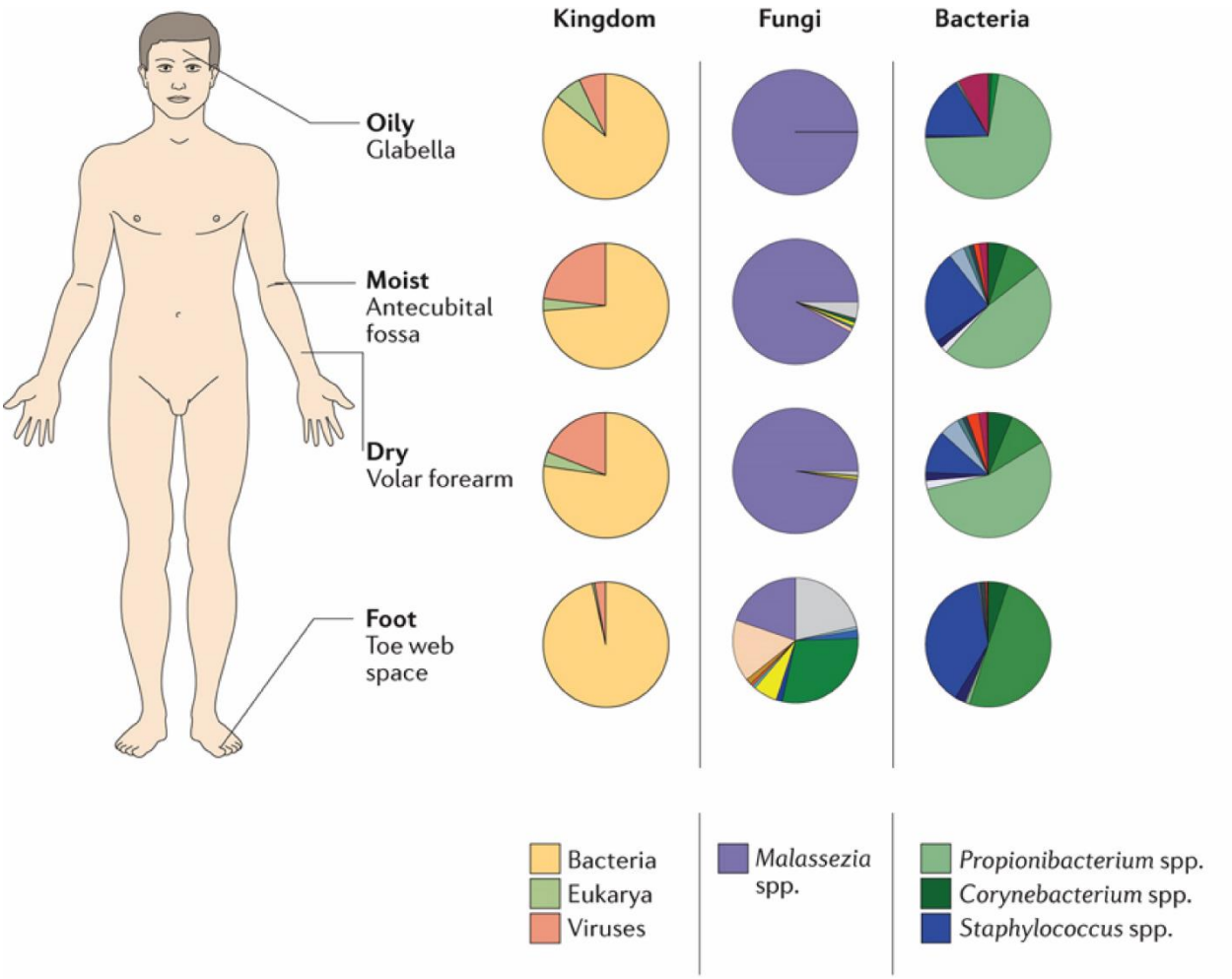


Figure 1-12. *Malassezia* on the human skin.
Pie charts indicating the relative consensus abundance of fungal and bacterial species across four major representative microenvironments in the human body (BYRD *et al.* 2018).

However, culture-independent metagenomic studies have detected sequences highly similar to the *Malassezia* species that infect humans, from extreme terrestrial and marine ecosystems such as hydrothermal vents, Antarctic soils, deep-sea sediments, and also in association with nematodes, sponges, and corals (ARENZ *et al.* 2006; LAI *et al.* 2007; LE CALVEZ *et al.* 2009; AMEND 2014). The prevalence of *Malassezia* species in such diverse host niches makes it an interesting system to understand the forces that shape their genomes to facilitate adaptation to these niches.

Genomic features

Most of the early sequencing efforts on *Malassezia* were restricted to regions that were routinely used for phylogenetic classification of the species (GUILLOT AND GUEHO 1995; GUEHO *et al.* 1996). The foundation for a genomics approach towards understanding the biology of these species was laid by a report in 2007 that described for the first time the features of dandruff associated *M. globosa* and *M. restricta* genomes (XU *et al.* 2007). This was followed by the analysis of the *M. sympodialis* genome predominantly associated with atopic eczema (GIOTI *et al.* 2013). Both studies estimated the genome sizes of the above species to be under 9 Mb, the smallest known for free-living fungi. A striking and defining feature of all the three species was the loss of genes encoding fatty acid synthases, and the amplification of genes coding for lipases, phospholipases, sphingomyelinases, and proteases. This observation was suggestive of a niche-specific evolution of these genomes, adapted to life on the skin where lipids are the primary source of energy. This also explained the requirement of supplementation of culture media with lipids and fatty acids to support the axenic growth of these species.

The above features were subsequently proved to be common to the *Malassezia* genus by a comparative genomics study that analyzed about 14 *Malassezia* species (WU *et al.* 2015). All of the *Malassezia* genomes analyzed were compact (<9 Mb) in line with the previous findings (Table 1-1). Functional analysis of these genomes identified 13 protein families/PFAM domains that are exclusive to *Malassezia* species and about 700 PFAM domains that were lost in the common ancestor. The *Malassezia* specific domains are yet to be characterized and are of unknown function. However, the set of genes lost encompasses a wide variety of functions, including carbohydrate metabolism, hydrolysis, and fatty acid synthases, supporting niche-specific evolution of these genomes. Besides the genes related to metabolic processes, the ones encoding the RNAi machinery were also lost in *Malassezia* species. This is supported by a significant reduction in the transposon density in these genomes.

The *Malassezia* genomes also show signatures of the presence of Horizontally Transferred Genes (HTGs) indicated by the presence of PFAMs absent in related fungi but detected in several bacterial species (WU *et al.* 2015; IANIRI *et al.* 2020). About six genes were proposed to be HTGs, predominantly associated with the oxidative stress response. Of these, three gene families were proposed to represent gain-of-function in *Malassezia* which include (i) a putative secreted glycosyl hydrolase (PF06742) that is detected in common pathogenic bacteria like *Mycobacterium tuberculosis* and *Listeria monocytogenes*, (ii) a bacterial catalase gene (PF00199), and (iii) a putative PrsW family protease (PF13367) that was detected in *Malassezia* species resident on human skin, which is required in bacteria and protists for response against anti-microbial peptides. While the expression of these genes is validated, their role in *Malassezia* biology will help understand better the evolution of these species.

Table 1-1. Genome sizes and chromosome number variations in the Malassezia species complex.

Reference indicates the study which reported these species. The number of scaffolds/contigs for each species is from the NCBI genomes portal. An asterisk indicates the genomes assembled as a part of this study. The chromosome number estimate provided is based on PFGE based karyotyping (references included in the text).

Species	Reference	Genome size	Genome assembly	Chromosome number estimate
M. cuniculi	(CABANES *et al.* 2011)	7.46 Mb	76 scaffolds	x
M. slooffiae	(GUEHO *et al.* 1996)	8.3 Mb	Complete*, 9 Chr	9
M. restricta	(GUEHO *et al.* 1996)	7.29 Mb	Complete, 9 Chr	9
M. globosa	(GUEHO *et al.* 1996)	9.1 Mb	Complete*, 9 Chr	x
M. pachydermatis	(WEIDMAN 1925)	8.15 Mb	91 contigs	6
M. equina	(CABANES *et al.* 2007)	7.65 Mb	117 scaffolds	x
M. dermatis	(SUGITA *et al.* 2002)	7.55 Mb	18 scaffolds	x
M. sympodialis	(SIMMONS AND GUEHO 1990)	7.75 Mb	Complete, 8 Chr	8
M. caprae	(CABANES *et al.* 2007)	7.58 Mb	229 scaffolds	x
M. nana	(HIRAI *et al.* 2004)	7.57 Mb	13 scaffolds	x
M. obtusa	(GUEHO *et al.* 1996)	7.84 Mb	1709 scaffolds	7
M. japonica	(SUGITA *et al.* 2003b)	8.3 Mb	16 scaffolds	x
M. yamatoensis	(SUGITA *et al.* 2004)	8.1 Mb	49 scaffolds	x
M. vespertilionis	(LORCH *et al.* 2018)	7.58 Mb	14 scaffolds	x
M. furfur (JLPK23)	(BAILLON 1889)	8.31 Mb	Complete*, 7 Chr	7
M. psittaci	(CABANES *et al.* 2016)	x	x	x
M. arunalokei	(HONNAVAR *et al.* 2016)	x	x	x
M. brasiliensis	(CABANES *et al.* 2016)	x	x	x

Chromosome number variations

Molecular characterization and genetic variations within different isolates of a species were studied mainly with techniques such as RAPD, RFLP, and AFLP in conjunction with PFGE analyses to

detect gross chromosomal changes (BOEKHOUT AND BOSBOOM 1994; BOEKHOUT *et al.* 1998; SENCZEK *et al.* 1999). This provided initial insights into the number of chromosomes present in several *Malassezia* species, and also revealed variations in the chromosome number between different *Malassezia* species that ranged between 6 and 10 (Table 1-1). The electrophoretic karyotyping analyses suggested that *M. pachydermatis* contained 6 chromosomes, *M. sympodialis* contained 8 chromosomes, and *M. slooffiae* along with *M. restricta* contained 9 chromosomes respectively. The estimates for *M. obtusa* and *M. globosa* were ambiguous, predicted to contain 7 and 8 chromosomes respectively. Analysis of *M. furfur* isolates revealed polymorphisms in both length and the number of chromosomes (BOEKHOUT AND BOSBOOM 1994; SENCZEK *et al.* 1999). This is supported by sequence-based studies that observed genome sizes for some *M. furfur* isolates that were double the size of other haploid *Malassezia* genomes, hinting at whole-genome duplication-like events or other hybridization events leading to heterogeneity (WU *et al.* 2015).

Recent efforts to develop whole-genome assemblies of these species with the help of long-reads based technologies (such as PacBio SMRT sequencing approach) helped in validating the chromosome number estimated previously. Such complete assemblies are presently available only for *M. restricta* (confirming the presence of 9 chromosomes), *M. sympodialis* (confirming the presence of 8 chromosomes), and *M. pachydermatis* (confirming the presence of 6 chromosomes)(KIM *et al.* 2018; CHO *et al.* 2019; MORAND *et al.* 2019). Besides these, the genomes of few other strains such as *M. nana*, *M. dermatis*, *M. vespertilionis*, and *M. japonica* available with the NCBI are assembled to less than 14 scaffolds, making them useful for comparative genomics and other studies including prediction of chromosome numbers. However, the cause and the role of such variations in the biology of these species remain elusive.

Genetic manipulation of *Malassezia*

Functional analyses of these candidate genes in *Malassezia* was hampered by the lack of genetic manipulation methods. This was ascribed mainly to its fastidious nature of growth *in vitro* and the failure of conventional methods used to engineer yeasts (IANIRI *et al.* 2016). An *Agrobacterium*-mediated transformation system could successfully integrate desired DNA fragments cloned within T-DNA border sequences into *Malassezia* (IANIRI *et al.* 2016; CELIS *et al.* 2017; IANIRI *et al.* 2017a). However, only three species, *viz*. *M. furfur*, *M. sympodialis*, and *M. pachydermatis* have been successfully engineered with this method. The advent of this method along with tools such as the development of CRISPR, *Agrobacterium*-mediated insertional mutagenesis, development of mouse models for infections enables functional genetics research in this medically relevant fungal species (AYKUT *et al.* 2019; IANIRI *et al.* 2019; LIMON *et al.* 2019; SPARBER *et al.* 2019; IANIRI *et al.* 2020).

Sexual cycle

The genomic locus that determines the sex/mating-type in fungi is called the *MAT* locus, containing genes or transcription factors that are required to regulate mating and post-mating developmental processes (NI *et al.* 2011). Mating is governed by two sets of mating-type specific genes- the first being the locus that encodes the pheromone or the cognate receptor (P/R) and the other encoding homeodomain transcription factors (HD) that regulate post-mating processes (CASSELTON AND OLESNICKY 1998; HSUEH AND HEITMAN 2008; SUN *et al.* 2019). Based on the arrangement of these two loci in the genome, two mating-type configurations were initially described for the basidiomycetes- bipolar and tetrapolar (BAKKEREN AND KRONSTAD 1994; HSUEH AND HEITMAN 2008). In the bipolar mating type, the regions containing the P/R and HD genes are present as a genetically linked single *MAT* locus. On the other hand, the two regions are genetically unlinked and segregate independently in mitosis in case of a tetrapolar configuration. Compatibility at both the loci is a prerequisite for mating in this case. A variation from these two configurations termed 'pseudo-bipolar' mating type was also identified within the basidiomycetes, wherein the P/R and HD loci appear to be linked like the case of bipolar systems; however, they exhibit features of tetrapolar systems such as multiallelism for HD locus and evidence for recombination between the P/R and HD loci (COELHO *et al.* 2010).

The *Malassezia* species were long thought to be devoid of a sexual cycle as it has not been observed to date. However, searches for circumstantial evidence for mating in these species shed light on the following features. The regions containing P/R and HD loci were found to be about 150 kb apart from each other in both *M. globosa* and *M. sympodialis* hinting at a possible bipolar mating system (XU *et al.* 2007; COELHO *et al.* 2013; GIOTI *et al.* 2013). However, sequencing several *M. sympodialis* isolates revealed a potential triallelism for HD locus and possible recombination between these loci. In line with these observations, a pseudo-bipolar mating configuration was assigned to *M. sympodialis*. Besides these, analysis of gene complement required for meiosis is also suggestive of a possible sexual cycle in the *Malassezia* species complex (COELHO *et al.* 2013; GIOTI *et al.* 2013). With the advent of genetic manipulation protocols, functional analysis of *MAT* loci, optimization of the right combination of strains, culture conditions to score for mating structures, will further our insights on the presence of an extant sexual cycle in *Malassezia*

Introduction to the model organism: *Candida albicans*

Phylum	: Ascomycota
Subphylum	: Saccharomycotina
Class	: Saccharomycetes

Order : Saccharomycetales
Family : *Debaromycetaceae*
Genus : *Candida*
Species : *albicans*

C. albicans is a diploid budding yeast belonging to the fungal phylum of Ascomycota. *C. albicans* and related *Candida* species constitute the CUG clade in which the tRNA corresponding to CUG codes for serine instead of leucine most of the time (MASSEY *et al.* 2003). It is known to be an opportunistic pathogen that survives as a harmless commensal in healthy individuals as a part of the microflora in the skin, gut, oral cavity, and urogenital tract (SOLL 2002). *C. albicans* is responsible for superficial infections like thrush, oropharyngeal and vaginal candidiasis, and systemic infections in hospitalized patients with debilitated immunity (SOLL 2002; BERMAN 2012). It is one of the most common fungal infections detected worldwide, with about 4 million cases reported every year (BROWN *et al.* 2012). *C. albicans* is known to be an obligate commensal of warm-blooded animals. However, a free-living form from oak trees was recently isolated as well (BENSASSON *et al.* 2019).

Phenotypic plasticity

C. albicans cells have the exceptional ability to switch between different morphogenetic states, and this ability has been attributed to successful colonization and infection (Figure 1-13) (MERSON-DAVIES AND ODDS 1989; SUDBERY *et al.* 2004). In routine laboratory growth conditions in rich media, *C. albicans* grows as an ellipsoid budding yeast. These cells are also referred to as white cells (described below in the context of mating). Under specific conditions, these yeast cells undergo a transition to true hyphal cells or pseudohyphal cells, each with distinct modes of division (BERMAN 2006). Pseudohyphal cells undergo cell cycle similar to the yeast cells. These cells exhibit unipolar budding and do not separate following cytokinesis resulting in a chain-like appearance. True hyphae result when the cells are grown at 37˚C in the presence of external stimuli such as the presence of serum in growth media. The polarized germ tube is a characteristic feature of hyphal cells. During division, the nuclear mass migrates to and divides across the preseptum. Several factors regulate the switching and growth in these forms (BERMAN 2006). Besides pseudohyphae and true hyphae, another polarized growth form called the elongated large budded cells exists in *C. albicans*. This phenotype is consequential of persistent arrest at any stage of the cell cycle or due to conditions of stress (BACHEWICH *et al.* 2003). Another yeast form called the ‘opaque’ form was first identified in the strain WO-1. The opaque form is the mating-competent form of yeast cells and is characterized by large oblong cells with pimples in the cell surface (MILLER AND JOHNSON 2002). The phenotypic switching to opaque cells is favored by conditions such as the presence of N-acetyl

glucosamine, CO_2, and acidic pH. Besides the opaque form, a gray phenotype was also reported to be an intermediate mating competent form of yeast cells (TAO *et al.* 2014). Together, the three yeast forms were proposed to be a part of a tristable phenotypic switching system playing critical roles in commensal and pathogenic traits. Passage of wild-type *C. albicans* cells through the mammalian gut resulted in the transition to another phenotypic form termed as the GUT (gastrointestinally induced transition) phenotype, which was transcriptionally poised to survive as a commensal in the gastrointestinal tract (PANDE *et al.* 2013).

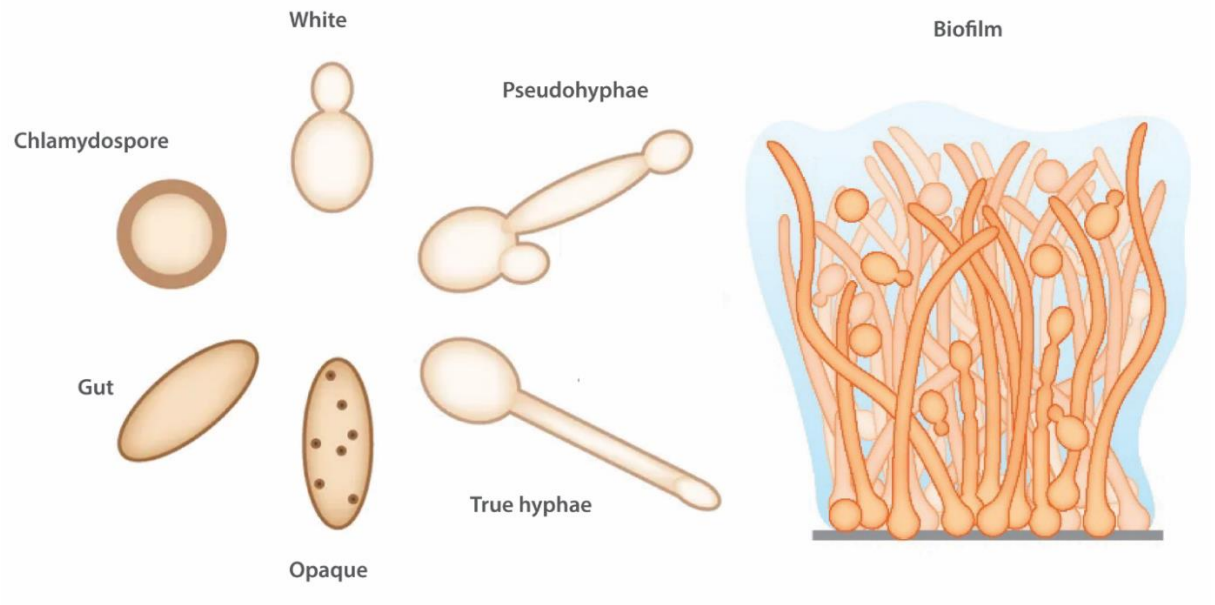


Figure 1-13. Morphological forms of *C. albicans*.
(A) White cells are essential for establishing systemic infections and dissemination. Elongated forms of pseudohyphae and true hyphae enable tissue penetration and evasion from the host immune system. Opaque cells are the mating competent yeast form of *C. albicans*. The gut form resembles opaque cells in shape and is adapted for commensal growth in the gastrointestinal tract. Chlamydospores arise in response to stress, but the role *in vivo* is unclear. (B) Schematic representation of a biofilm structure comprising of yeast and elongated cells along with an extracellular matrix. [figures adapted from (GOW 2013; NOBILE AND JOHNSON 2015)]

In response to harsh environmental conditions, *C. albicans* cells switch to the chlamydospore form, which are large spherical cells with a thickened cell wall (WHITEWAY AND BACHEWICH 2007). Besides these solitary growth forms, *C. albicans* cells can also transit to community-dwelling by the formation of biofilm structures (NOBILE AND JOHNSON 2015). These structures consist of several forms of *C. albicans* cells encapsulated in an extracellular matrix. These biofilms show drug resistance and can cause secondary seeding of the infection by dispersing cells from the biofilms, making it difficult to treat.

Genome plasticity

The size of the diploid *C. albicans* genome is about 32 Mb, distributed into 8 pairs of chromosomes. The observed variations in the genome include copy number variations, SNPs, segmental, or whole chromosome ploidy changes, each of which contributes to fitness advantage in diverse stress conditions. About 30% of clinical isolates show altered electrophoretic mobility indicating the presence of gross chromosomal changes in the genome (ASAKURA *et al.* 1991). Different forms of repetitive loci in the genome are thought to contribute to the genomic variations. Major Repeat Sequence present at least in part in each chromosome consists of RPS repeats flanked by HOK and RPB-2 elements (CHIBANA *et al.* 1994; LEPHART AND MAGEE 2006). These sequences mediate inter and intra-locus rearrangements generating polymorphisms in chromosome length and chimeric chromosome arms (SELMECKI *et al.* 2010). The 23 bp telomeric repeats and the telomere-associated gene families present in the subtelomeric regions can mediate telomere-telomere fusions. Centromeres in three chromosomes are flanked by inverted repeat structures (SANYAL *et al.* 2004). In the event of breakage within this region, mechanisms like break-induced replication can result in the formation of isochromosomes (discussed below), a form of trisomy (SELMECKI *et al.* 2010). Besides these, the *C. albicans* genome also contains LTRs and retrotransposons, which show variations in the copy number across different isolates (GOODWIN AND POULTER 2000; TODD *et al.* 2019). Short tandem repeats like trinucleotide repeats have also been detected in ORFs involved in traits such as commensalism and virulence wherein they can contribute to phenotype plasticity by introducing errors in alleles due to recombination or slippage between the repeat units (BRAUN *et al.* 2005).

Stress conditions are known to augment the genetic variations in the *C. albicans* genome, as evidenced by the loss of heterozygosity and aneuploidy (FORCHE *et al.* 2011). These LOH events are primarily effected by non-disjunction or by recombination/repair pathways. A classic example is the formation of isochromosome 5L or trisomy chromosome 4 upon exposure to fluconazole (SELMECKI *et al.* 2006; ANDERSON *et al.* 2017). Amplification of *ERG11* and *TAC1* as a consequence of this phenomenon is said to impart fluconazole resistance (SELMECKI *et al.* 2008; SELMECKI *et al.* 2009). Similarly, growth on media containing L- sorbose or D- arabinose as the sole carbon source is known to result in loss of a homolog of chromosome 5 and gain of chromosome 6 respectively (JANBON *et al.* 1998).

Mating

Analysis of the genome of *C. albicans*, thought to be devoid of sex for a long time (GRASER *et al.* 1996; BOUGNOUX *et al.* 2008), suggests that this organism is competent for mating with a conserved

mating-type like locus (*MTL* locus) encoding for transcription factors or regulators of mating specific genes (HULL AND JOHNSON 1999; MAGEE AND MAGEE 2000; BUTLER *et al.* 2009). However, no meiosis has been observed experimentally or in nature. Homologs of genes like Ime3 and Spo13 that are critical for the completion of Meiosis 1 are yet to be detected in this organism (BUTLER *et al.* 2009). Divergence in the function of meiosis-specific genes has also been proposed. For instance, a meiosis-specific transcription factor Ndt80 is involved in regulating the biofilm gene circuit in *C. albicans* (NOBILE AND JOHNSON 2015). Another such gene product Ume6 is implicated in autophagy and hyphal growth regulation in *C. albicans* (BANERJEE *et al.* 2008).

C. albicans is well-known for an alternative mode of mating, the parasexual cycle (BENNETT AND JOHNSON 2003; LACHKE *et al.* 2003; DUMITRU *et al.* 2007). Strains with homozygosity in the MTL locus were shown to mate with those containing the opposite mating type to generate tetraploid progeny, which undergoes a form of ploidy reduction called concerted chromosome loss. These two steps together constitute the parasexual cycle (BENNETT 2015). Serval aneuploidies, rearrangements, and recombination between homologous chromosomes have been reported to be a consequence of the parasexual cycle and is proposed to confer fitness advantages in diverse niches (FORCHE *et al.* 2008; HICKMAN *et al.* 2015; BERMAN 2016).

Centromeres structure and regulation

The centromeres of *C. albicans* is one of the first examples of epigenetically short regional centromeres. The CENP-A^{Cse4} enriched regions span about 3-4.5 kb in length that maps to an ORF free regions of up to 14 kb in the genome (SANYAL *et al.* 2004). Unlike the stabilization of ARS plasmids upon introduction of centromere DNA sequences observed in *S. cerevisiae* or *S. pombe*, no functional significance could be attributed to the underlying DNA sequence (BAUM *et al.* 2006). The requirement of a preexisting centromere-associated CENP-A^{Cse4} for centromere function is suggestive of epigenetic regulation of centromere function in this species. The clustered nature of centromeres in this organism and the resulting CENP-A cloud at the vicinity was proposed to make the centromere-proximal regions as favorable sites for neocentromere formation (THAKUR AND SANYAL 2013; SREEKUMAR *et al.* 2019a). It was also observed that the activated neocentromere subsequently reorganized in space to be a part of the centromere cluster in this species (BURRACK *et al.* 2016). The centromere identity is maintained across cell cycles by a concerted effort of the CENP-A chaperone HJURP^{Scm3} and the members of replication/repair machinery such as Orc4, Mcm2, and the Rad51-Rad52 complex (MITRA *et al.* 2014; SREEKUMAR *et al.* 2019b). The loss of function of any of the above components resulted in reduced CENP-A levels at the centromere cluster.

With a cryptic RNAi pathway and a lack of conventional heterochromatin marks, the pericentromeres were poorly defined in this organism. The centromere flanking regions show features of both euchromatin and heterochromatin (FREIRE-BENEITEZ *et al.* 2016). By combining the results from Hi-C and reversible silencing assays, the pericentromere was defined based on the presence of CEN-flanking compact chromatin that spans 25 kb clustered around the centromere (SREEKUMAR *et al.* 2019a). The presence of enhanced cis- and trans- chromosomal contacts segregated this region from the bulk chromatin.

Kinetochore assembly

Similar to the observations in the budding yeast *S. cerevisiae*, the kinetochore proteins are localized in a clustered manner throughout the cell cycle and are constitutively associated with the microtubules (SANYAL AND CARBON 2002; ROY *et al.* 2011; THAKUR AND SANYAL 2011). The homologs of the core kinetochore ensemble found in *S. cerevisiae* remains conserved in *C. albicans,* barring the CBF3 complex. In both the budding and fission yeasts, a master regulator for the assembly of kinetochore subcomplexes has been identified- Ndc10 in *S cerevisiae* and Mis6 in *S. pombe* (TAKAHASHI et al. 2000; HAYASHI et al. 2004; CAMAHORT et al. 2007; CHO AND HARRISON 2011*)*. These proteins are known to dictate the localization of CENP-A^{Cse4} and all other kinetochore proteins. On the contrary, an interdependent regulation of kinetochore protein localization was observed in *C. albicans* with no single protein forming the base of localization hierarchy (THAKUR AND SANYAL 2012). While the inner kinetochore proteins are essential for the recruitment of the outer kinetochore subunits, the localization of outer kinetochore proteins was also found essential for the localization of inner kinetochore proteins including CENP-A^{Cse4}. Further, perturbation of any of the kinetochore subunits resulted in the disintegration of the kinetochore ensemble, followed by proteasome-mediated degradation of CENP-A^{Cse4} (Figure 1-14).

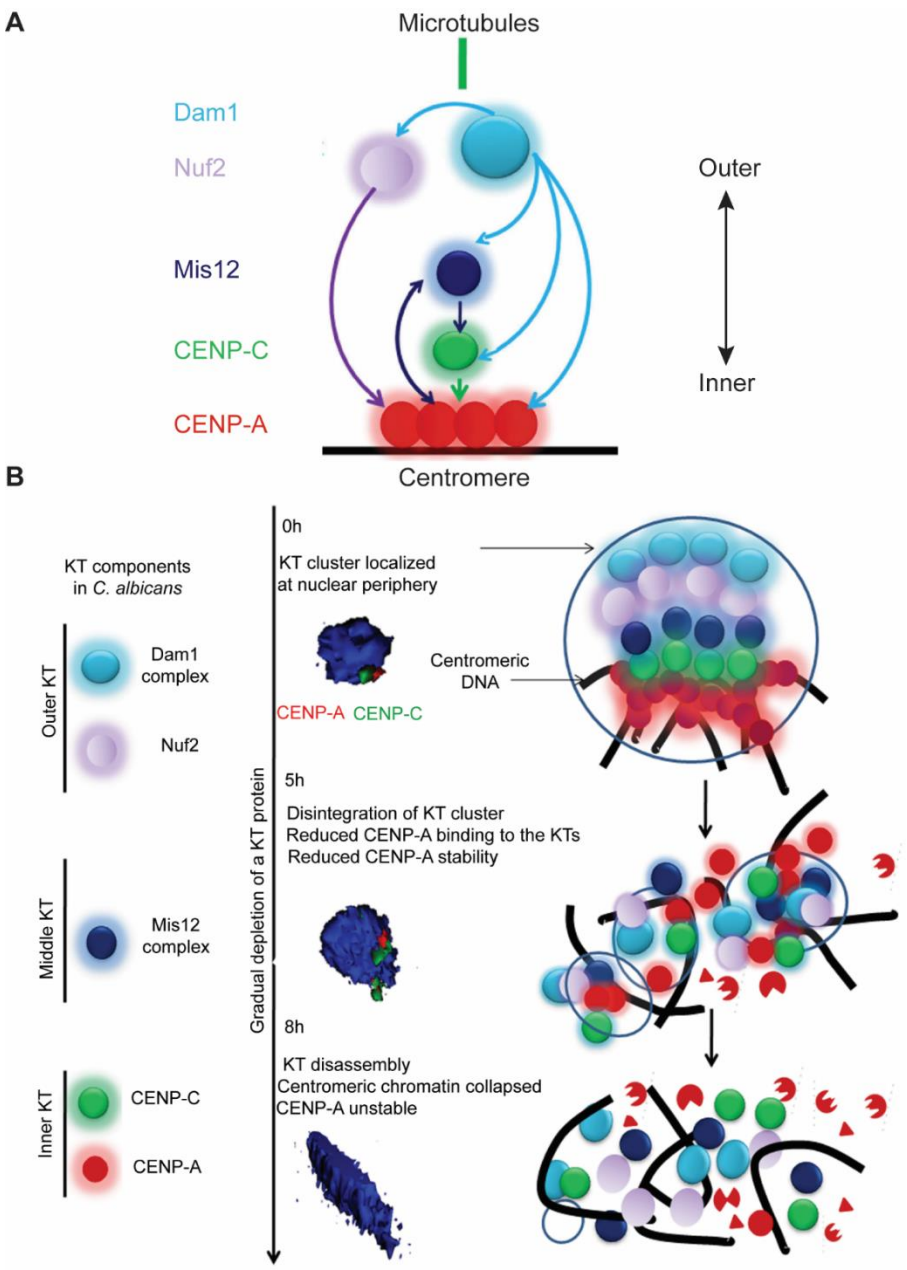


Figure 1-14. Regulation of kinetochore protein localization in *C. albicans*.
(A) Schematic representation of localization dependencies of inner and outer kinetochore proteins as reported in *C. albicans*. (B) Depletion of any of the proteins from the inner, middle or outer kinetochore results in declustering of the kinetochores followed by disassembly of the kinetochore ensemble. Disassembled kinetochores fail to protect the centromeric chromatin. CENP-A^{Cse4} molecules that are no longer assembled into centromeric chromatin eventually get degraded by the proteasomal mediated pathway (THAKUR AND SANYAL 2012).

The core kinetochore proteins- CENP-A^{Cse4}, Mif2, Mtw1, and the members of the Dam1 complex, have been found essential for viability and chromosome segregation (SANYAL AND CARBON 2002; BURRACK *et al.* 2011; ROY *et al.* 2011; THAKUR AND SANYAL 2011). The depletion of any of these proteins resulted in mitotic arrest consequential an active SAC. Mutants show aberrant spindle structures such as short spindles in most cases with hyper elongated spindles in a smaller proportion in specific mutants. Despite harboring a larger CENP-A^{Cse4} enriched region as compared to *S. cerevisiae*, both these species have one microtubule per kinetochore (JOGLEKAR *et al.* 2008). The essentiality of the Dam1 complex in *C. albicans* was found reduced when multiple microtubule associations per kinetochore were promoted, suggesting that the essentiality of Dam1 complex is correlated with one MT/kinetochore association (BURRACK *et al.* 2011; THAKUR AND SANYAL 2011).

Rationale and objectives

The primary objective of the first half of this study was towards understanding the mechanisms that could result in the observed karyotypic diversity within the members of the *Malassezia* species complex.

It is known that the genomes of all known *Malassezia* species are under 9 Mb, one of the smallest known for free-living fungi, and are distributed into 6 to 9 chromosomes in the known 18 species of the complex (BOEKHOUT *et al.* 1998; SENCZEK *et al.* 1999; WU *et al.* 2015). The *Malassezia* yeasts are ubiquitous skin commensals that have been isolated from diverse animals ranging from humans to bats (CABANES 2014; FINDLEY AND GRICE 2014; THEELEN *et al.* 2018). Besides the human skin, *Malassezia* species were also found to exacerbate disease conditions like Crohn's disease and pancreatic cancers (AYKUT *et al.* 2019; LIMON *et al.* 2019). Further, sequences closely related to the *Malassezia* species have been detected from extreme environmental niches such as corals, hydrothermal vents, and deep-sea sediments(AMEND 2014). Genomic rearrangements are known to facilitate the adaptive evolution of species, enabling them to thrive in diverse conditions (EICHLER AND SANKOFF 2003; GUSA AND JINKS-ROBERTSON 2019). The significance of such rearrangements is profound in organisms that have no observed or a rare sexual cycle wherein they can also aid in speciation (NOOR *et al.* 2001; RIESEBERG 2001). While the analysis of *Malassezia* genomes indicate that these species are mating competent, no mating has been observed in nature or laboratory conditions (XU *et al.* 2007; COELHO *et al.* 2013; GIOTI *et al.* 2013; ZHU *et al.* 2017). This makes it important to understand the mechanisms that result in genetic diversity (chromosome number variation in this case) within these species.

The most commonly observed mode of chromosome number variation is the telomere-telomere fusion of chromosomes. Well known examples include the formation of extant human chromosome 2 by fusion of two ancestral chromosomes (YUNIS AND PRAKASH 1982; JDO *et al.* 1991) and the reduction in karyotype observed within the members of the Saccharomycotina in the case of fungi (GORDON *et al.* 2011). Breakage at the centromere has been suspected of having driven the evolution of *Ashbya gossypii* from the pre-WGD ancestor of *S. cerevisiae* (GORDON *et al.* 2011). Insertion of an entire chromosome into another, termed nested chromosome insertions are commonly observed in plants(MURAT *et al.* 2010).

To identify if any of the above or alternate mechanisms resulted in rearrangements and chromosome number variations in the *Malassezia* species, we sought to identify centromeres in species with different numbers of chromosomes and trace the evolutionary trajectory of changes in the karyotype. The results pertaining to these objectives are presented in Chapter 2 of this thesis.

In the following part of the thesis, we studied the role of Dad2, a Dam1 complex subunit, in mediating interactions with microtubules in human fungal pathogen *C. albicans*.

The outer kinetochore, comprising the KMN (KNL1C-MIS12C-NDC80C) network, forms the primary microtubule-binding module of the kinetochore ensemble (CHEESEMAN *et al.* 2006; CHEESEMAN AND DESAI 2008). The efficiency of this association is further enhanced by another outer kinetochore complex called the Dam1 complex (LAMPERT *et al.* 2010; TIEN *et al.* 2010). The Dam1 complex is detected exclusively in fungi and is replaced by a functional homolog called the Ska1 complex in higher eukaryotes (VAN HOOFF *et al.* 2017b). The Dam1 complex subunits were shown essential for viability in many pathogenic fungi such as *C. albicans*, *C. tropicalis*, and *C. neoformans*, as the conditional mutants show severe viability defects and massive chromosome missegregation (THAKUR AND SANYAL 2011; CHATTERJEE *et al.* 2016; SRIDHAR *et al.* 2019). This makes the Dam1 complex subunits an attractive target for the development of safe and potent antifungals.

Dad2, a subunit of the Dam1 complex, is essential for viability in *C. albicans*, with the mutants showing elevated levels of chromosome missegregation (THAKUR AND SANYAL 2011). Multiple sequence alignment of various Dad2 sequences led us to identify a conserved 10-amino acid long stretch in the C-terminus that we labeled as the Dad2 signature sequence (DSS). The presence of conserved positively charged residues in this domain is reminiscent of the microtubule-interacting domains of other kinetochore proteins like Ndc80, Dam1, and Ska1 (CHEESEMAN *et al.* 2002; CHEESEMAN *et al.* 2006; ABAD *et al.* 2014). We surmise that the DSS could be a previously unknown microtubule-binding motif in Dad2/Dam1 complex.

We aimed to dissect the role of this conserved 10-amino acid stretch in Dad2, in mediating kinetochore-microtubule attachments in *C. albicans.* A brief introduction to how kinetochore subcomplexes cooperatively associate with the microtubule attachments is provided at the beginning of Chapter 3, followed by our findings related to this objective.

Chapter 2

Loss of centromere function drives karyotype evolution in closely related *Malassezia* species

Results

Chromosome number varies in the *Malassezia* species complex

Previous reports based on pulsed-field gel electrophoresis (PFGE) have suggested that chromosome number varies within the *Malassezia* species complex. The early diverged species *M. slooffiae* of Clade C was reported to have 9 chromosomes (BOEKHOUT *et al.* 1998). Clade B *Malassezia* species are reported to have 9 (*M. globosa* and *M. restricta)*, 8 (*M. sympodialis*), or 6 chromosomes (*M. pachydermatis*). Among the Clade A species, *M. obtusa* and *M. furfur* CBS14141 were reported to have 7 chromosomes each (BOEKHOUT AND BOSBOOM 1994; ZHU *et al.* 2017). A high-quality reference genome is a prerequisite to understanding the rearrangements associated with chromosome number variation. Additionally, this will also assist in resolving ambiguities in PFGE-based estimates of chromosome number when similar-sized chromosomes are present. Complete genome assemblies were not available for many of the species with reported numbers of chromosomes. To obtain better-assembled reference genomes, we sequenced the genomes of *M. slooffiae* and *M. globosa* as representatives of the 9-chromosome state, and *M. furfur* as a representative of the 7-chromosome state using the PacBio SMRT sequencing technology.

The *M. globosa* genome was completely assembled into 9 contigs with telomeres on both ends (BioSample accession SAMN10720087). We validated these by assigning each band on the pulsed-field gel with the sizes from the genome assembly and further confirmed these by chromoblot analysis following PFGE. This analysis shows that chromosome 5 contains the rDNA locus and migrates higher than the expected size of 902 kb, as a diffuse ensemble of different sizes along with chromosome 3 (Figure 2-1A). The assembled genome of *M. slooffiae* has 14 contigs of which 9 contigs have telomeres on both ends, indicative of 9 chromosomes (BioSample accession SAMN10720088). Each of the 9 contigs could be assigned to the bands observed in the pulsed-field gel (Figure 2-1B). For *M. furfur*, the final genome assembly consisted of 7 contigs with telomeres on both ends and matched the expected chromosome sizes obtained from an earlier PFGE analysis of CBS14141 (Figure 2-1C). The complete genome assembly of *M. sympodialis* reported earlier is distributed into 8 chromosomes with telomere repeats on both ends (Figure 2-1D) and serves as a representative of an 8-chromosome state in this study. The assembly statistics are described in Table 2-1.

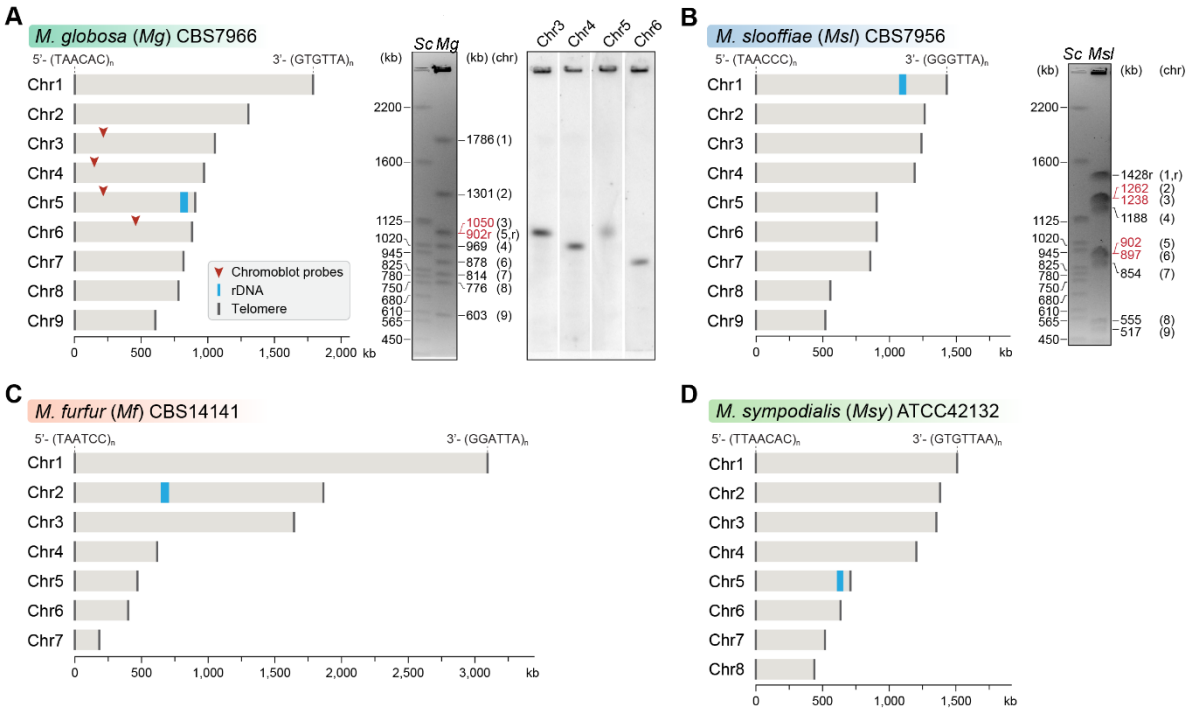


Figure 2-1. Genome assembly and karyotype diversity in representative *Malassezia* species. The genomes of *M. globosa* (A), *M. slooffiae* (B), and *M. furfur* (C) were sequenced and assembled in this study, while the genome assembly of *M. sympodialis* (D) was reported earlier (Zhu et al., 2015) and is shown for comparison. In each panel, bar plots represent the assembled chromosomes of the indicated *Malassezia* species, with the telomeres and the ribosomal DNA (rDNA) marked as dark grey and blue bars, respectively. Telomere-repeat motifs are shown at the 5'- and 3'-ends of the Chr1 in each species. Electrophoretic karyotypes of *M. globosa* (Mg) and *M. sloffiiae* (Msl), are shown in (A) and (B) respectively, with chromosome sizes estimated from the genome assembly. Chromosomes of *S. cerevisiae* (Sc) served as size markers. The chromosome containing the rDNA (marked with an "r"), in *M. globosa*, co-migrates with Chr3. This was assessed by chromoblot hybridization using unique sequences from Chr3, Chr4, Chr5, and Chr6 as probes (regions indicated by red arrowheads). Chromosomes of similar size (denoted in red) migrate together in the gel and appear as a doublet band (i.e., MgChr3-MgChr5, MslChr2-MslChr3, and MslChr5-MslChr6).

The genome sequencing and assembly were obtained from Tom Dawson as a part of a collaboration for this project. The chromoblot analysis to validate the chromosome number was performed in collaboration with Joe Heitman's Lab (by Sheng Sun, Duke University).

Changes in chromosome number are always associated with birth or loss of centromeres to stabilize the karyotype in organisms with monocentric chromosomes. To understand the transitions between these different karyotypic states observed in the *Malassezia* species complex, we sought to validate centromeres in species representative of each karyotype experimentally.

Table 2-1. Statistics of the genome assemblies of *M. globosa*, *M. slooffiae*, and *M. furfur* generated in this study.

	M. globosa	*M. slooffiae*	*M. furfur*
Strain	CBS7966	CBS7956	CBS14141
Sequencing platform	PacBio	PacBio	PacBio
No. of reads	122,622	92,622	57,425
Coverage	90X	140X	64X
Size (bp)	9,119,538	8,938,001	8,311,773
GC content (%)	52.09	66.31	64.9
Assembly status			
Total contigs	10	14	8
Nuclear genome	**Complete** 9 contigs telomere on both ends	**Complete** 9 contigs telomere on both ends	**Complete** 7 contigs telomere on both ends
Gaps	none	none	none
Assembly polished with Illumina reads	No	No	No

Identification of kinetochore protein homologs in *M. sympodialis*

Kinetochore proteins, by their exclusive centromere localization, serve as useful tools in the identification of centromere DNA sequences. To experimentally identify centromeres in this species, we looked for conserved kinetochore proteins in *M. sympodialis* by BLAST analysis. Protein sequences of homologs from *C. neoformans* were used as the query for most proteins unless stated otherwise. Schematic representation of the conserved kinetochore proteins we detected in *M. sympodialis* is shown (Figure 2-2A, Table 2-2).

Point centromere-specific proteins viz. Ndc10, Cep3, and Ctf13 of the Cbf3 complex initially identified in *S. cerevisiae* were not detected in *M. sympodialis*. Their absence suggested that despite small compact genomes, it was less likely to possess point centromeres. However, we could detect homologs of CENP-A, CENP-C, and most of the outer kinetochore proteins of the KMN network in the genome of *M. sympodialis.* Similar to the observation in *C. neoformans*, no proteins from the **c**onstitutive **c**entromere **a**ssociated **n**etwork (CCAN) were detected in this species as well (Table 2-2).

Table 2-2. Identification of kinetochore proteins in *M. sympodialis* by BLAST.
Sequences of *C. neoformans* homologs were used as the query for BLAST. Asterisk (*) indicates cases where sequences from *Ustilago maydis* homologs were used as query. For the CBF3 complex, sequences of *S. cerevisiae* homologs were used as the query. ND- not detected

	Sub Complex	Query	Protein ID (SHOxxxxx.x)	% Identity	E- value	Query coverage
Outer kinetochore	Dam1 complex	Ask1	79565.1	27.4	4.84e-22	65.64
		Dad1	75772.1	47.1	4.44e-01	27.36
		Dad2*	79986.1	40	9.9e-23	82.89
		Dad3	77770.1	42.5	3.62e-10	78.75
		Dad4	79701.1	39.1	1.24e-09	95.83
		Dam1	79667.1	43.7	6.44e-08	45.45
		Duo1	76611.1	27.4	1.19e-04	17.63
		Spc19	N.D.			
		Spc34	N.D.			
		Hsk3*	77065.1	41.9	1.6e-15	92.68
	Ndc80 complex	Ndc80	79901.1	25.6	1.15e-36	80.33
		Nuf2	78898.1	25.8	7.64e-14	74.59
		Spc24	N.D.			
		Spc25	78263.1	25.8	1.1e-07	42.18
	Mtw1 complex	Mis12	76526.1	25.1	5.67e-09	70.91
		Dsn1	78762.1	24.4	3.94e-08	24.05
		Nnf1*	77715.1	40.8	2.10E-19	35.9
		Nsl1*	79581.1	28.3	1.68E-06	42
	KNL1	Spc105	75936.1	19.9	5.36e-11	40.59
Inner kinetochore	Constitutive Centromere Associated Network (CCAN)	Cnn1/Wip1/ Mhf1/Mhf2	N.D.			
		Mcm16/Mcm22/ Ctf3	ND.			
		Okp1/Ame1/ Ctf19/Mcm21	ND.			
		Chl4/Iml3	ND.			
	CENP-C	Mif2	79930.1	35.8	5.77e-29	22.67
	CENP-A	Cse4	76408.1	70.8	2.77e-46	66.67
Point CEN specific complex	Cbf3 complex	Ndc10	ND.			
		Cep16	N.D.			
		Ctf3	N.D.			

While CENP-A^{Cse4} is an ideal candidate to identify centromeres in an organism, it is often rendered non-functional upon epitope tagging. To avoid such technical shortcomings, especially in a haploid species like *M, sympodialis*, we resorted to tagging the Mis12 homolog Mtw1. The Mis12 complex subunits are evolutionarily conserved kinetochore proteins that link the chromatin-associated inner kinetochore proteins to the microtubule-associated outer kinetochore proteins. Members of the Mis12 complex localize to centromeres in many yeasts (GOSHIMA *et al.* 1999; GOSHIMA *et al.* 2003; WESTERMANN *et al.* 2003; ROY *et al.* 2011). Recent studies suggest that the protein domains associated with the Mis12 complex members are exclusive to kinetochore proteins and are not detected in any other proteins making them attractive tools to identify centromere sequences (TROMER *et al.* 2019).

We functionally expressed an N-terminally GFP-tagged Mtw1 protein (Protein ID: SHO76526.1) from its native promoter, and the expression of the fusion protein was confirmed by western blotting (Figure 2-2B). Upon staining with anti-GFP antibodies and DAPI (4′,6-diamidino-2-phenylindole), we could detect punctate localization of Mtw1 at the nuclear periphery (Figure 2-2C) consistent with the clustered kinetochore localization observed in other yeasts (GOSHIMA *et al.* 1999; EUSKIRCHEN 2002; ROY *et al.* 2011). Live-cell images of MSY001 (GFP-*MTW1*) cells revealed that the kinetochores (GFP-Mtw1) remained clustered throughout the cell cycle, starting from unbudded G1 cells in interphase to large budded cells in mitosis (Figure 2-2D). The GFP-Mtw1 tagged *M. sympodialis* strain was generated in Joe Heitman's lab (by Giuseppe Ianiri) as a part of our collaboration on this project.

Mtw1 is localized to a single region at the GC minima of each *M. sympodialis* chromosome

Having identified Mtw1 as an authentic kinetochore protein, we performed ChIP-sequencing using the GFP-Mtw1 expressing strain of *M. sympodialis* (MSY001). Mapping the reads to the reference genome of *M. sympodialis* strain ATCC42132 (ZHU *et al.* 2017) identified one significantly enriched locus on each of the eight chromosomes (Figure 2-3A). The length of the Mtw1-enriched centromere regions identified from the ChIP-seq analysis ranges between 3167 bp and 5143 bp with an average length of 4165 bp (Table 2-3). However, the region of maximum Mtw1 enrichment on each chromosome

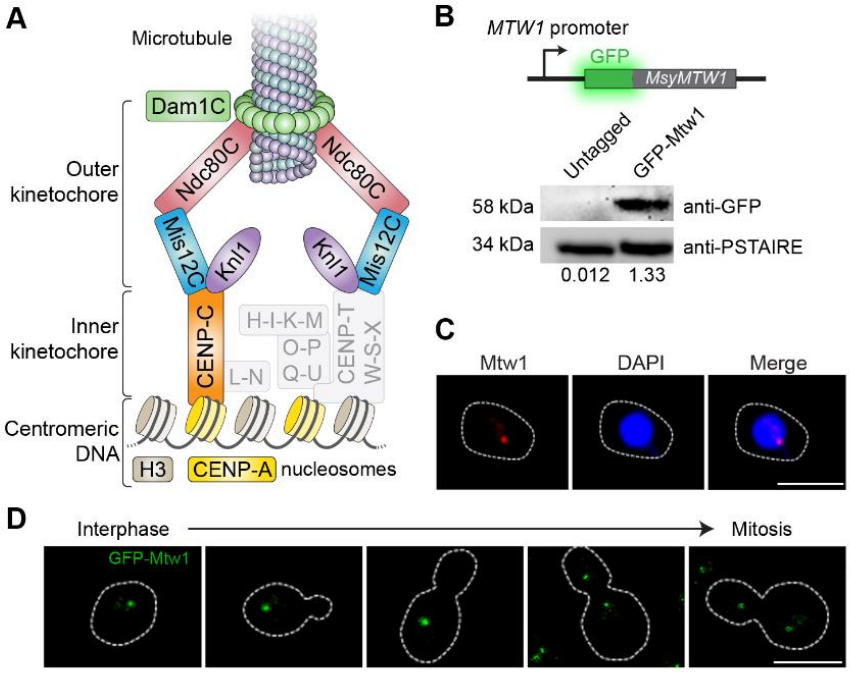


Figure 2-2. Kinetochores cluster and localize at the nuclear periphery in *M. sympodialis*.
(A) Schematic of the kinetochore organization of *M. sympodialis*. Gray boxes indicate proteins absent in *M. sympodialis*. The outer kinetochore protein Mtw1 (a component of Mis12C) served as the kinetochore marker in the present study. (B) Line diagram representation of Mtw1 tagged with GFP at the N-terminus. Immunoblot analysis of whole-cell lysates prepared from the untagged *M. sympodialis* strain (ATCC42132) and GFP-Mtw1 expressing cells (MSY001) probed with anti-GFP antibodies and anti-PSTAIRE antibodies. PSTAIRE served as a loading control. Relative intensity values normalized to PSTAIRE are indicated below each lane. (C) Logarithmically grown MSY001 cells expressing GFP-Mtw1 were fixed and stained with DAPI (blue) and anti-GFP antibodies (pseudo-colored in red). Scale bar, 2.5 µm. (D) Cell cycle stage-specific localization dynamics of GFP-Mtw1. Scale bar, 2.5 µm.

(based on the number of sequenced reads aligned) mapped to the intergenic region harboring the GC trough (approximately 1 kb long), which was previously predicted to be the centromeres of *M. sympodialis* (Figure 2-3B) (ZHU *et al.* 2017). The regions of Mtw1 enrichment span beyond the core centromeres and include active genes located proximal to these troughs (Figure 2-3B and 2-3C). However, these ORFs do not show consensus features such as the orientation of transcription or functional classification. We validated this enrichment by ChIP-qPCR analysis with primers homologous to the predicted centromeres compared to a control region distant from the centromere (Figure 2-4A).

Histone H3 is depleted at the core centromere with active genes at the pericentric regions in *M. sympodialis*

The presence of CENP-A nucleosomes at the kinetochore should result in reduced histone H3 enrichment at the centromeres when compared to a non-centromeric locus. To test this, we performed ChIP with anti-histone H3 antibodies and analyzed the immunoprecipitated (IP) DNA by qPCR. As compared to a control ORF region unlinked to the centromere (190 kb away from *CEN1*), the pericentric regions flanking the core centromere showed a marginal reduction in histone H3 enrichment that was further reduced at the core that maps to the GC trough with the highest enrichment of the kinetochore protein Mtw1. That the core centromere region showing the maximum depletion of histone H3 coincided with the regions most enriched with Mtw1 further supports that histone H3, in these regions, is possibly replaced by its centromere-specific variant CENP-A (Figure 2-4B).

Table 2-3. Coordinates of centromeres and their GC content in *M. sympodialis*.
Coordinates and length of Mtw1-enriched regions in comparison with that of the core centromeres in *M. sympodialis*. Genome average GC content (in %): 58.5

Chromosome number	Core centromere				Full-length centromere		
	Coordinates		Length (bp)	%GC	Coordinates		Length (bp)
	Start	End			Start	End	
1	786541	787061	520	16.4	784833	788599	3767
2	355760	355841	81	20	354218	357486	3269
3	237534	238686	1152	15.6	235615	239940	4326
4	418202	418728	526	15.2	415985	420656	4672
5	125056	125220	164	18	123219	127284	4066
6	101950	102502	552	14.4	100342	105251	4910
7	431542	431987	445	13.2	430028	433194	3167
8	24694	25564	870	18.4	22334	27476	5143

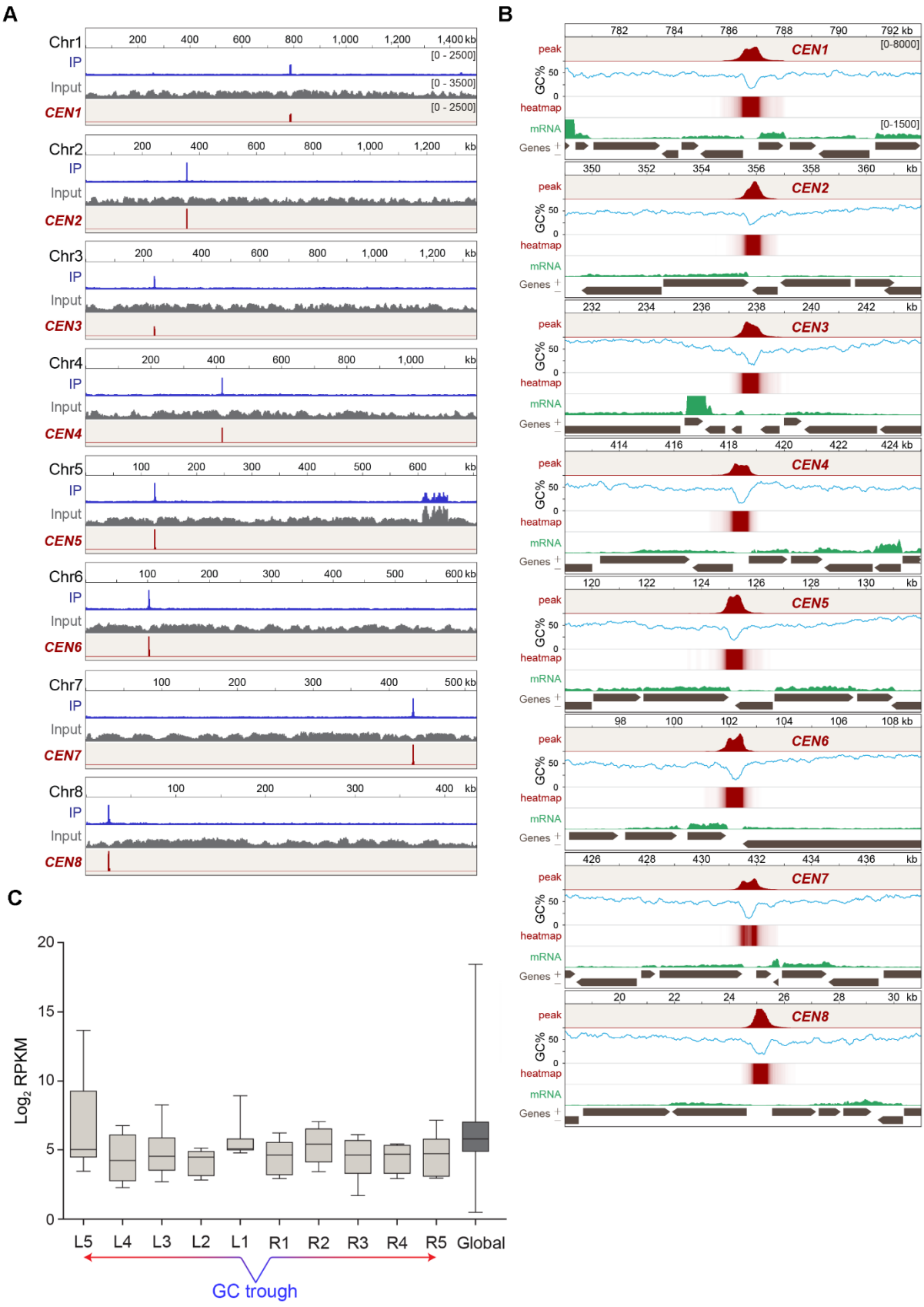

A
Chr1
IP
Input
CEN1
[0 - 2500]
[0 - 3500]
[0 - 2500]
Chr2
CEN2
Chr3
CEN3
Chr4
CEN4
Chr5
CEN5
Chr6
CEN6
Chr7
CEN7
Chr8
CEN8
B
peak
GC%
heatmap
mRNA
Genes
[0-8000]
[0-1500]
C
Log2 RPKM
L5
L4
L3
L2
L1
R1
R2
R3
R4
R5
Global
GC trough

Figure 2-3. Localization of Mtw1 to single-peaks identifies centromeres on each of the eight chromosomes of *M. sympodialis*.
(A) ChIP-seq reads obtained after immunoprecipitation with GFP-trap beads were mapped to each chromosome. The *x*-axis indicates chromosomal coordinates (in kb), and the *y*-axis indicates read depth. Input: reads from total DNA; IP: reads from the immunoprecipitated sample; *CEN*, Mtw1-enriched regions derived by subtracting input reads from those of the IP sample (peak values 0-2500). Additional peaks observed in both IP and input tracks on Chr5 are from the rDNA locus. (B) A 13 kb-window of Mtw1 enrichment profile (*CEN*, represented as peaks and heat-map in two different tracks, red) plotted along with the GC content (%GC, blue) and regions of transcription (RNA-seq, green). The numbers mentioned in the topmost track in every panel indicate chromosomal coordinates (in kb). The scales in the *y*-axis are as follows: *CEN* (0-8000), %GC (0-75), RNA-seq reads (0-1500). Gray arrows in each panel indicate the predicted ORFs based on RNA-seq data with arrowheads pointing the direction of transcription of the corresponding gene (also marked as '+' and '-' in the axis label). (C) Normalized read counts (RPKM) for transcripts from ORFs flanking the centromere as compared with the genome average RPKM value calculated from RNA-seq data of *M. sympodialis* (ATCC42132). The *x*-axis represents ORFs labeled L1 - L5 and R1 - R5, with L1 and R1 being proximal to the GC trough of the centromere. The genome average considered for comparison is labeled as global. The *y*-axis represents Log_2 RPKM values.

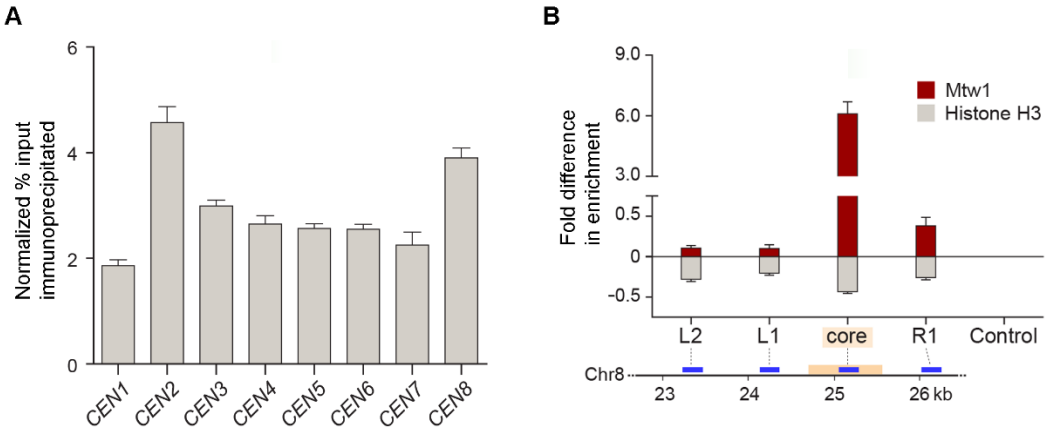
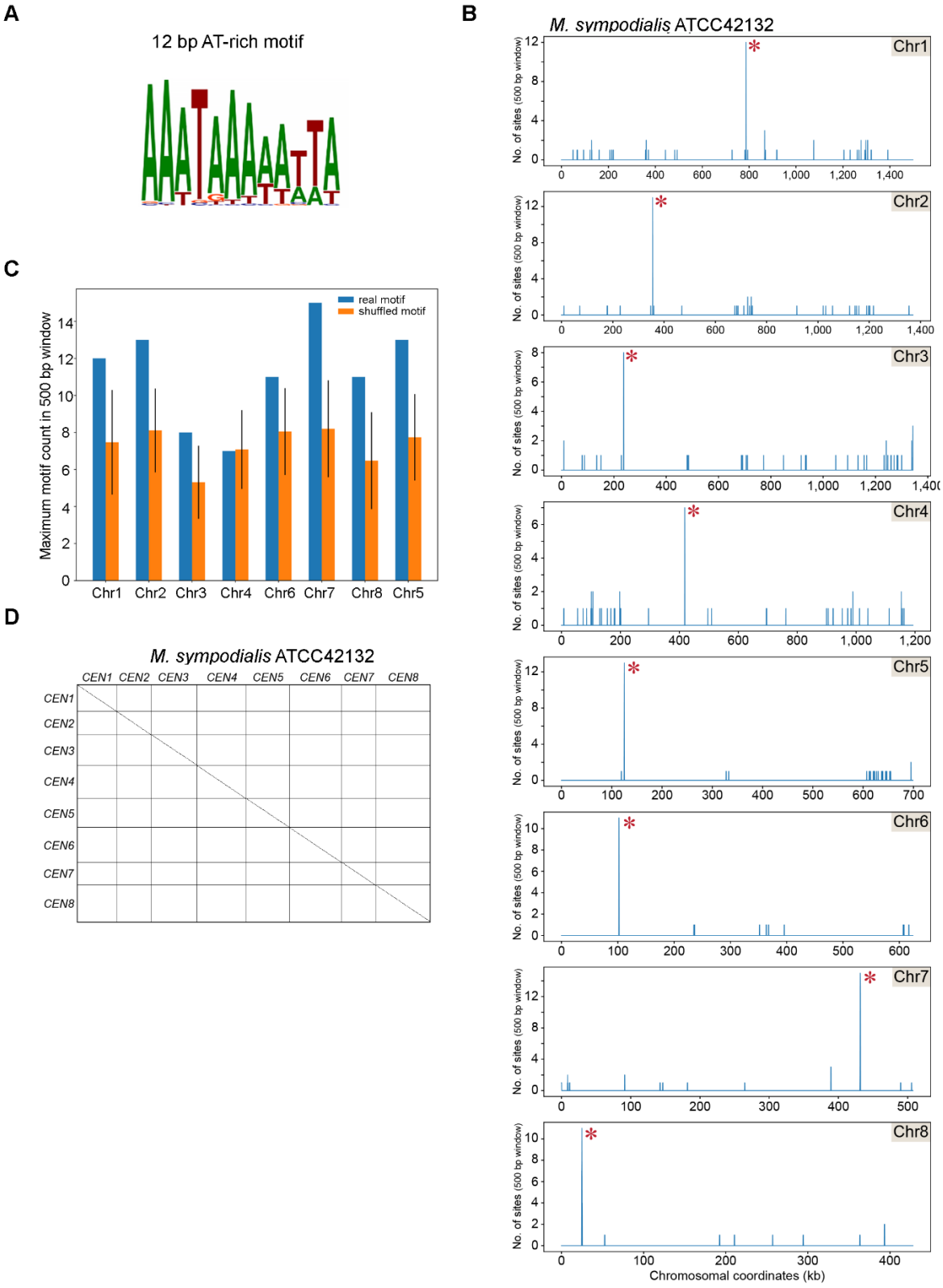


Figure 2-4. Mtw1-enriched regions in *M. sympodialis* show reduced histone H3 levels.
(A) ChIP-qPCR assays validating the enrichment of Mtw1 at the centromeres identified by ChIP-sequencing analysis. The *x*-axis indicates individual *CEN* regions assayed with primers mapping to the chromosomal GC troughs, and the *y*-axis indicates enrichment of Mtw1 over a centromere unlinked control region as "normalized % input immunoprecipitated". Values from three experiments, each performed with three technical replicates were used to generate the plot. Error bars indicate the standard deviation (SD). (B) Fold difference in Mtw1 and histone H3 enrichment at *CEN8* as compared to a non-centromeric control region (190 kb away on the right of *CEN1*) by qPCR analysis. Schematic of a 4 kb region of Chr8 with *CEN8* core (yellow box) is depicted below the graph. Blue lines indicate regions assayed by PCR: core-region corresponding to the GC trough, L1 and R1- 750 bp away from the core, L2- 1500 bp away from the core, and non-centromeric control region (190 kb away from centromere in Chr1). The *x*-axis indicates regions across the *CEN8* probed by PCR and *y*-axis indicates fold difference in the enrichment of Mtw1 and histone H3 as compared to the control region. Error bars indicate the standard deviation (SD). Values from three experiments, each performed with three technical replicates were used to generate the plot.

The short regional centromeres of *M. sympodialis* are enriched with a 12 bp long AT-rich consensus sequence motif

To understand the features of *M. sympodialis* centromeres, we analyzed the centromere DNA sequences for the presence of consensus motifs or structures such as inverted repeats. PhyloGibbs-MP (SIDDHARTHAN *et al.* 2005; SIDDHARTHAN 2008) predicted a 12 bp long AT-rich motif common to all of the centromere sequences of *M. sympodialis* (Figure 2-5A). We swept the Position Weight Matrix (PWM) from the PhyloGibbs-MP output across each chromosome of *M. sympodialis* and counted the number of motif predictions in sliding 500 bp window, sliding by 100 bp at a time. Sites with log-likelihood-ratio (LLR) of >7.5 were counted as motif predictions. The LLR is the natural logarithm of the ratio of the likelihood of the 12 bp substring arising as a sample from the PWM to the likelihood of it being generic "background". In each case, the global peak coincides with the centromere (Figure 2-5B). In each chromosome, the centromere region shows between 7 and 13 motif matches, while no other 500 bp window shows more than 3 matches. This suggests that the AT-rich motif is more enriched at the centromeres than at any other region in the *M. sympodialis* genome (Figure 2-5B). To ensure that this is not an artifact of the GC-poor nature of the centromere, we repeated the analysis with a synthetic shuffled PWM, created by scrambling the order of the columns of the original PWM (that is, scrambling the positions in the motif while keeping the corresponding weight vectors the same). This shuffled motif showed more matches in the centromeres than what is seen in the non-centromeric genomic sequence, but significantly fewer than the authentic centromeric sequences in most chromosomes (Figure 2-5C). A dot-plot analysis was performed to detect the presence of any direct or inverted-repeat structure associated with the centromeres in *M. sympodialis*. Analysis of all of the centromere sequences along with 5 kb flanking sequences using SyMap confirmed the lack of direct/inverted repeat structures (Figure 2-5D).

Figure 2-5. Sequence features of centromeres in *M. sympodialis.*
(A) Logo of the consensus DNA sequence identified from *M. sympodialis* centromeres, graphically represented with the size of the base corresponding with the frequency of occurrence. (B)The genome of *M. sympodialis* was scanned for matches to the 12 bp AT-rich motif using a 500 bp sliding window. Hit counts (*y*-axis) were plotted against the chromosomal coordinates (*x*-axis, in kb). Red asterisks near the line corresponding to maximum enrichment in every chromosome mark the regions predicted as centromeres. (C) Comparison of the maximum motif counts in a 500 bp window in each *M. sympodialis* chromosome (*y*-axis) estimated with the PWM of the 12 bp motif (blue bars, labeled real) and a scrambled version of the matrix (orange bars, labeled scrambled). The *x*-axis indicates individual chromosomes in *M. sympodialis* (D) Dot-plot generated by plotting sequences that include 5 kb flanking the Mtw1-enriched regions in *M. sympodialis* against themselves is depicted. The coordinates of regions used in *M. sympodialis* are *CEN1*: 781716-791716, *CEN2*: 350852-360852, *CEN3*: 232777-242777, *CEN4*: 413320-423320, *CEN5*: 120251-130251, *CEN6*: 97796-107796, *CEN7*: 426611-436611 and *CEN8*: 19905-29905.


A
12 bp AT-rich motif
B
M. sympodialis ATCC42132
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
No. of sites (500 bp window)
Chromosomal coordinates (kb)
C
real motif
shuffled motif
Maximum motif count in 500 bp window
Chr1 Chr2 Chr3 Chr4 Chr6 Chr7 Chr8 Chr5
D
M. sympodialis ATCC42132
CEN1 CEN2 CEN3 CEN4 CEN5 CEN6 CEN7 CEN8

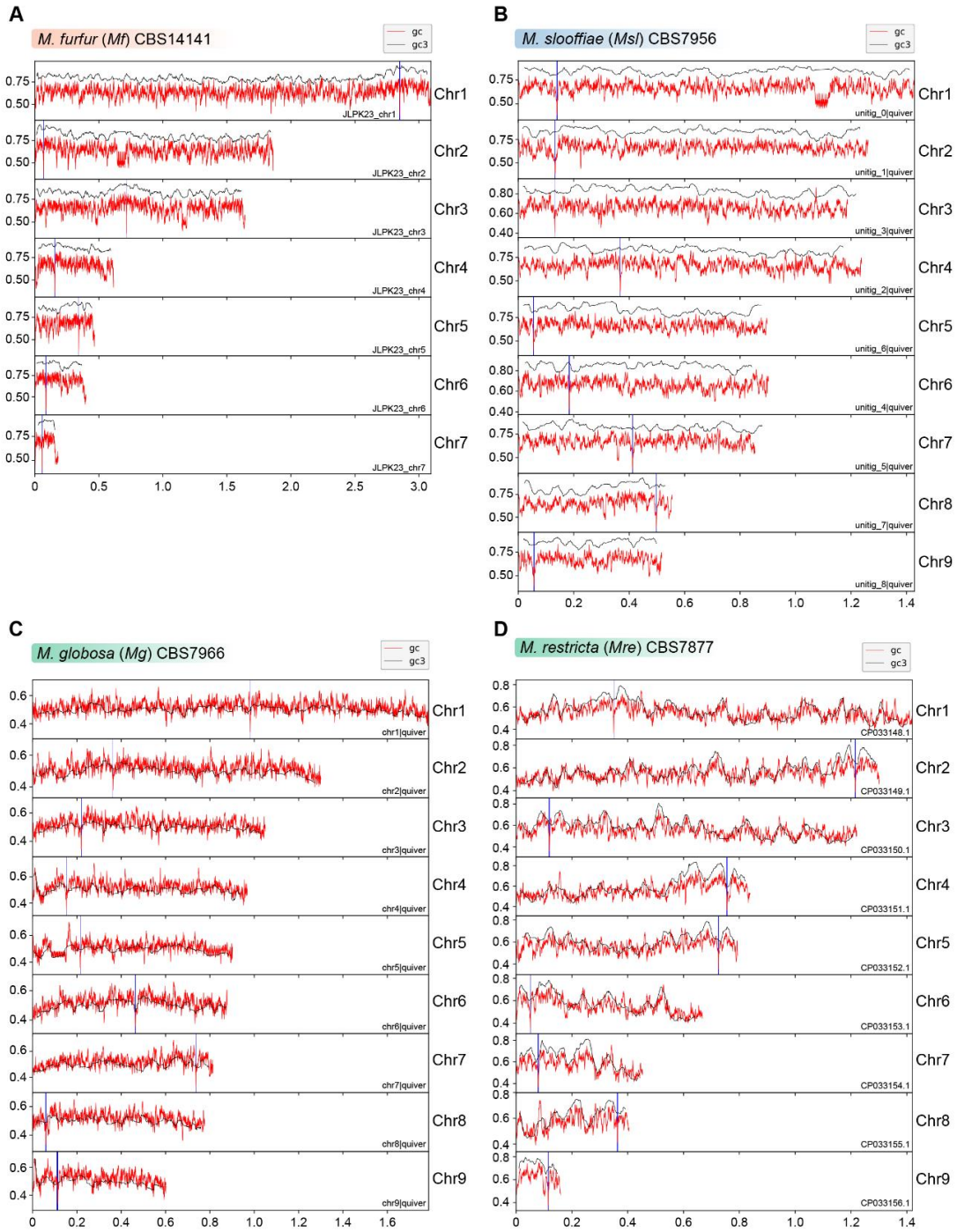


Figure 2-6. Putative centromeres of *M. furfur*, *M. slooffiae*, *M. globosa*, and *M. restricta* map to global GC troughs in each chromosome.

(A-D) Graphs indicating GC content (red lines) and GC3 content (black lines, corresponding to the GC content at the third codon position in genes) of each chromosome of *M. furfur*, *M. slooffiae*, *M. globosa*, and *M. restricta*. The position of putative centromeres mapping to chromosomal GC minima are marked in blue. The *x*-axis indicates chromosomal coordinates in Mb.

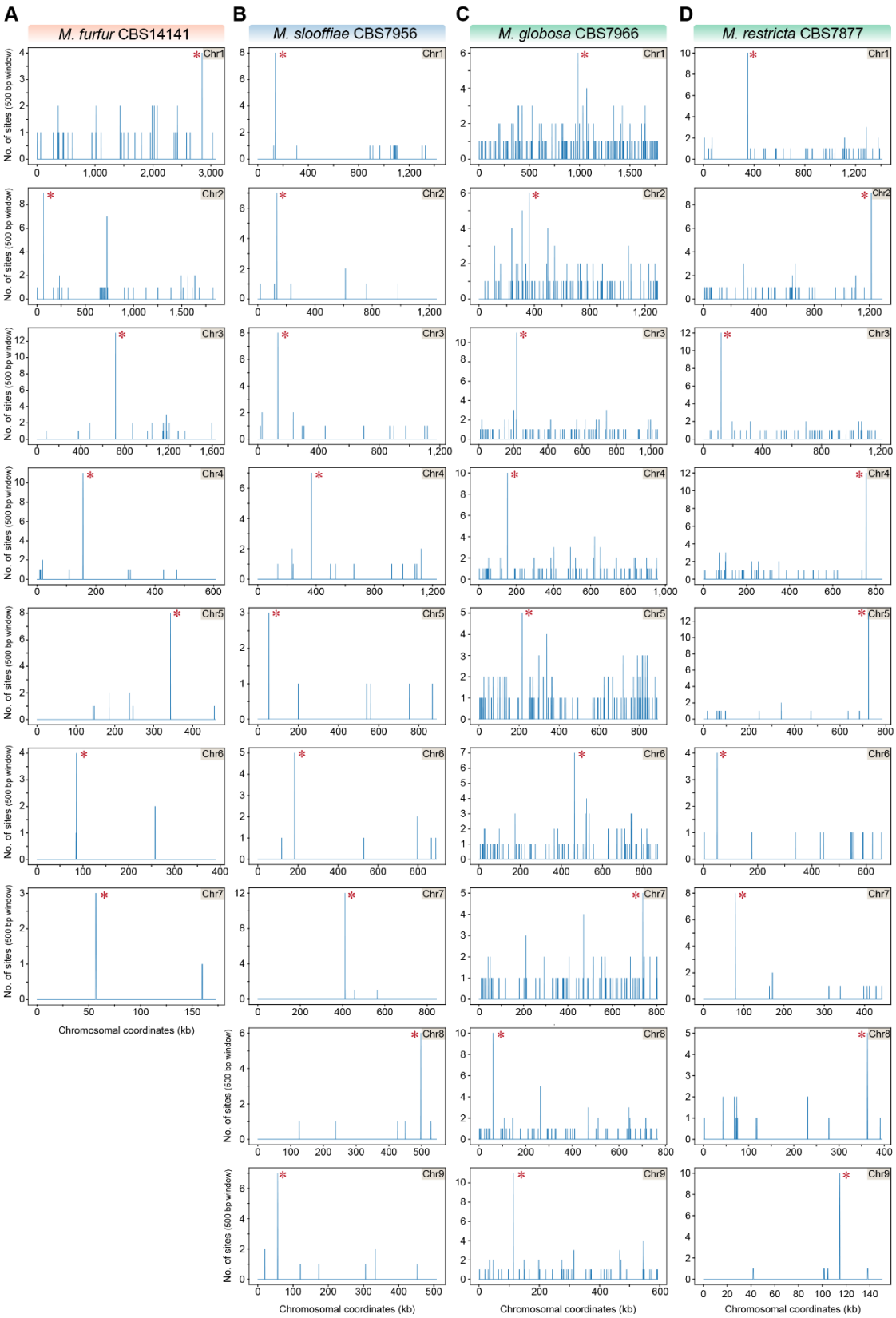

A
M. furfur CBS14141
B
M. slooffiae CBS7956
C
M. globosa CBS7966
D
M. restricta CBS7877
No. of sites (500 bp window)
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chromosomal coordinates (kb)

Figure 2-7. The 12 bp AT-rich motif is enriched at the putative centromeres of *M. furfur*, *M. slooffiae*, *M. globosa*, and *M. restricta*.
(A-D) The genomes of *M. furfur*, *M. slooffiae*, *M. globosa*, and *M. restricta* were scanned for matches to the 12 bp AT-rich motif using a 500 bp sliding window. Hit counts (*y*-axis) were plotted against the chromosomal coordinates (*x*-axis, in kb) for each of the above species. Red asterisks near the line corresponding to maximum enrichment in every chromosome or scaffold mark the regions predicted as centromeres in each species.

In the absence of any centromere exclusive DNA sequence, the unique and distinguishing features of centromere regions in *M. sympodialis* are an AT-rich core region of <1 kb (average AT content of 78% as compared to the genome average of 41.5%) enriched with the 12 bp motif in a kinetochore protein-bound region of 3 to 5 kb. As expected, the kinetochore bound region contains a reduced level of histone H3. The motif count plots and the GC/GC3 plots in the following pages were generated by Rahul Siddharthan and Rakesh Netha (IMSc, Chennai), as a part of a collaboration for this project.

Centromeres in *M. furfur*, *M. slooffiae*, and *M. globosa* map to chromosomal GC minima

Using the unique centromere features identified in *M. sympodialis*, we predicted one centromere locus on each of the 7 *M. furfur* chromosomes, and these all map to chromosomal GC troughs (Figure 2-6A, Table 2-4). Similar to *M. furfur*, we predicted the centromeres in *M. slooffiae*, *M. globosa*, and *M. restricta*, each of which contains 9 chromosomes (Figure 2-6B to 2-6D, Table 2-4). Each of the predicted centromeres showed enrichment of the 12 bp AT-rich motif identified with *M. sympodialis* centromere sequences as compared to other regions in the genomes (Figure 2-7).

To experimentally validate the centromeric loci in *M. furfur*, we functionally expressed the centromeric histone H3 variant CENP-A with a 3xFLAG tag at the C-terminus (Figure 2-8A). We performed ChIP in strain MF001 (CENP-A-3xFLAG) and analyzed immunoprecipitated DNA by qPCR using primers specific to each of the seven predicted centromeres and a centromere-unlinked control locus 1.3 Mb away from *CEN1*. Enrichment of CENP-A at all seven centromeres over the control locus confirmed that the regions predicted are indeed centromeres in *M. furfur* CBS14141 (Figure 2-8B and 2-8C).

Given the lack of genetic manipulation methods for *M. slooffiae* and *M. globosa*, we tested the enrichment of histone H3 at the predicted centromeres in these two species. All of the nine centromeric loci in these two species showed a reduced histone H3 level when compared to a control locus unlinked to centromeres (Figure 2-8D and 2-8E). Furthermore, upon analyzing the enrichment profile at one

centromere (*CEN1*) in *M. slooffiae*, we observed a reduction in the enrichment levels of histone H3 at the GC troughs as compared to the flanking regions (Figure 2-8F). In the case of *M. globosa*, the regions spanning a centromere (*CEN2*) also depicted a similar reduction in the histone H3 levels (Figure 2-8G). Taken together, the significant reduction in the histone H3 levels at the predicted centromeres, indicative of the presence of CENP-A, suggests that these putative centromere regions are indeed *bona fide* centromeres in these species.

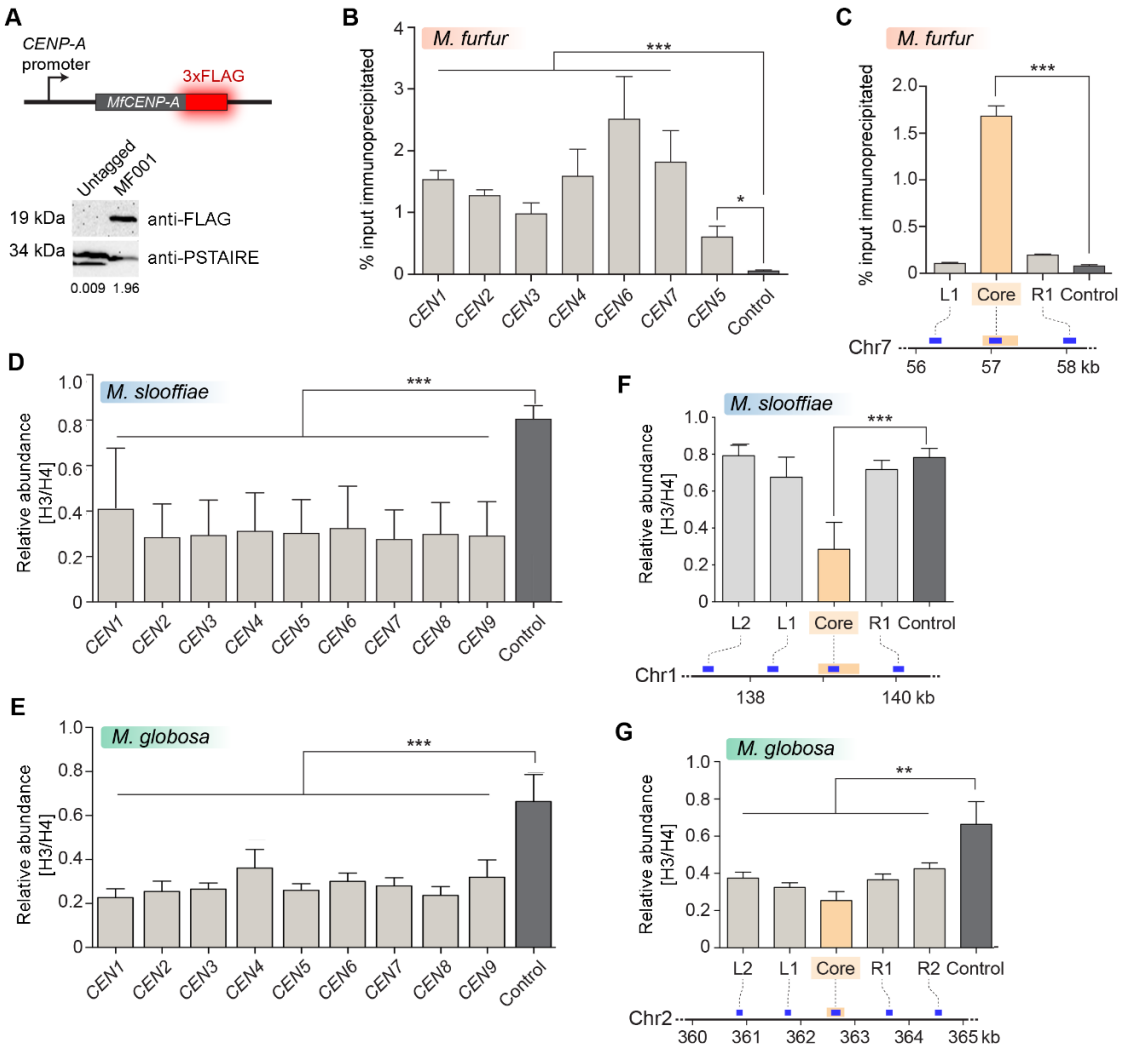


Figure 2-8. Centromeres in *M. furfur*, *M. slooffiae*, and *M. globosa* map to global GC troughs on each chromosome.
(A) Schematic of epitope tagging of CENP-A with 3xFLAG at the C-terminus is shown. (Bottom) Immunoblot analysis using whole-cell lysates prepared from the untagged wild-type *M. furfur* (CBS14141) cells and CENP-A-3xFLAG expressing cells (MF001) probed with anti-FLAG antibodies and anti-PSTAIRE antibodies. PSTAIRE served as the loading control. (B) The abundance of CENP-A at each of the predicted *M. furfur* centromeres by qPCR analysis of DNA immunoprecipitated with anti-FLAG affinity gel in MF001 cells expressing CENP-A-3xFLAG. The *x*-axis indicates individual

CEN regions assayed with primers homologous to the GC troughs on each chromosome that were predicted as centromeres. The non-centromeric control maps to a region 1.3 Mb away from predicted CEN1. The *y*-axis indicates enrichment of CENP-A estimated as the percentage of input immunoprecipitated. (C) The abundance of CENP-A across *MfCEN7* by ChIP-qPCR analysis in MF001 cells. A schematic representation of a 2 kb region is shown below the graph. The yellow bar indicates the centromere core of *CEN7* corresponding to the GC trough. Blue bars indicate regions analyzed by qPCR: L1 and R1- 750 bp away from the centromere core. 'Control' refers to a region 1.3 Mb away from *CEN1*. The CENP-A enrichment is plotted along the *y*-axis as the percentage of input immunoprecipitated. (D, E) Comparison of relative abundance of histone H3 compared to histone H4 at the predicted centromeres to a non-centromeric control locus in *M. slooffiae* and *M. globosa* respectively. Enrichment was estimated as the percentage of input immunoprecipitated with histone H3 and histone H4 antibodies, and their ratio, as indicated in the *y*-axis, is plotted as relative enrichment. The *x*-axis indicates centromeres in each species, and these were assayed with primers homologous to GC troughs in each chromosome that were predicted as the centromeres. The control region unlinked to the centromere corresponds to a locus 630 kb away from *CEN1* in *M. slooffiae* and 416 kb away from *CEN2* in *M. globosa* respectively. (F) The relative abundance of histone H3 compared to histone H4 across *MslCEN1* by qPCR analysis of the DNA immunoprecipitated using histone H3 and histone H4 antibodies. The schematic of *MslCEN1* locus is shown below the graph. The yellow bar indicates the *CEN1* core region corresponding to the GC trough. Blue bars indicate regions analyzed by qPCR: L1 and L2 map to regions 750 bp and 1.5 kb to the left of the *CEN1* core; R1 maps to a region 750 bp to the right of the *CEN1* core. The control region corresponds to a locus 630 kb away from the *CEN1* core. The ratio of enrichment of histone H3 to that of histone H4 is plotted as the relative enrichment in *y*-axis. (G) The relative abundance of histone H3 compared to histone H4 across *MgCEN2* by qPCR analysis of the DNA immunoprecipitated using histone H3 and histone H4 antibodies. The schematic of a 5 kb region containing *MgCEN2* is shown below the graph. The yellow bar indicates the *CEN2* core region corresponding to the GC trough. Blue bars indicate regions analyzed by qPCR: L1 and L2 indicate regions 750 bp and 1.5 kb to the left of the *CEN2* core; R1 and R2 indicate regions 750 bp and 1.5 kb to the right of the *CEN2* core. The control region corresponds to a locus 416 kb away from the *CEN2* core. The ratio of enrichment of histone H3 to that of histone H4 is plotted as the relative enrichment in the *y*-axis. Values from three experiments, each performed with three technical replicates were used to generate the plots in figures B-G. Error bars indicate the standard deviation (SD). Statistical significance was tested by one-way ANOVA: * significant at $P < 0.05$, *** significant at $P < 0.001$.

Table 2-4. Coordinates, length, and GC content (in %) of the centromeres predicted in *M. furfur*, *M. globosa*, *M. slooffiae*, and *M. restricta*.

	Chr./Scaffold	***CEN***	**Core centromere**				**% GC Genome**
			Start	**End**	**Length (bp)**	**% GC**	
M. furfur CBS14141	Chr1	*CEN1*	2,850,135	2,850,402	268	15.7	64.9
	Chr2	*CEN2*	68,763	68,931	168	15.4	
	Chr3	*CEN3*	717,557	718,084	528	22.9	
	Chr4	*CEN4*	155,897	156,301	405	18.3	
	Chr5	*CEN5*	342,885	343,372	488	21.5	
	Chr6	*CEN6*	86,112	86,832	721	27	
	Chr7	*CEN7*	56,894	57,339	445	20.9	

M. globosa CBS7966	Chr1	*CEN1*	981894	982242	349	17.7	52.05
	Chr2	*CEN2*	362480	362807	327	25.9	
	Chr3	*CEN3*	219647	220121	474	27.2	
	Chr4	*CEN4*	152635	152994	359	18.3	
	Chr5	*CEN5*	215437	215595	158	17	
	Chr6	*CEN6*	464007	464114	107	32.4	
	Chr7	*CEN7*	736701	737015	314	18.1	
	Chr8	*CEN8*	59472	59817	345	19.7	
	Chr9	*CEN9*	114080	114535	455	23.5	
M. slooffiae CBS7956	Chr1	*CEN1*	138,919	139,465	547	26	66.31
	Chr2	*CEN2*	132,717	133,193	477	23.1	
	Chr3	*CEN3*	367,665	368,177	513	23.8	
	Chr4	*CEN4*	130,942	131,501	560	27	
	Chr5	*CEN5*	183,442	183,981	540	28.5	
	Chr6	*CEN6*	411,984	412,552	569	27.4	
	Chr7	*CEN7*	54,307	54,889	583	30	
	Chr8	*CEN8*	497,637	498,149	513	24	
	Chr9	*CEN9*	55,948	56,479	532	26.7	
M. restricta CBS877	Chr1	*CEN1*	347,813	348,406	594	29.7	55.73
	Chr2	*CEN2*	87,190	87,806	617	33.3	
	Chr3	*CEN3*	1,101,494	1,102,083	590	33.9	
	Chr4	*CEN4*	754,356	754,989	634	34.2	
	Chr5	*CEN6*	621,177	621,863	687	31.7	
	Chr6	*CEN7*	390,657	391,286	630	35.1	
	Chr7	*CEN8*	362,842	363,381	540	32	
	Chr8	*CEN9*	117,021	117,603	583	32.8	
	Chr9	*CENR*	70,306	70,913	608	36.3	

Centromere loss by breakage resulted in chromosome number reduction in *M. sympodialis*

Synteny of genes across centromeres is mostly conserved in closely related species (BYRNE AND WOLFE 2005; PADMANABHAN *et al.* 2008; YADAV *et al.* 2018b). To understand the transition between different chromosome number states, we analyzed the conservation of gene synteny across centromeres in these species. By mapping gene synteny at the centromeres of *M. globosa* and *M. slooffiae* (each carries 9 chromosomes), compared with *M. sympodialis* (containing 8 chromosomes), we found complete gene synteny conservation in 8 of the 9 centromeres (Figure 2-9A and 2-9B). Thus, syntenic regions of all 8 *M. sympodialis* centromeres are present in the genomes of *M. globosa* and *M. slooffiae*. In the case of *M. restricta*, 7 putative centromeres are completely syntenic with *M. sympodialis* centromeres while one centromere retained partial gene synteny (Table 2-5). However, no gene synteny conservation was observed at the centromere of Chr2 in *M. globosa,* Chr5 in *M. slooffiae,* or Chr8 in *M. restricta* (Table 2-5). This was indicative of the loss of a centromere during the transition from a 9-chromosome state to an 8-chromosome state.

The GC trough corresponding to *MgCEN2/MslCEN5* is flanked by genes that map to MsyChr2 on one arm and MsyChr4 on the other (Figure 2-9C and 2-9D). The centromere region in each of MgChr2 and MslChr5 marks a synteny breakpoint showing no homologous region in the *M. sympodialis* genome indicative of a loss of this centromere DNA sequence. We also observed that the genes flanking the breakpoint are conserved in *M. sympodialis*, suggesting that the specific intergenic region was involved (Figure 2-9E and 2-9F). Evidence for internalization of telomere adjacent ORFs or the presence of interstitial telomere repeats indicative of telomere-telomere fusions was not detected in the *M. sympodialis* genome. These observations strongly support our hypothesis that breakage of *MgCEN2/MslCEN5* (or the orthologous ancestral *CEN*) and fusion of the two acentric arms to other chromosomes resulted in the chromosome number reduction observed between these species. The circos plots and chromosome bar plots indicating synteny break in figures 2-9 to 2-11 were generated by Marco Coelho as a part of a collaboration with Joe Heitman's lab (Duke University).

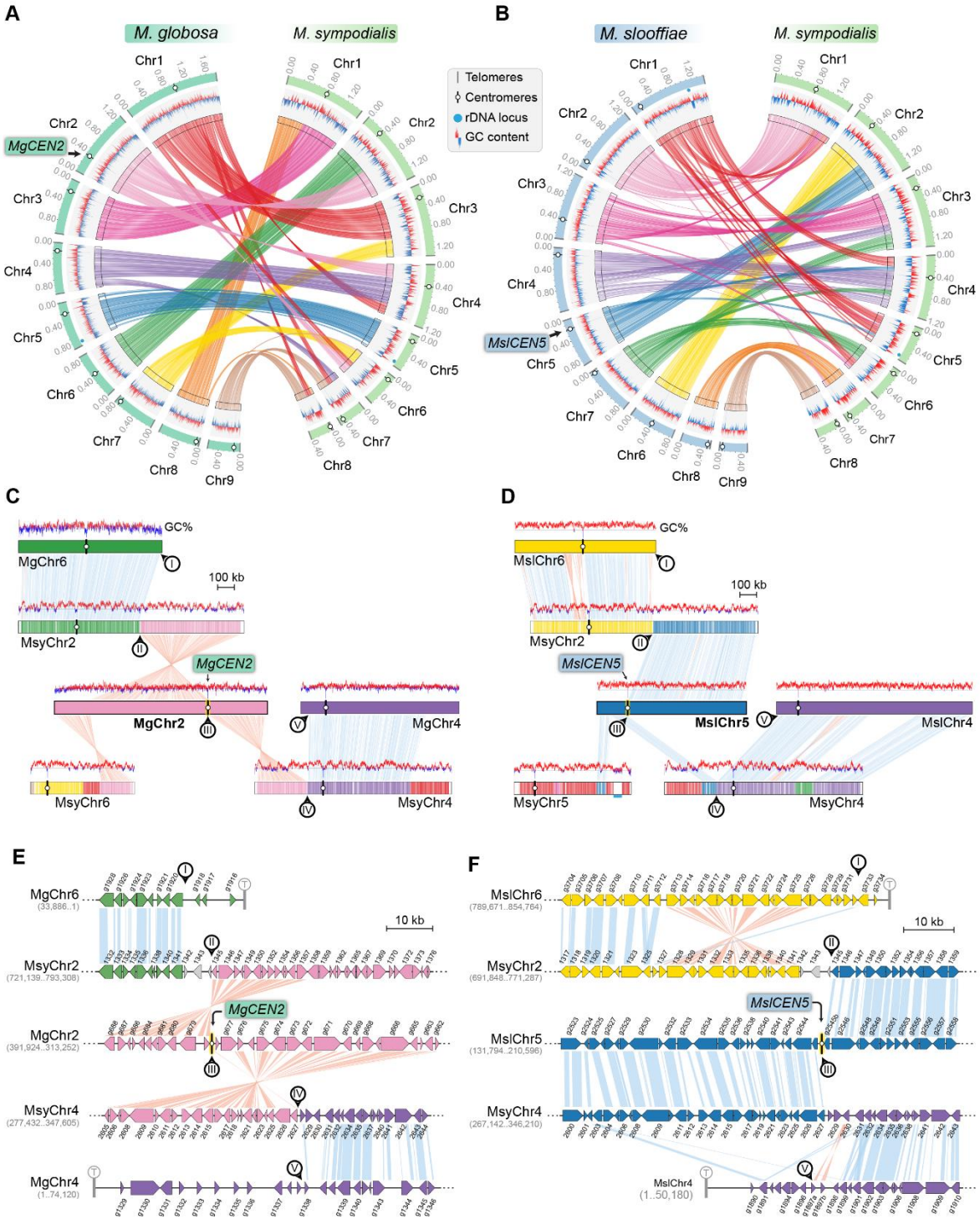


Figure 2-9. *MgCEN2* and *MslCEN5* map to a synteny breakpoint in *M. sympodialis*.
(A, B) Circos plots depicting the conserved gene synteny blocks between *M. globosa* and *M. sympodialis* chromosomes and *M. slooffiae* and *M. sympodialis* chromosomes. Tracks from outside to inside represent positions of centromeres and telomeres, GC content (plotted as blue and red lines indicating GC content below or above genome average, calculated in 0.4 kb non-overlapping windows), and colored connectors indicate regions of conserved gene synteny between the two species.

(C, D) Linear chromosome plots depicting syntenic regions between Chr2 of *M. globosa* and Chr5 of *M. slooffiae*, respectively, with chromosomes of *M. sympodialis*. GC content (in %) is shown as red/blue lines above each chromosome. Labels in black circles mark the gene synteny breakpoints. Synteny breakpoint at *MgCEN2* and *MslCEN5* is marked as III. The regions on MsyChr2 and MsyChr4 where the homologs of ORFs flanking the breakpoint are located are marked II and IV respectively. Labels I and V indicate gene synteny conservation on the other side of the fusion points II and IV on MsyChr2 and MsyChr4 as compared to *M. globosa* and *M. slooffiae* chromosomes respectively. (E, F) Zoomed-in image of the gene synteny breakpoint at *MgCEN2* and *MslCEN5* representing the conservation of genes flanking these centromeres in *M. sympodialis* chromosomes at the ORF level.

Table 2-5. Synteny of centromeres across all of the *Malassezia* species analyzed in this study.
'BP' indicates the presence of a gene synteny break, 'Inactivated' indicates centromere inactivation due to sequence divergence and erosion of AT-richness.

Clade C	**Clade B1**			**Clade B2**		**Clade A**		
M. slooffiae	***M. dermatis***	***M. nana***	***M. sympodialis***	***M. globosa***	***M. restricta***	***M. vespertilionis***	***M. japonica***	***M. furfur***
9 Chr	**8 Chr**	**8 Chr**	**8 Chr**	**9 Chr**	**9 Chr**	**9 Chr**	**9 Chr**	**7 Chr**
CEN2	*CEN1*	*CEN1*	*CEN1*	*CEN3*	*CEN5* partial	*CEN6*	*CEN8*	*CEN4*
CEN6	*CEN2*	*CEN2*	*CEN2*	*CEN6*	*CEN7*	*CEN9*	*CEN4* partial	*CEN2* partial
CEN3	*CEN3*	*CEN3*	*CEN3*	*CEN1*	*CEN3*	*CEN7* partial	*CEN6* partial	*CEN3* partial
CEN4	*CEN4*	*CEN4*	*CEN4*	*CEN4*	*CEN1*	*CEN1*	*CEN2*	*CEN7* partial
CEN1	*CEN6*	*CEN6*	*CEN5*	*CEN5*	*CEN2*	*CEN4*	*CEN7*	*CEN1*
CEN7	*CEN5*	*CEN5*	*CEN6*	*CEN7*	*CEN6*	*CEN5* partial	*CEN5* partial	*CEN6*
CEN8	*CEN7*	*CEN7*	*CEN7*	*CEN9*	*CEN4*	*CEN2*	*CEN9* partial	**In-activated**
CEN9	*CEN8*	*CEN8*	*CEN8*	*CEN8*	*CEN9*	*CEN8*	*CEN1*	**In-activated**
CEN5	**BP**	**BP**	**BP**	*CEN2*	*CEN8* partial	*CEN3*	*CEN3* partial	*CEN5*

Centromere inactivation by sequence divergence and loss of AT-richness resulted in chromosome number reduction in *M. furfur*

To understand the basis of the change in chromosome number from 9 to 7 in *Malassezia* species, we compared the synteny of ORFs flanking the *M. slooffiae* or *M. globosa* centromeres with that *of M. furfur*. Of the 9 centromeres in *M. slooffiae*, three centromeres belonged to conserved gene synteny blocks, and four other centromeres showed partial gene synteny conservation in *M. furfur* (Figure 2-10A). A similar pattern of gene synteny conservation was observed between *M. globosa* and *M. furfur* (Figure 2-10B, Table 2-5).

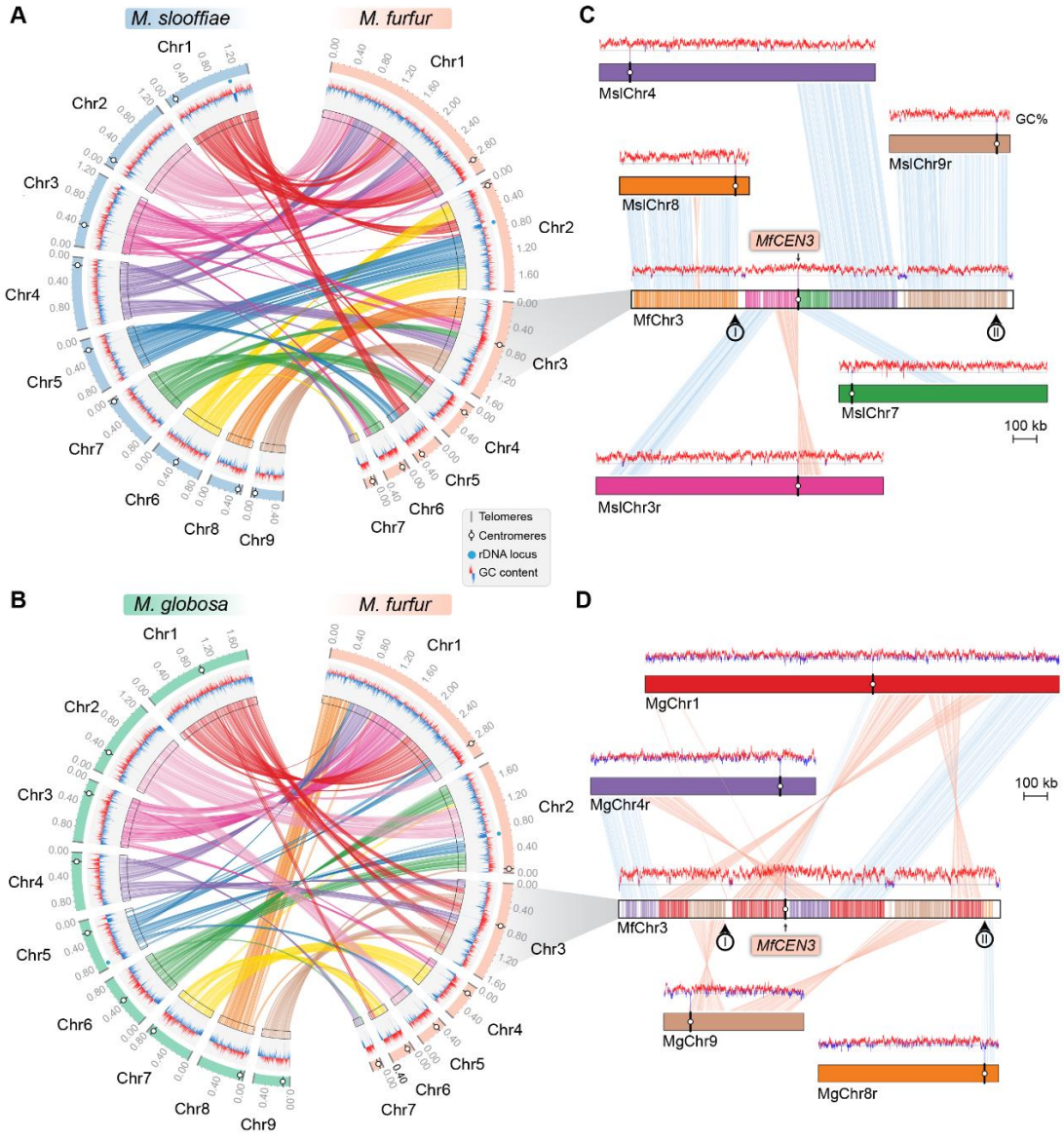
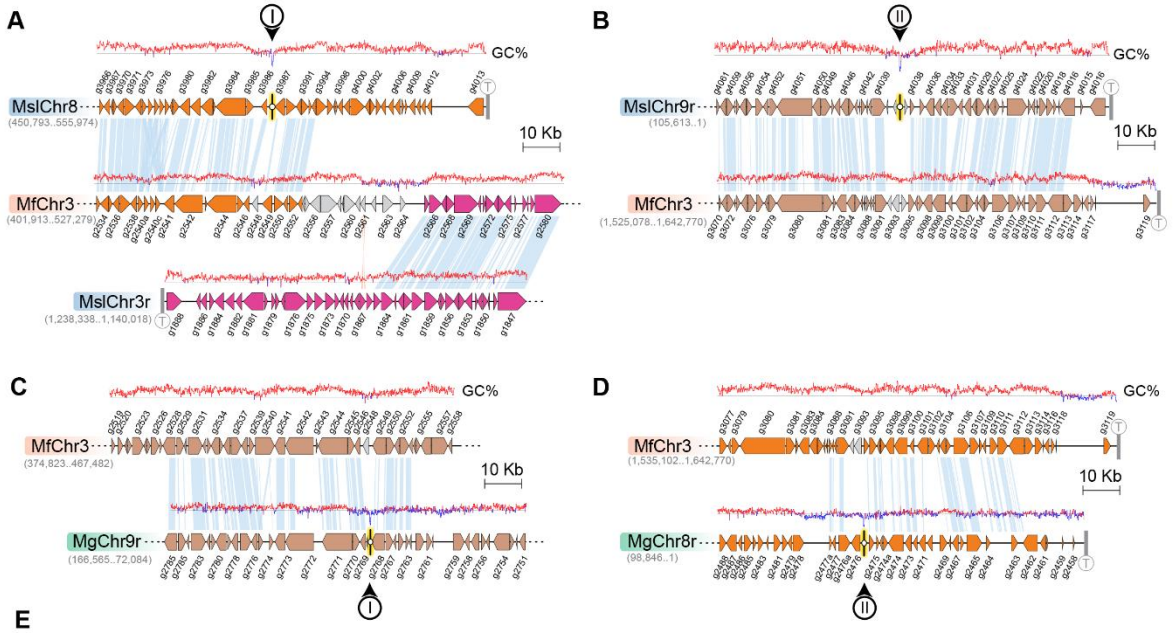


Figure 2-10. Inactivation of *CEN8* and *CEN9* of *M. slooffiae* and *M. globosa* in MfChr3 resulted in chromosome number reduction in *M. furfur*.
(A, B) Circos plots depicting the conserved gene synteny blocks between the *M. slooffiae* and *M. globosa* chromosomes respectively with *M. furfur* chromosomes. Tracks from outside to inside represent positions of centromeres and telomeres, GC content (plotted as blue and red lines indicating GC content below or above genome average, calculated in 0.4 kb non-overlapping windows), and colored connectors indicate regions of conserved gene synteny between the two species. (C) Linear chromosome plot depicting the syntenic regions between chromosome 3 of *M. furfur* and chromosomes 3, 4, 7, 8, and 9 of *M. slooffiae*. GC content (in %) is shown as red/blue lines above each chromosome. Regions corresponding to *MslCEN8* and *MslCEN9* in MfChr3 are marked I and II respectively. (D) Linear chromosome plot depicting the gene synteny conservation between chromosome 3 of *M. furfur* and chromosomes 1, 4, 8, and 9 of *M. globosa*. Regions corresponding to *MgCEN9* and *MgCEN8* in MfChr3 are marked I and II respectively.



E

	Length in *M. globosa*	Length of the corresponding intergenic region in *M. furfur*	GC content in *M. globosa*	GC content of the region in *M. furfur*
MgCEN8	345 bp	730 bp	19.7%	60%
MgCEN9	455 bp	2413 bp	23.5%	54%

Figure 2-11. Loss of AT-content in regions corresponding to *CEN8* and *CEN9* of *M. slooffiae* and *M. globosa* in MfChr3.
(A, B) Zoomed-in view of regions marked I and II in figure 2-10C, corresponding to inactivated *MslCEN8* and *MslCEN9* in *M. furfur* along with the regions of gene synteny is represented. GC content (in %) is shown as a red/blue line diagram above each chromosome. (C, D) Zoomed-in view of the regions marked I and II in figure 2-10D, corresponding to inactivated *MgCEN9* and *MgCEN8* in *M. furfur* along with the regions of gene synteny is represented. GC content (in %) is shown as red/blue lines above each chromosome. (E) The change is GC content observed during inactivation of Mg*CEN8* and Mg*CEN9* is tabulated.

The genes flanking the remaining two centromeres (*CEN8* and *CEN9*) in *M. slooffiae* were present in conserved gene synteny blocks in the two arms of MfChr3 (see synteny blocks corresponding to labels I and II in Figure 2-10C). However, the intergenic regions corresponding to *CEN8* and *CEN9* in *M. slooffiae*, marked as I and II in the ORF-level synteny plots, appear to have evolved to decreased AT-richness in *M. furfur* (Figure 2-11A and B). A similar centromere inactivation mechanism was observed when *CEN8* and *CEN9* of *M. globosa* were compared to the corresponding syntenic regions in *M. furfur* (Figure 2-10D, Figure 2-11C and D). The extent of change in %GC when the intergenic region is present as a centromere (MgCEN8, MgCEN9, MslCEN8, and MslCEN9) as compared to the sequence when they are at the arms of MfChr3 is suggestive of centromere inactivation by changes in the centromeric DNA sequence in these species (Figure 2-11).

The common ancestral *Malassezia* species contained 9 chromosomes

To trace the ancestral karyotype in *Malassezia*, we predicted centromeres and inferred chromosome numbers in other species of clades A and B based on GC troughs and gene synteny. We identified putative centromeres in *M. dermatis* and *M. nana* in Clade B because of their relatively better-assembled genomes distributed in 18 and 13 scaffolds respectively. Of these, we could predict 8 centromeric regions marked by GC troughs that were also enriched with the 12 bp motif in each species (Figure 2-12A, 2-12B, Figure 2-13A, 2-13B, and Table 2-6). Furthermore, in both of these species, the 8 putative centromeres shared complete gene synteny conservation with the regions spanning *M. sympodialis* centromeres, indicating that their common ancestor had 8 chromosomes (green circle in Figure 2-14). To map the common ancestor in Clade B *Malassezia* species, we analyzed regions flanking centromeres of Chr2 of *M. globosa* and Chr8 of *M. restricta*, both of which mapped to the gene synteny breakpoint of the genome of the *Malassezia* species with 8 chromosomes suggesting their common ancestor, named as Ancestor B (Anc. B), also had 9 chromosomes (Figure 2-14, Table 2-5). Based on our centromere predictions in Clade B species and synteny analysis, we propose that centromere breakage would have occurred in the common ancestor of *M. sympodialis*, *M. nana*, and *M. dermatis* after divergence from the common ancestor of *M. globosa* and *M. restricta* which retained a 9-chromosome configuration (Figure 2-14).

As mentioned earlier, *M. furfur* and *M. obtusa* of Clade A contain 7 chromosomes each (BOEKHOUT AND BOSBOOM 1994; BOEKHOUT *et al.* 1998). To further understand the karyotype variations within this clade, we predicted the chromosome number in *Malassezia vespertilionis* and *Malassezia japonica* as their genomes are relatively well assembled (SUGITA *et al.* 2003a; LORCH *et al.* 2018). In both of these species, we could predict 9 GC troughs indicative of centromeres of 9

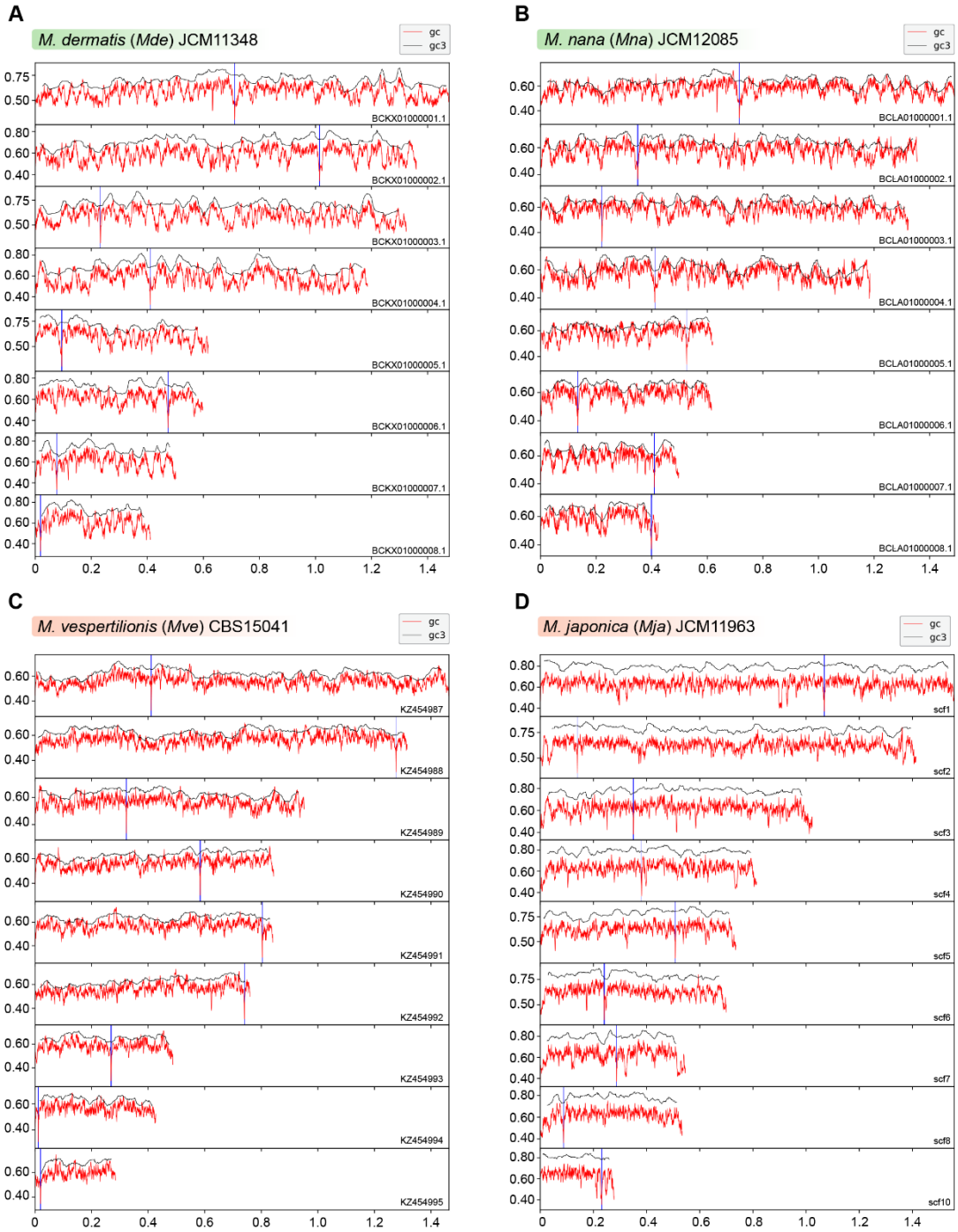


Figure 2-12. Putative centromeres of *M. dermatis*, *M. nana*, *M. vespertilionis*, and *M. japonica* map to global GC troughs in each chromosome.
(A-D) Graphs indicating GC content (red lines) and GC3 content (black lines) of each chromosome of *M. dermatis*, *M. nana*, *M. vespertilionis*, and *M. japonica.* The position of putative centromeres mapping to chromosomal GC minima are marked in blue. The *x*-axis indicates chromosomal coordinates in Mb.

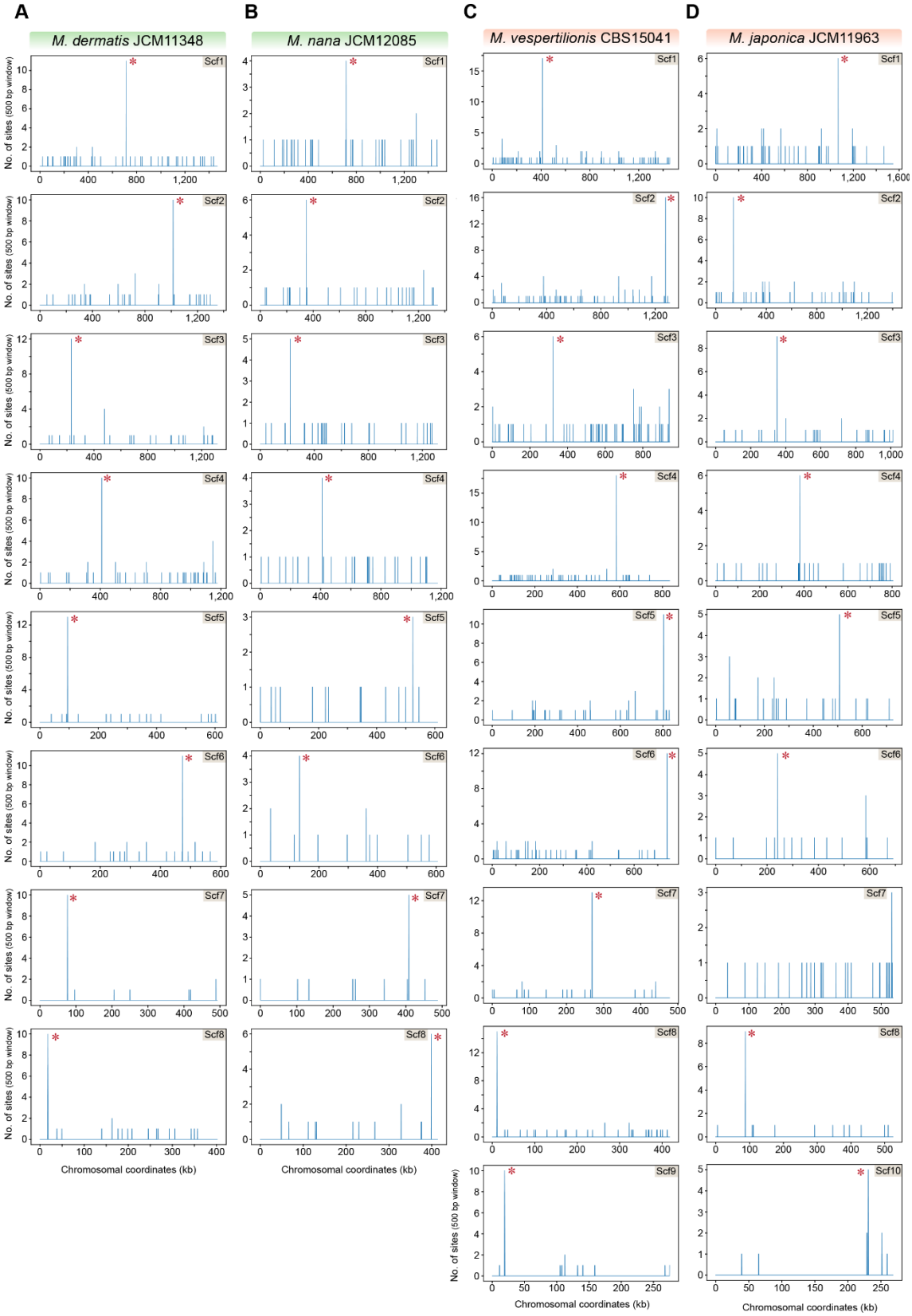

A
M. dermatis JCM11348
B
M. nana JCM12085
C
M. vespertilionis CBS15041
D
M. japonica JCM11963
No. of sites (500 bp window)
Scf1
Scf2
Scf3
Scf4
Scf5
Scf6
Scf7
Scf8
Scf9
Scf10
Chromosomal coordinates (kb)

Figure 2-13. The 12 bp AT-rich motif is enriched at the putative centromeres of *M. dermatis, M. nana, M. vespertilionis,* and *M. japonica*.
(A-D) The genomes of *M. dermatis, M. nana, M. vespertilionis,* and *M. japonica* were scanned for matches to the 12 bp AT-rich motif using a 500 bp sliding window. Hit counts (*y*-axis) were plotted against the chromosomal coordinates (*x*-axis, in kb) for each of the above species. Red asterisks near the line corresponding to maximum enrichment in every chromosome or scaffold mark the regions predicted as centromeres in each species.

chromosomes (Figure 2-12C and 2-12D). The 12 bp motif was found to be enriched in all of the predicted centromeres of *M. vespertilionis* (Figure 2-13C). In the case of *M. japonica*, all of the predicted centromeres (except the centromere of scaffold 7) showed enrichment of the 12 bp motif (Figure 2-13D). The presence of species with 9 chromosomes in Clade A suggests that the ancestral state in this clade, named Ancestor A (Anc. A), also contained 9 chromosomes (Figure 2-14).We identified 9 centromeres in *M. slooffiae*, the only species in Clade C with a well-assembled genome. The presence of species with 9 chromosomes in each of the three clades of *Malassezia* species, the conservation of gene synteny across orthologous centromeres, and the similar centromere features shared by all 9 species analyzed in this study suggest that *Malassezia* species diverged from a common ancestor that had 9 chromosomes with short regional centromeres enriched with the 12 bp AT-rich DNA sequence motif. The phylogenetic tree was provided by Marco Coelho from Joe Heitman's lab (Duke University) as a part of a collaboration on this project.

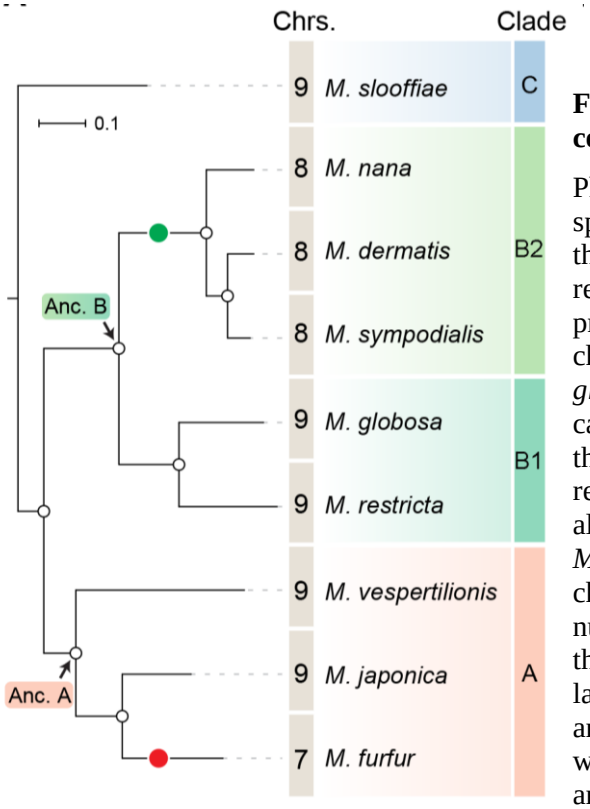


Figure 2-14. Karyotype evolution by loss of centromere function in *Malassezia* species.

Phylogenetic relationships between the *Malassezia* species analyzed in this study are represented and their chromosome numbers shown. Species representing each clade are color-coded based on previous reports (Theelen et al. 2018). The chromosome numbers for *M. slooffiae* and *M. globosa* are based on results from this study. In the case of *M. sympodialis*, *M. restricta*, and *M. furfur*, the chromosome numbers are based on previous reports (Boekhout and Bosboom 1994; Senczek et al. 1999; Zhu et al. 2017). For *M. dermatis*, *M. nana*, *M. vespertilionis*, and *M. japonica*, the number of chromosomes was estimated from the predicted number of centromeres. The node corresponding to the ancestral state of Clade A and Clade B are labeled 'Anc. A' and 'Anc. B' respectively. Green and red circles indicate the origins of karyotypes with 8 and 7 chromosomes respectively from an ancestral state of 9 chromosomes.

Table 2-6. Coordinates, length, and GC content (in %) of the centromeres predicted in *M. nana, M. dermatis, M. vespertilionis*, and *M. japonica.*

.	Chr./Scaffold	*CEN*	Core centromere				% GC Genome
			Start	End	Length (bp)	% GC	
M. nana JCM12085	BCLA01000001.1 (Scf1)	*CEN1*	715,036	715,592	557	27.8	57.95
	BCLA01000002.1 (Scf2)	*CEN2*	349,428	350,120	693	33	
	BCLA01000003.1 (Scf3)	*CEN3*	220,773	221,345	573	27.9	
	BCLA01000004.1 (Scf4)	*CEN4*	410,594	411,387	794	33.2	
	BCLA01000005.1 (Scf5)	*CEN5*	524,594	525,105	512	24.8	
	BCLA01000006.1 (Scf6)	*CEN6*	133,647	134,324	678	33.6	
	BCLA01000007.1 (Scf7)	*CEN7*	408,363	409,067	705	34.2	
	BCLA01000008.1 (Scf8)	*CEN8*	398,756	399,423	668	32.5	
M. dermatis JCM11348	BCKX01000001.1 (Scf1)	*CEN1*	711,456	711,978	523	22.8	59.05
	BCKX01000002.1 (Scf2)	*CEN2*	1,014,281	1,014,977	697	31.7	
	BCKX01000003.1 (Scf3)	*CEN3*	232,065	232,795	731	29.3	
	BCKX01000004.1 (Scf4)	*CEN4*	409,839	410,631	793	29.5	
	BCKX01000005.1 (Scf5)	*CEN5*	94,520	95,018	499	18.2	
	BCKX01000006.1 (Scf6)	*CEN6*	473,487	474,334	848	30.4	
	BCKX01000007.1 (Scf7)	*CEN7*	76,361	76,975	615	26	
	BCKX01000008.1 (Scf8)	*CEN8*	17,893	18,540	648	26.4	
M. vespertilionis CBS15041	KZ454987.1 (Scf1)	*CEN1*	410,820	411,340	521	15.7	56.6
	KZ454988.1 (Scf2)	*CEN2*	1,275,509	1,276,238	730	25.8	
	KZ454989.1 (Scf3)	*CEN3*	322,361	323,277	917	38.2	
	KZ454990.1 (Scf4)	*CEN4*	583,450	584,319	870	28.9	
	KZ454991.1 (Scf5)	*CEN5*	802,843	804,042	1,200	28.8	
	KZ454992.1 (Scf6)	*CEN6*	739,896	740,558	663	22.3	
	KZ454993.1 (Scf7)	*CEN7*	268,699	269,626	928	28.8	
	KZ454994.1 (Scf8)	*CEN8*	10,985	11,865	881	28	
	KZ454995.1 (Scf9)	*CEN9*	19,047	19,724	678	29.1	
M. japonica JCM11963	BCKY01000001.1 (Scf1)	*CEN1*	1068050	1068614	564	25.1	62.35
	BCKY01000002.1 (Scf2)	*CEN2*	139423	139920	497	20.3	
	BCKY01000003.1 (Scf3)	*CEN3*	350068	350603	535	24.3	
	BCKY01000004.1 (Scf4)	*CEN4*	380877	381439	562	24.2	
	BCKY01000005.1 (Scf5)	*CEN5*	507632	508230	598	24.5	
	BCKY01000006.1 (Scf6)	*CEN6*	240968	250550	582	23.7	
	BCKY01000007.1 (Scf7)	*CEN7*	286711	287234	523	24.9	
	BCKY01000008.1 (Scf8)	*CEN8*	87314	87873	559	23.9	
	BCKY01000010.1 (Scf10)	*CEN9*	230906	231456	530	24.1	

Discussion

In this study, we experimentally validated the chromosome number in *M. slooffiae* and *M. globosa* by PFGE analysis. We sequenced and assembled the genomes of *M. slooffiae*, *M. globosa*, and *M. furfur* and compared each one with the genome of *M. sympodialis* to understand the karyotype differences observed in members of the *Malassezia* species complex. These species represent each of the three major clades of *Malassezia* species, with chromosome numbers ranging from 7 to 9. Because centromere loss or gain directly influences the chromosome number of a given species, we experimentally identified the centromeres of these representative species to understand the mechanisms of karyotype diversity. Kinetochore proteins are useful tools in identifying the centromeres of an organism. Localization of the evolutionarily conserved kinetochore protein Mtw1 suggested that kinetochores are clustered throughout the cell cycle in *M. sympodialis*. The ChIP-sequencing analysis identified short regional (< 5 kb long) centromeres in *M. sympodialis* that are depleted of histone H3 and are enriched with an AT-rich sequence motif. Identification of centromeres in *M. slooffiae*, *M. globosa*, and *M. furfur* further suggested that centromere properties are conserved across these *Malassezia* species. By predicting putative centromeres in five other species along with four species with experimentally mapped centromeres described above across three clades of *Malassezia*, we concluded that an AT-rich centromere core of < 1 kb in length enriched with the 12 bp sequence motif is a potential signature of centromeres in nine *Malassezia* species analyzed in this study. Comparative genomics analysis revealed two mechanisms of centromere inactivation resulting in karyotype change. The presence of a 9-chromosome state in each of the three clades along with conserved centromere features and conserved gene synteny across these centromeres, helped us to infer that the ancestral *Malassezia* species had 9 chromosomes with short regional centromeres with an AT-rich core enriched with the 12 bp sequence motif.

Centromeres in the *Malassezia* species complex represent the first example of short regional centromeres in basidiomycetes. Centromeres reported in other basidiomycetes, such as those in *Cryptococcus* and *Ustilago* species, are of the large regional type (YADAV *et al.* 2018b). *Malassezia* species analyzed in this study have a significantly smaller genome (< 9 Mb) as compared to other basidiomycetes and lack RNAi machinery. The occurrence of short regional centromeres in RNAi-deficient *Malassezia* species is in line with a previous finding wherein a reduction in centromere size was observed in RNAi-deficient basidiomycete species as compared to their RNAi-proficient relatives (YADAV *et al.* 2018b). With the presence of clustered kinetochores across cell cycle stages, and the absence of key genes encoding the RNAi machinery, these *Malassezia* species resemble ascomycetes such as many of the CTG clade species with short regional centromeres (SANYAL *et al.* 2004; NAKAYASHIKI *et al.* 2006; PADMANABHAN *et al.* 2008; KAPOOR *et al.* 2015; CHATTERJEE *et al.* 2016; YADAV *et al.* 2018a), rather than the basidiomycetes with large regional centromeres. By combining these

features, we conclude that the genome size and the presence of complete RNAi machinery could be the determinants of the centromere type of a species irrespective of the phylum it belongs to.

Based on the binding patterns of the kinetochore protein across the *M. sympodialis* genome, the 3 to 5 kb long region can be divided into two domains: (a) an AT-rich *CEN* core that maps to the intergenic region containing the GC trough showing maximum kinetochore binding (< 1 kb) and (b) the regions flanking the core that shows basal levels of kinetochore protein binding. We observed conservation of the 12 bp AT-rich motif in the centromere core across the nine *Malassezia* species. It should also be noted that the 12 bp motif is significantly enriched at the centromeres but not exclusive to the centromeres as it is detected across the chromosomes at a lower frequency. Also, it is evident from the sequence logo of the motif that it is not a unique sequence, rather, degeneracy is observed in almost all positions in this 12 bp motif. This makes it difficult to ascribe a genetic basis of centromere specification (sequence directed kinetochore protein recruitment) as observed in *S. cerevisiae* where the centromere DNA sequences are bound by a cognate CBF3 complex. Taken together with the loss of AT-richness in the inactivated centromeres in *M. furfur*, these results highlights AT-richness as a defining, centromere exclusive feature in *Malassezia* species. The functional significance of the AT-rich core, the motif or the flanking sequences can be dissected by testing their ability to confer mitotic stability to minichromosomes. Alternatively, centromeres can be experimentally deleted, and the fate of the chromosomes can be monitored to detect the formation of neocentromeres. Such systems will enable us to understand the requirements of a functional centromere in *Malassezia*.

However, testing these domains for centromere function in various *Malassezia* species is challenging at present due to technical limitations. Other than *M. sympodialis*, *M. pachydermatis*, and *M. furfur*, no other *Malassezia* species have been successfully transformed (IANIRI *et al.* 2016; CELIS *et al.* 2017; IANIRI *et al.* 2017b). Moreover, in all of these cases, the genetic manipulations are performed by *Agrobacterium*-mediated transconjugation, which is not amenable for the introduction of circular plasmids. Deletion of centromeres in a haploid species can be challenging given the lack of tools like Cre-lox or Flp-frt system that enable conditional deletion. Hence, the functional significance of the 12 bp motif and other domains remain as important questions to be addressed towards understanding the centromere function in these species in the future.

While the presence of the 12 bp motif at non-centromeric loci makes it unlikely to be a direct kinetochore protein recruitment site, it can facilitate centromere function by other modes as described below. The centromeres in *M. sympodialis* contain transcribed ORFs, as documented in centromeres of rice, maize, and *Zymoseptoria tritici* (NAGAKI *et al.* 2004; WANG *et al.* 2014a). In contrast to these cases, our read count analysis did not reveal any significant difference in the transcription (RPKM values) of

centromere associated ORFs to that elsewhere in the genome of *M. sympodialis*. We posit that the 12 bp AT-rich motif sequences could facilitate the transcription of these sequences by recruiting transcription factors having a possible role in kinetochore assembly/function. This hypothesis stems from the following observations. A GATA- type transcription factor Ams2 was found to regulate centromere function in *S. pombe* (HEMMERICH *et al.* 2000). This zinc finger domain containing protein binds to a GATA-core consensus DNA sequence, which were detected at the central core in *S. pombe* centromeres apart from gene promoters. This sequence directed association of Ams2 to the centromere DNA was proposed to regulate recruitment of CENP-A either by direct association or indirectly by modulating the transcriptional status of the centromeres. (HEMMERICH *et al.* 2000; CHEN *et al.* 2003). A similar role for transcriptional modulators in regulating the genetically defined *S. cerevisiae* centromere function have been identified. Proteins like Cbf1, Ste12 associate with the centromeres and were found essential for generation of transcripts from this region (OHKUNI AND KITAGAWA 2011; LING AND YUEN 2019). In line with this function, binding sites for these factors were found in the centromeres (for Cbf1) and adjacent pericentric regions (for Ste12) in addition to other regions of the genome. Neither these proteins or their cognate binding sites are exclusively present at the centromere, but were still shown to modulate chromosome segregation fidelity (OHKUNI AND KITAGAWA 2011; LING AND YUEN 2019). While these studies reinforce the role of transcription in centromere function irrespective of the centromere structure, they also act as a precedence towards identifying a functional role for the 12 bp AT-rich motifs which are enriched but not exclusive to the centromeres.

In this study, we report three high-quality chromosome-level genome assemblies and identified centromeres in nine *Malassezia* species, representing all of the three *Malassezia* clades with differing numbers of chromosomes. This will serve as a rich resource for comparative genomics in the context of niche adaptation and speciation. Analysis of gene synteny conservation across centromeres using these genomes revealed breakage at the centromere as one of the mechanisms resulting in a karyotype change between closely related species - those with 9 chromosomes, such as *M. slooffiae* and *M. globosa,* and those with 8 chromosomes like *M. sympodialis* (Figure 2-15). Gene synteny breakpoints adjacent to the centromeres have been reported in *C. tropicalis* that has seven chromosomes - one less than *C. albicans* (CHATTERJEE *et al.* 2016). Centromere loss by breakage was proposed to have reduced the *Ashbya gossypii* karyotype by one when compared to the pre-whole genome duplication ancestor (GORDON *et al.* 2011). Breakpoints of conserved gene synteny between mammalian and chicken chromosomes were also mapped to the centromeres (INTERNATIONAL CHICKEN GENOME SEQUENCING 2004). Similar consequences in the karyotype have been reported in cases where centromeres were experimentally excised. Besides neocentromere formation, survival by fusion of acentric chromosome arms has been

shown in *S. pombe* (ISHII *et al.* 2008). Such fusions are detected upon deletion of centromeres in another basidiomycete, *C. deuterogattii* (SCHOTANUS AND HEITMAN 2019). By comparing the ancestral state (*M. slooffiae*) and other *Malassezia* species with either the same number or fewer chromosomes, we observed gene synteny breaks adjacent to centromeres (indicated by partial synteny conservation) apart from the break observed at *MgCEN2* or *MslCEN5*. Is this suggestive of the fragile nature of *Malassezia* centromeres? We advance the following hypothesis to explain the observed breaks at centromeres.

Studies of the common fragile sites in the human genome suggest different forms of replication stress as a major source of instability and subsequent breakage at these sites (HELMRICH *et al.* 2011; LETESSIER *et al.* 2011; OZERI-GALAI *et al.* 2011). The resolution of the resulting replication fork stall is critical for the stability of these fragile sites (SCHWARTZ *et al.* 2005). Studies on the fragile site FRA16D in humans show that the AT-rich DNA (Flex1) results in fork stalling as a consequence of cruciform or secondary structure formation (ZHANG AND FREUDENREICH 2007). Centromeres are natural replication fork stall sites in the genome (GREENFEDER AND NEWLON 1992; SMITH *et al.* 1995; MITRA *et al.* 2014). The AT-rich core centromere sequence in *M. globosa* is also predicted to form secondary structures, which can be facilitated by the inherent replication fork stall at the centromeres. Whenever these secondary structures are unresolved and the fork restart fails, DSBs can occur at the centromeres. Chromosomal breakage and aneuploidy resulting from such defects are known in cancers (KOPS *et al.* 2005). In mammals, centromeric DSBs are repaired efficiently compared to regions elsewhere in the genome, primarily due to the presence of several homology tracts in the form of repetitive DNA sequences and the stiffness provided by the inherent heterochromatic state to facilitate ligation (RIEF AND LOBRICH 2002). *Malassezia* species are haploid in nature and lack typical pericentric heterochromatin marks. While the efficiency of centromeric DSB repair in the absence of long tracts of homologous sequences is not known in this species complex, we propose that the AT-rich core sequences, by secondary structure formation during DNA replication, could occasionally undergo DNA breakage at the centromere in *Malassezia* species.

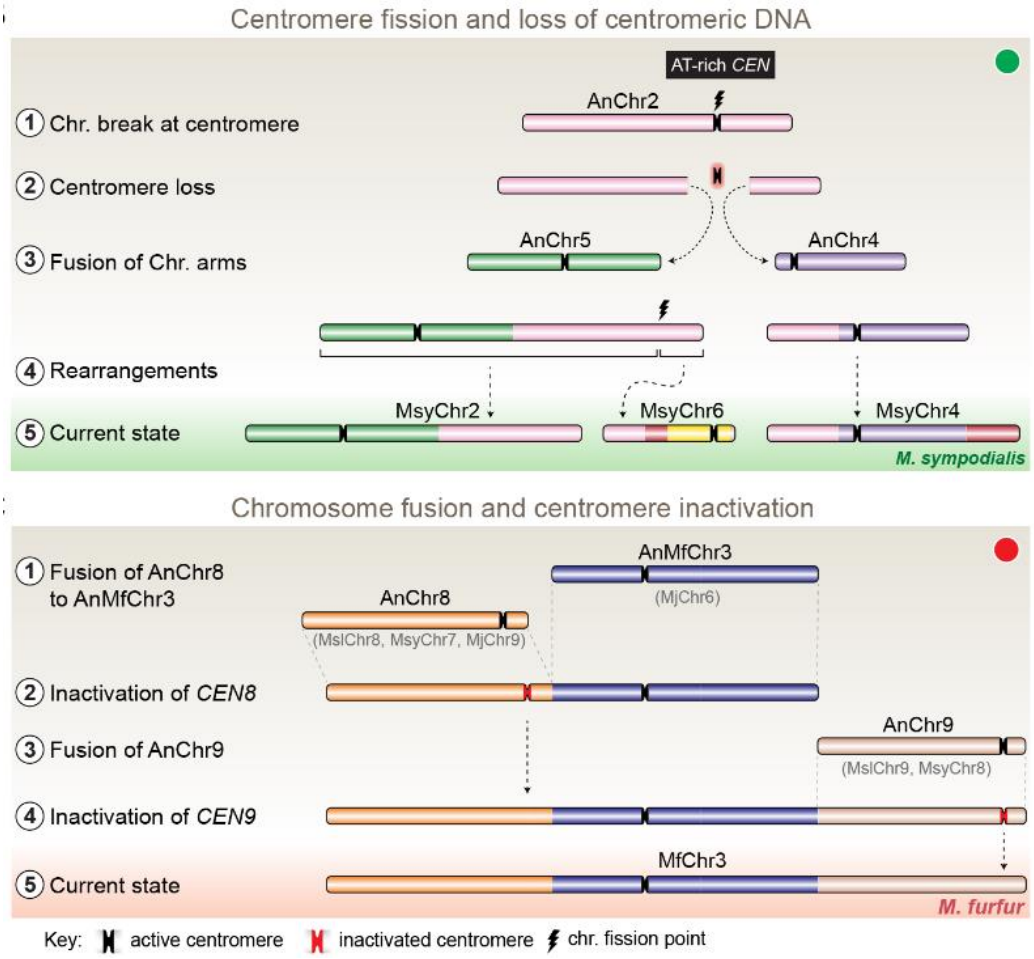


Figure 2-15. Mechanisms of chromosome number reduction in *Malassezia* species. Schematic of the centromere loss by breakage and the resulting reduction in chromosome number as observed in *M. sympodialis* (represented as the current state). A karyotype with 9 chromosomes (as shown for *M. globosa*) is depicted as the ancestral state. (C) Proposed model of centromere inactivation observed in *M. furfur*. The fusion of AnChr8 and AnChr9 to the AnMfChr3 equivalent, resulting in a 7-chromosome configuration is depicted. The fusion product corresponding to extant MfChr3 is represented as the current state

The second mechanism of chromosome number reduction based on our comparative analyses of the *M. furfur* genome with genomes of *M. slooffiae* or *M. globosa* involves the inactivation of centromeres in the process of transition from a 9-chromosome state to a 7-chromosome state (Figure 2-15). Centromere inactivation occurs in cases involving the fusion of centric chromosomal fragments, stabilizing the fusion product, and generating a functionally monocentric chromosome, as seen in case of the origin of human Chr2 from the shared ancestor with the great apes (YUNIS AND PRAKASH 1982; JDO

et al. 1991). A larger proportion of known centromere inactivation events were shown to be mediated by epigenetic modification wherein inactivated centromeres are enriched with marks such as H3K9me2/3, H3K27me2/3, or DNA methylation, emphasizing the role of heterochromatin in this process (ZHANG *et al.* 2010; KOO *et al.* 2011; SATO *et al.* 2012). Deletion of the centromere sequence corresponding to kinetochore binding has also been reported as an alternate mechanism, albeit less frequent, in both humans and yeasts (STIMPSON *et al.* 2010; GORDON *et al.* 2011; SATO *et al.* 2012). Unlike the above two modes, we observed divergence in the sequences corresponding to the inactivated centromeres (*CEN8* and *CEN9* of both *M. slooffiae* and *M. globosa*) in the arms of *M. furfur* Chr3, resulting in the loss of AT-richness of these centromere core regions (Figure 2-15). This is also suggestive of a functional role for AT-rich DNA in centromere function in these species.

A change in chromosome number between two closely related species like *C. albicans* and *C. tropicalis* is associated with a change in centromere structure, between unique short regional centromeres in *C. albicans* that are epigenetically regulated, as compared to a genetically defined homogenized inverted repeat-associated centromere in *C. tropicalis* (CHATTERJEE *et al.* 2016). Strikingly, in both the transitions described above for *Malassezia* species, we did not observe any change in the centromere structure. The emergence of evolutionarily new centromeres, as seen in the case of primates, was not detected in *Malassezia* species (ROCCHI *et al.* 2009; KALITSIS AND CHOO 2012). This is particularly striking in the absence of conservation of any specific centromere-exclusive DNA sequence. This suggests that a strong driving force helps maintain the highly conserved centromere properties in closely related *Malassezia* species descended from their common ancestor, even after extensive chromosomal rearrangements involving centromeres that might have driven speciation. Furthermore, centromere inactivation/loss of centromere function seems to be a conserved theme mediating chromosome number variation from unicellular yeast species to metazoans, including primates.

Chapter 3

Functional analysis of a conserved motif in an outer kinetochore protein Dad2 in *Candida albicans*

The Dam1 complex and kinetochore-microtubule interactions

The kinetochore complex physically links the replicated sister chromatids to spindle microtubule polymers for the proper segregation of sister chromatids into daughter cells. Kinetochores need to sustain contacts with the microtubules through constant cycles of polymerization and depolymerization at the microtubule ends. This interaction helps the separation of the sister chromatids. From a mechanical standpoint, such stable associations with the dynamic microtubule ends can be better achieved with a cooperative effect of multiple low-affinity contacts as compared to high-affinity/low turnover binding.

Studies across eukaryotic systems identified the KMN network as the principal microtubule-binding component of the kinetochore (CHEESEMAN *et al.* 2006; CHEESEMAN AND DESAI 2008). The Ndc80 complex, from both yeasts and humans, bind to and shows biased diffusion along microtubules *in vitro.* The ability of Ndc80 complex to remain stably associated and track the depolymerizing ends in the presence of cargo required an array of Ndc80 molecules instead of a single Ndc80 complex (MCINTOSH *et al.* 2008; POWERS *et al.* 2009; UMBREIT *et al.* 2012). In line with these observations, about 6-20 molecules of the Ndc80 complex was estimated to be present at the kinetochores. However, the efficiency of Ndc80 in each of the above assays diminished upon raising the salt concentrations to physiologically relevant levels, suggestive of the requirement for additional couplers to achieve this process *in vivo*.

The Dam1 complex in yeasts was initially identified as a regulator of the mitotic spindle and its attachment with kinetochores (BUTTRICK AND MILLAR 2011). The complex adopts a ring-like structure around the microtubules *in vitro* and *in vivo,* a conformation favorable to glide along the microtubule walls and remain attached to under the curved protofilaments at the plus end of the microtubules. In isolation, the purified Dam1C can bind to microtubules, track, and transport cargo along with the depolymerizing ends (MIRANDA *et al.* 2005; WESTERMANN *et al.* 2005; WESTERMANN *et al.* 2006). Indeed, tethering of this outer kinetochore complex was sufficient for stable inheritance of an acentric plasmid (KIERMAIER *et al.* 2009; LACEFIELD *et al.* 2009). The redundancy of an additional microtubule-binding entity was resolved when it was found that the addition of the Dam1 complex enhanced the efficiency of Ndc80C- microtubule interactions. In the presence of the Dam1 complex, Ndc80C remained stably associated with the microtubules for a longer time and showed tip-tracking ability even in the absence of external oligomerization (LAMPERT *et al.* 2010; TIEN *et al.* 2010). While the Dam1 complex is not essential for the establishment of initial attachments to microtubules (which requires the Ndc80C), it is essential for bi-orientation consistent with its ability to impart stability in the presence of tension. These factors make the Dam1 complex function more relevant in systems that harbor a single microtubule attachment per kinetochore.

In line with the above *in vitro* findings, mutants of any of Dam1 complex subunits in several fungi show delayed mitotic progression, defects in the spindle structure, and massive nuclear missegregation (CHEESEMAN *et al.* 2001b; JANKE *et al.* 2002; LIU *et al.* 2005; THAKUR AND SANYAL 2011; CHATTERJEE *et al.* 2016; SHAH *et al.* 2019). The essentiality of the Dam1 complex proteins, along with their fungus-exclusive presence, makes them attractive targets for therapeutic intervention.

Dad2 in *C. albicans*

As mentioned above, proteins of the Dam1 complex, including Dad2, were found to be essential for viability in *C. albicans* (THAKUR AND SANYAL 2011). The depletion of Dad2 resulted in a mitotic arrest mediated by the spindle assembly checkpoint. Strikingly, persistent depletion resulted in the morphogenetic transition of yeast cells to elongated large budded cells that eventually lost viability. Analysis of cells arrested at large-budded stage revealed defects in the spindle structure in the form of short spindles in the majority of cells and long, broken spindles in a minor fraction of the population. An increased frequency of nuclear missegregation, similar to the case of Dad2 mutants in *S. cerevisiae,* was also observed. The depletion of Dad2 also resulted in the disintegration of the kinetochore ensemble and the proteasomal degradation of CENP-A^{Cse4} (THAKUR AND SANYAL 2012). Additionally, of the Dam1 complex subunits studied, only Dad2 was found to localize to the spindle mid-zone in *C. albicans,* in a lower proportion of cells. Upon overexpression, the localization of Dad2 along the entire mitotic spindle was also detected in some cells. The molecular basis of this localization and its functional significance remains elusive (THAKUR AND SANYAL 2011), given no interaction between Dad2 and tubulin subunits has been mapped to date (LEGAL *et al.* 2016; JENNI AND HARRISON 2018).

In this study, we identified a 10 amino acid- long motif at the C-terminus of CaDad2 consisting of conserved arginine residues that we propose to be the microtubule-binding domain of Dad2. Despite poor sequence identity between fungal Dad2 sequences, this motif that we named the Dad2 signature sequence (DSS) remains conserved. In the following part of the thesis, we discuss the results from mutagenesis based genetic and biochemical assays towards understanding the role of this conserved DSS in mediating *Ca*Dad2 function.

Results

Identification of Dad2 signature sequence

To test the microtubule-binding ability of Dad2, we analyzed the primary amino acid sequence of Dad2 from several fungal species belonging to the phyla Ascomycota and Basidiomycota. From the multiple sequence alignment of 32 Dad2 sequences, we identified a conserved N-terminal region that corresponds to the Hidden Markov Model (HMM) profile of the Dad2 family of proteins (black bar below the alignment, Figure 3-1A). In addition to this, we detected a second conserved region spanning 10 amino acids near the C-terminus of Dad2 protein sequences (red bar below the alignment, Figure 3-1A) that could not be detected in a pairwise alignment of Dad2 sequences from *C. albicans* and *S. cerevisiae*. The amino acid logo of this sequence motif is highlighted (Figure 3-1B). This region in CaDad2 consists of one highly conserved proline and arginine residues (P87 and R92) and another arginine residue (R94) that is relatively less conserved but majorly represented by positively charged or neutral residues at this position (Figure 3-1B). The presence of conserved positively charged residues in this region provided the first hint at a possible microtubule-binding role of this region. This stems from two observations made from biochemical and cell biological studies on the microtubule-binding proteins of the kinetochores. First, the tubulin monomers are known to contain a stretch of negatively charged glutamate residues at their C-terminus (commonly known as E-hooks). These regions are known to mediate electrostatic interactions between microtubules and complexes such as the Dam1C and the Ndc80C, which form the load-bearing module of the kinetochore ensemble (WESTERMANN *et al.* 2005). Second, the two subunits of the ternary Ska complex-Ska1 and Ska3, the functional ortholog of the Dam1 complex in metazoans, were found to contain conserved positive charged residues composed of arginine and lysine that were shown to mediate the binding of the Ska complex to microtubules (WELBURN *et al.* 2009; ABAD *et al.* 2014; ABAD *et al.* 2016).

To experimentally address the function of the conserved DSS motif, we developed a reintegrant system based on a conditional mutant of Dad2 in *C. albicans* (J108: *dad2/PCK1pr-DAD2*) (Figure 3-2A) (THAKUR AND SANYAL 2011). The transcription of the only allele of *DAD2* is regulated by the *PCK1* promoter that transcribes when J108 cells are grown in the presence of succinate and repressed when grown in the presence of dextrose (Figure 3-2B).

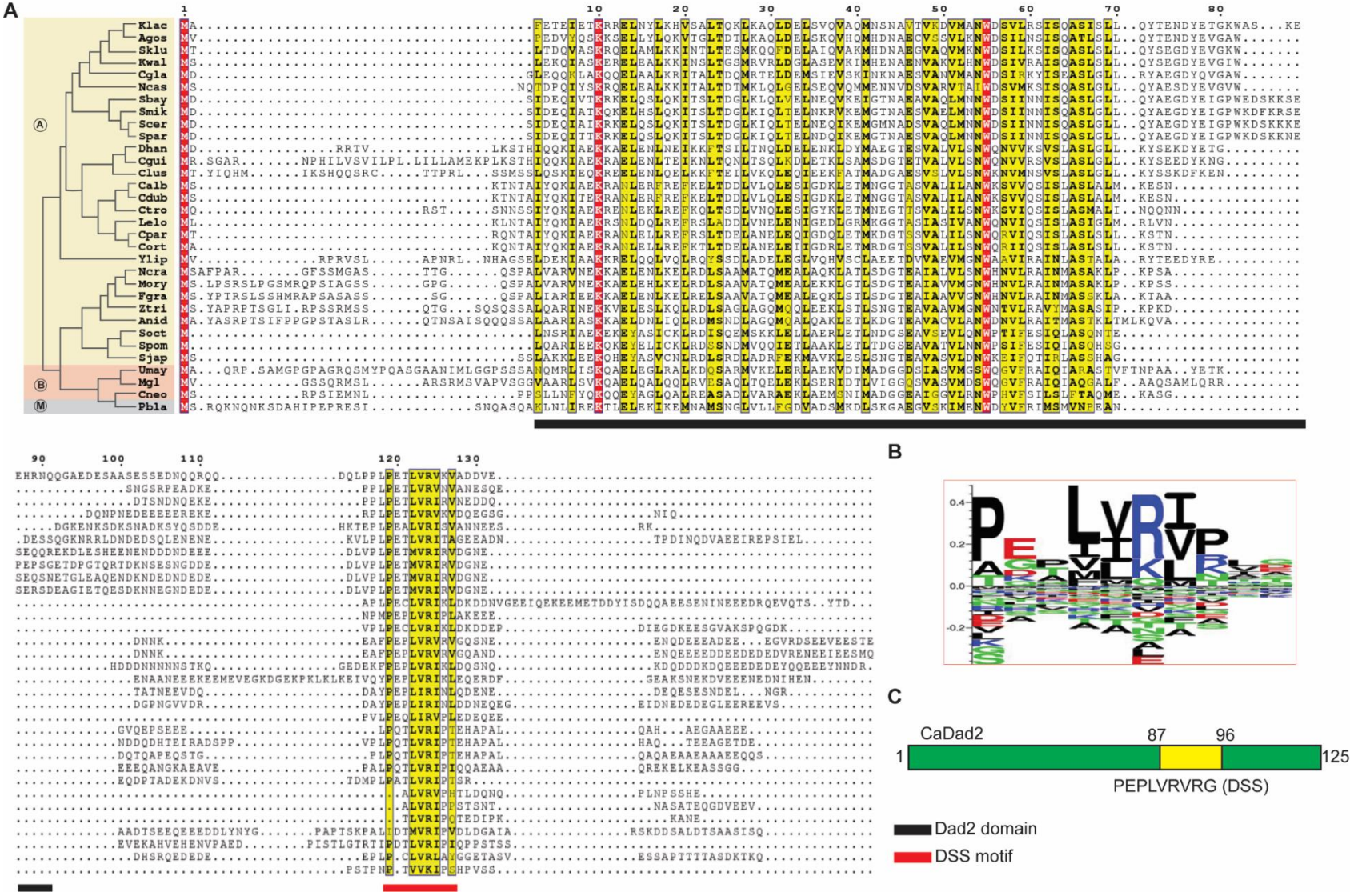

A
Klac
Agos
Sklu
Kwal
Cgla
Ncas
Sbay
Smik
Scer
Spar
Dhan
Cgui
Clus
Calb
Cdub
Ctro
Lelo
Cpar
Cort
Ylip
Ncra
Mory
Fgra
Ztri
Anid
Soct
Spom
Sjap
Umay
Mgl
Cneo
Pbla
B
C
CaDad2
87
96
1
125
PEPLVRVRG (DSS)
Dad2 domain
DSS motif

Figure 3-1. Identification of a conserved Dad2 signature sequence (DSS) at the C-terminus of Dad2.
(A) A maximum-likelihood phylogenetic tree and multiple sequence alignment of Dad2 from 32 fungal species are collectively represented. Species from different fungal sub-phyla are color-coded and labeled as A (Ascomycota), B (Basidiomycota), and M (Mucoromycota). Two conserved blocks were identified- the Dad2 domain at the N-terminus (in the black bar, marked for CaDad2 residues 3-74 based on HMM for Dad2 family), and Dad2 Signature Sequence (DSS, the gray bar below the alignment marked for CaDad2 residues 87-96) at the C-terminus. (B)The amino acid logo of DSS based on the MSA, with the size of each amino acid indicating the probability of occurrence. Amino acids are color-coded based on the inherent charge contributed by the side chain: Black- neutral, blue-positive, and red- negative. (C) Schematic representation of CaDad2 polypeptide in its linear form highlighting the position and sequence of DSS in this species.

At an ectopic but neutral *RPS1* locus in this strain background, we reintegrated a Protein A-tagged (TAP-) Dad2 that either contains the DSS motif (CaSR100- *dad2/PCK1pr-DAD2 RPS1/ RPS1 (DAD2pr-DAD2^{FL}-TAP-NAT*)) or lacks it (CaSR101- *dad2/PCK1pr-DAD2 RPS1/ RPS1 (DAD2pr-DAD2$^{\Delta DSS}$-TAP-NAT*)). The reintegrant strains, when grown in media containing dextrose, will shut down the expression of endogenous Dad2 and express only the protein A-tagged Dad2 (Dad2^{FL} in CaSR100 or Dad2$^{\Delta DSS}$ in CaSR101) from the *RPS1* locus (Figure 3-2B and 3-2C). Based on previous studies, the reintegrant strains were grown in dextrose for 8 h in all of the following experiments, to ensure maximum depletion of endogenous Dad2.

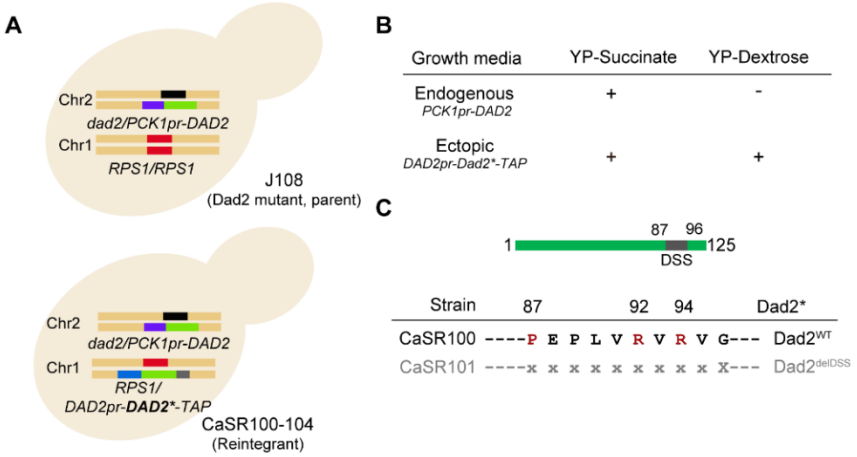


Figure 3-2. Schematic of the experimental setup to study DSS function in *C. albicans*.
(A) Schematic of Dad2 genomic locus (Chr2) in the parent strain J108 and the reintegrant strains in which Protein-A tagged Dad2$^{FL/\Delta DSS}$ will be expressed from the *RPS1* locus (Chr1). (B) Expression state of the endogenous Dad2 under the *PCK1* promoter as present in J108 (*dad2/PCK1prDAD2*) in the presence of succinate or dextrose as the carbon source is schematically shown. Conditional repression of endogenous Dad2 using dextrose in the reintegrant strains will facilitate the analysis of ectopic Dad2 (either FL/ΔDSS). (C) The green bar represents a linear form of Dad2, 125 amino acid (aa) long. Gray bar indicates the region corresponding to DSS (87-96 aa). The conserved proline and arginine residues

present in wild type Dad2 is shown below (red). CaSR100 and CaSR101 correspond to strains expressing Dad2^{FL} and Dad2$^{\Delta DSS}$ tagged with protein-A epitope respectively.

DSS is not essential for viability but delays cell cycle progression in *C. albicans*

The depletion of Dad2 has been shown to result in loss of viability in *C. albicans*. To study the role of the conserved DSS region in Dad2 function and its essentiality, we compared the ability of the strain CaSR100 (*dad2/PCK1pr-DAD2 RPS1/ RPS1 (DAD2pr-DAD2^{FL}-TAP-NAT*)) or CaSR101 (*dad2/PCK1pr-DAD2 RPS1/ RPS1 (DAD2pr-DAD2$^{\Delta DSS}$-TAP-NAT*)) generated as described above, to grow on the non-permissive media containing dextrose as the sole carbon source. The ability of both the strains to grow and form colonies when the expression of endogenous Dad2 was repressed suggested that the conserved DSS region is not essential for viability in *C. albicans* (Figure 3-3A). Phenotypic analysis of the CaSR101 cells expressing Dad2$^{\Delta DSS}$ upon depletion of endogenous Dad2 expressed from its native locus for 8 h revealed the presence of cells from all stages of the cell cycle, similar to the case of CaSR100 cells expressing Dad2^{FL} (Figure 3-3B). We observed a very modest increase in the number of cells at the large budded stage indicative of a transient arrest at metaphase or a delay in the cell cycle in these mutants. The distribution of cells at different budding stages was quantitated from the microscopic images and plotted (Figure 3-3C). In the case of the control strain J108EV (*dad2/PCK1pr-DAD2 RPS1/ RPS1(NAT)),* where all Dad2 alleles are repressed, about 90% of cells are in the large budded stage. In the case of CaSR100 cells expressing Dad2^{FL}, about 45% of cells are in this stage. However, in CaSR101 cells expressing Dad2$^{\Delta DSS}$, 64% of cells remained large budded.

Kinetochore integrity is unperturbed upon deletion of DSS in *C. albicans*

Depletion of outer kinetochore proteins resulted in the disintegration of the kinetochore ensemble, loss of CENP-A^{Cse4} localization, and its subsequent proteasomal degradation in *C. albicans* (THAKUR AND SANYAL 2012). This suggested that a clustered localization of any kinetochore protein was indicative of an intact kinetochore ensemble. To score for the integrity of the kinetochore ensemble, we tagged Ndc80 with mCherry at the C-terminus in the strains J108EV, CaSR100, and CaSR101 resulting in the strains J108N (*dad2/PCK1pr-DAD2 RPS1/ RPS1(NAT) NDC80/NDC80-mCherry),* CaSR200 (*dad2/PCK1pr-DAD2 RPS1/ RPS1 (DAD2pr-DAD2^{FL}-TAP-NAT) NDC80/NDC80-mCherry*), and CaSR201 (*dad2/PCK1pr-DAD2 RPS1/ RPS1(DAD2pr-DAD2$^{\Delta DSS}$-TAP-NAT) NDC80/NDC80-mCherry*) respectively. As reported previously, depletion of Dad2 for 8 h resulted in the loss of clustered localization of Ndc80 in the vector control strain J108N without any ectopic Dad2 expression (representative image in Figure 3-3D). Analysis of mean fluorescence intensity of Ndc80-mCherry per

kinetochore cluster in CaSR200 cells (expressing Dad2^{FL}, Ndc80-mCherry) and CaSR201 cells (expressing Dad2$^{\Delta DSS}$, Ndc80-mCherry) after depletion of endogenous Dad2 in non-permissive media for 8 h showed no significant difference (Figure 3-3D). Wild-type like localization of Ndc80-mCherry in the absence of DSS suggested that this region is not involved in the Dam1 complex assembly, which otherwise would have resulted in the loss of localization of other kinetochore proteins including Ndc80 (as observed for the vector control strain J108N). These observations suggest that the DSS is dispensable for the kinetochore integrity in *C. albicans*.

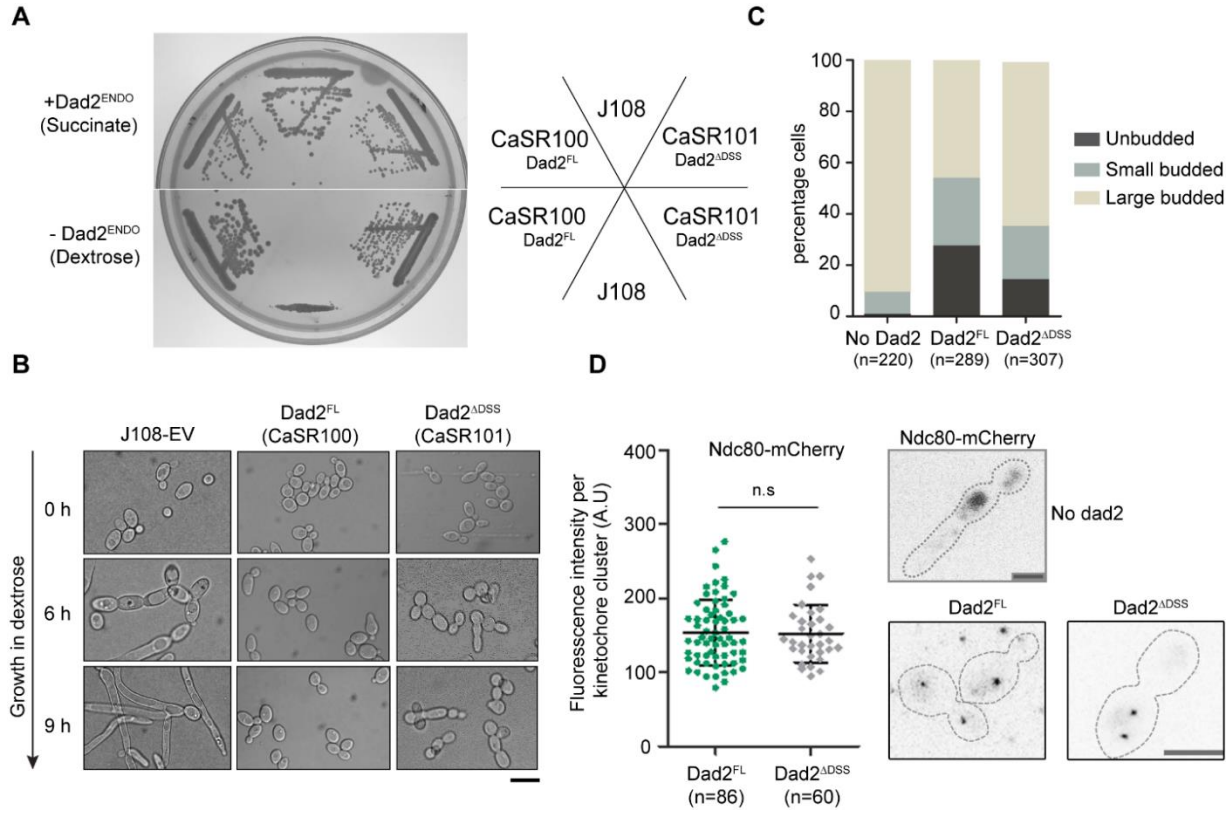


Figure 3-3. DSS is not essential for viability in *C. albicans*.
(A) Cells from the parent strain J108 along with reintegrant strains CaSR100 (Dad2^{FL}) and CaSR101 (Dad2$^{\Delta DSS}$) were streaked on plates containing media permissive and non-permissive for *PCK1* promoter regulating the expression of endogenous Dad2. Plates were imaged post-incubation at 30˚C for 48 h. (B) Light microscopic images of the indicated strains grown in the non-permissive medium for 6 h and 9 h after they were seeded from an overnight grown starter culture in permissive media (0 h). Scale bar, 5 µm. (C) Distribution of cells in different stages of budding (in percentage, *y*-axis) in strains J108, CaSR100, and CaSR101 when grown in YPDU for 6 h. (D) Kinetochore integrity in Dad2 mutants measured by the localization of Ndc80-mCherry in CaSR200 cells (Dad2^{FL}, Ndc80-mCherry) and CaSR201 (Dad2$^{\Delta DSS}$, Ndc80-mCherry) cells. Fluorescence intensity per kinetochore cluster in each strain is plotted, 'n.s.' indicates no significant difference. Representative images from both these strains are shown on the right. Scale bar, 5µm.

Deletion of DSS results in improper nuclear segregation

We next investigated the role of DSS in the process of nuclear segregation. We scored for the percentage of cells showing defective nuclear segregation when endogenous Dad2 was depleted for 8 h and complemented with ectopic expression of $Dad2^{FL}$ or $Dad2^{\Delta DSS}$ (Figure 3-4A). More than 90% of cells in the vector control strain J108EV (no ectopic Dad2) showed defective segregation in the form of nuclear mass stuck at the mother-daughter bud neck. Ectopic expression of $Dad2^{FL}$ in CaSR100 (*dad2/PCK1pr-DAD2 RPS1/ RPS1 (DAD2pr-DAD2^{FL}-TAP-NAT*)) rescued this effect, showing defective segregation in 21% of the large-budded cells. On the other hand, in the strain CaSR102 expressing $Dad2^{\Delta DSS}$ (*dad2/PCK1pr-DAD2 RPS1/ RPS1 (DAD2pr-DAD2$^{\Delta DSS}$-TAP-NAT*)), about 84% of the large-budded cells showed defective segregation, a four-fold increase in the frequency of defective segregation as compared to CaSR100. While an undivided nuclear mass at the bud-neck was predominantly observed, a stretched nucleus in the mother bud and unsegregated nucleus in the mother bud was also observed in the DSS mutants in 23% and 5% of the large budded cells respectively (Figure 3-4B). The increase in the nuclear missegregation frequency despite the presence of an intact kinetochore suggested that the kinetochore-microtubule interactions could be compromised in the DSS mutants.

DSS mutants show abnormal spindle structure

To understand the spindle structure in DSS mutants, we tagged the alpha-tubulin subunit (Tub1) of the microtubules with mCherry at the C-terminus in the strains CaSR100, CaSR101, and J108EV. We compared the spindle structure in resulting strains CaSR300 (*dad2/PCK1pr-DAD2, RPS1/RPS1 (DAD2pr-Dad2^{FL}-TAP-NAT), TUB1/TUB1-mCherry*), CaSR301 (*dad2/PCK1pr-DAD2, RPS1/RPS1(DAD2pr-Dad2$^{\Delta DSS}$-TAP-NAT), TUB1/TUB1-mCherry*), and J108T (*dad2/PCK1pr-DAD2, RPS1/RPS1(NAT) TUB1/TUB1-mCherry)*. CaSR300 cells (expressing $Dad2^{FL}$, Tub1-mCherry) showed a wild-type like anaphase spindle in large-budded cells (budding index >0.6) as described previously for *C. albicans*. The short spindle structure was present in 32% of the large budded cells in this strain (Figure 3-4C and 3-4D). However, CaSR301 cells (expressing $Dad2^{\Delta DSS}$, Tub1-mCherry) showed the short spindle phenotype in more than 85% of the large budded cells. The short spindle structure is classically associated with loss of kinetochore protein function as exemplified in studies involving conditional mutants of essential kinetochore proteins. The depletion of these proteins results in loss of contact points for the microtubules at the kinetochore locus, activating the SAC due to unattached kinetochore. Given an intact kinetochore in the DSS mutants, these observations strongly support our hypothesis that DSS might play a key role in mediating attachment of the kinetochore to microtubules.

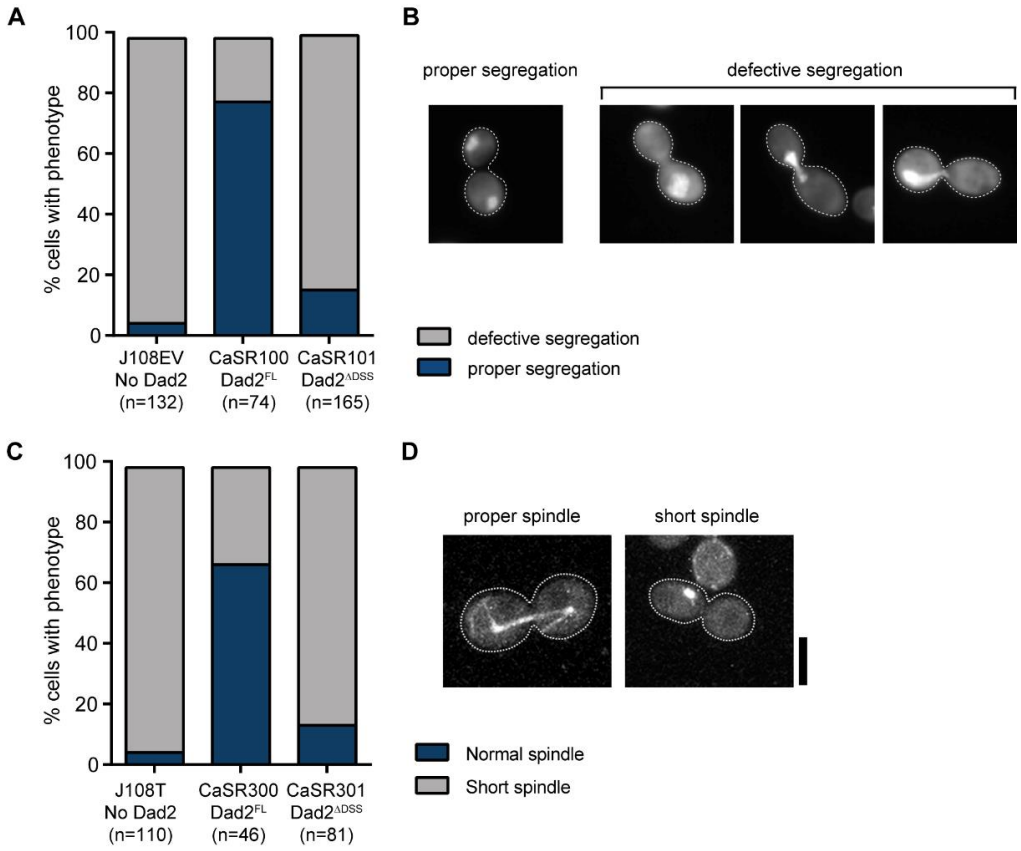


Figure 3-4. The deletion of DSS results in an increase in cells with nuclear missegregation and short spindle structure.
(A) J108EV, CaSR100 (Dad2^{FL}), and CaSR101 (Dad2$^{\Delta DSS}$) strains grown in YPSU for 12 h were washed and re-inoculated into YPDU at 0.2 OD. Cells were harvested after 8 h growth in YPDU at 30˚C, stained with DAPI (nucleus), and imaged under a fluorescence microscope. The percentage of cells showing defective segregation in each case is plotted (*y*-axis). (B) Representative images of nuclear segregation types that were considered normal and defective are shown. (C) CaSR300 (Dad2^{FL}, Tub1-mCherry) and CaSR301 (Dad2$^{\Delta DSS}$, Tub1-mCherry) cells grown as described in (A) were harvested, washed, and examined under a confocal laser scanning microscope. The percentage of large-budded cells with budding index >0.6 showing short spindle structure in each strain is plotted (*y*-axis). (D) Representative images of spindle structures that were considered proper and short are shown. Scale bar, 5µm.

The conserved arginine residues in DSS are essential for proper nuclear segregation and maintenance of spindle structure

It is known that conserved positively charged residues mediate interactions between outer kinetochore complex proteins containing these residues and the spindle microtubules. While no interaction between Dad2 and the microtubules have been mapped to date, we observed two conserved

arginine residues (R92 and R94) along with a conserved proline residue (P87) in DSS. To dissect the contribution of these conserved residues in Dad2 function, we generated reintegration strains that independently contain the following mutations in Dad2- P87A, R92A, and R94A. Similar to the DSS deletion mutant, we generated strains expressing Dad2 with the above mutations from an ectopic *RPS1* locus (Figure 3-5A).

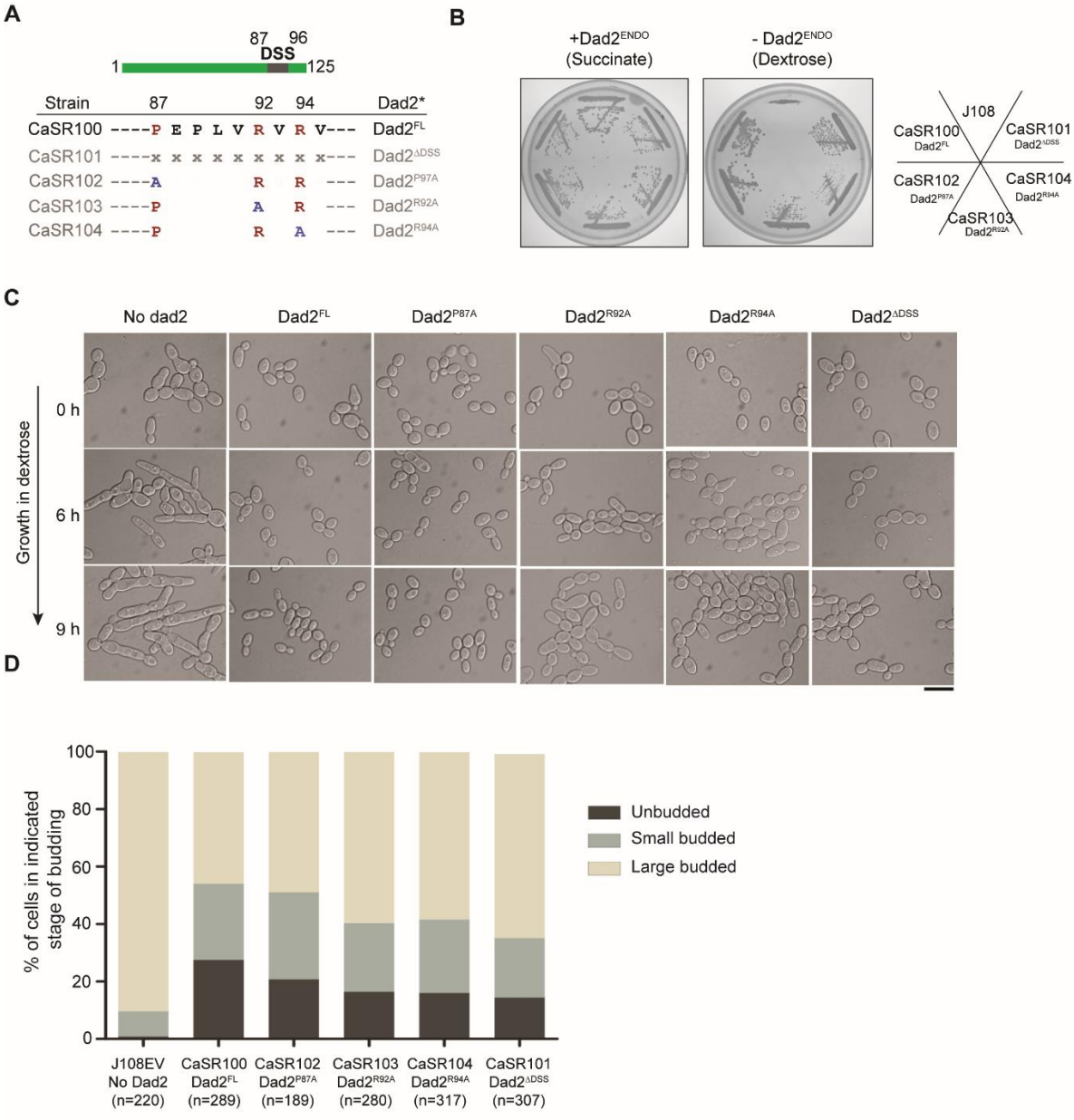


Figure 3-5. Mutations in conserved residues of the DSS does not affect the viability of *C. albicans*.

(A) Schematic of the mutations in the DSS analyzed in this study. The green bar represents a linear form of Dad2, 125 aa long. Gray bar indicates the region corresponding to DSS (87-95 aa). The conserved proline and arginine residues present in wild-type Dad2 is shown below (red). Each of these residues was individually mutated to alanine (blue). (B) Cells from the parent strain J108 along with reintegrant strains CaSR100 ($Dad2^{FL}$), CaSR101 ($Dad2^{\Delta DSS}$), CaSR102 ($Dad2^{P87A}$), CaSR103 ($Dad2^{R92A}$), and CaSR104 ($Dad2^{R94A}$) were streaked on plates containing media permissive and non-permissive for the *PCK1* promoter regulating the expression of endogenous Dad2. Plates were imaged post-incubation at 30˚C for 48 h. (C) Light microscopic images of the cells expressing indicated versions of Dad2 grown in the non-permissive medium for 6 h and 9 h after they were seeded from a starter culture in permissive media (0 h). (D) Distribution of cells in different stages of budding (in percentage, *y*-axis) in strains expressing the indicated versions of Dad2 after growth in YPDU at 30˚C for 6 h.

Strains expressing each of these point mutants- CaSR102 (*dad2/PCK1pr-DAD2, RPS1/RPS1(DAD2pr-DAD2^{P87A}-TAP-NAT*)), CaSR103 (*dad2/PCK1pr-DAD2, RPS1/RPS1(DAD2pr-DAD2^{R92A}-TAP-NAT*)), and CaSR104 (*dad2/PCK1pr-DAD2, RPS1/ RPS1(DAD2pr-DAD2^{R94A}-TAP-NAT)*) were able to grow on media containing dextrose, non-permissive for the expression of endogenous Dad2 (Figure 3-5B). Analysis of these mutants repressed for 9 h revealed a modest increase in the fraction of cells in the large-budded stage in CaSR103 and CaSR104 cells (expressing $Dad2^{R92A}$ and $Dad2^{R94A}$ respectively), as observed in CaS101 cells (expressing $Dad2^{\Delta DSS}$) (Figure 3-5C and 3-5D).

To dissect the contribution of the conserved proline and arginines in the defects observed in CaSR101 cells (expressing $Dad2^{\Delta DSS}$), we compared the frequency of nuclear missegregation in CaSR102-104 cells (expressing $Dad2^{P87A}$, $Dad2^{R92A}$, and $Dad2^{R94A}$ respectively) to that observed in CaSR100 (expressing $Dad2^{FL}$) and CaSR101 cells (expressing $Dad2^{\Delta DSS}$). Upon depletion of endogenous Dad2 for 8 h by growing in media containing dextrose, the frequency of nuclear missegregation was comparable between CaSR100 (expressing $Dad2^{FL}$) and CaSR102 cells (expressing $Dad2^{P87A}$) at 18% and 21% respectively. On the contrary, the frequency of defective segregation in CaSR103 (expressing $Dad2^{R92A}$) and CaSR104 cells (expressing $Dad2^{R94A}$), estimated to be 78% and 91% respectively was comparable to that of CaSR101 cells (expressing $Dad2^{\Delta DSS}$), strongly suggesting a role for the arginine residues in DSS function (Figure 3-6A).

To score for the kinetochore integrity in the Dad2 point mutant strains, we tagged Ndc80 with mCherry at the C-terminus in CaSR102, CaSR103, and CaSR104 to generate strains CaSR202 (*dad2/PCK1pr-DAD2, RPS1/RPS1(DAD2pr-DAD2^{P87A}-TAP-NAT), NDC80/NDC80-mCherry*), CaSR203 (*dad2/PCK1pr-DAD2, RPS1/RPS1(DAD2pr-DAD2^{R92A}-TAP-NAT), NDC80/NDC80-mCherry*), and CaSR204 (*dad2/PCK1pr-DAD2, RPS1/RPS1(DAD2pr-DAD2^{R94A}-TAP-NAT), NDC80/NDC80-mCherry*). As observed in the case of DSS deletion strain CaSR201 (expressing $Dad2^{\Delta DSS}$, Ndc80-mCherry), none of the point mutants showed significantly altered levels of Ndc80-mCherry when grown in media non-permissive for the expression of endogenous Dad2 for 8 h. An intact kinetochore ensemble in CaSR202-

204 cells suggested that any of the point mutations in the DSS (P87A, R92A, or R94A) did not hamper kinetochore assembly/ integrity (Figure 3-6B and 3-6C).

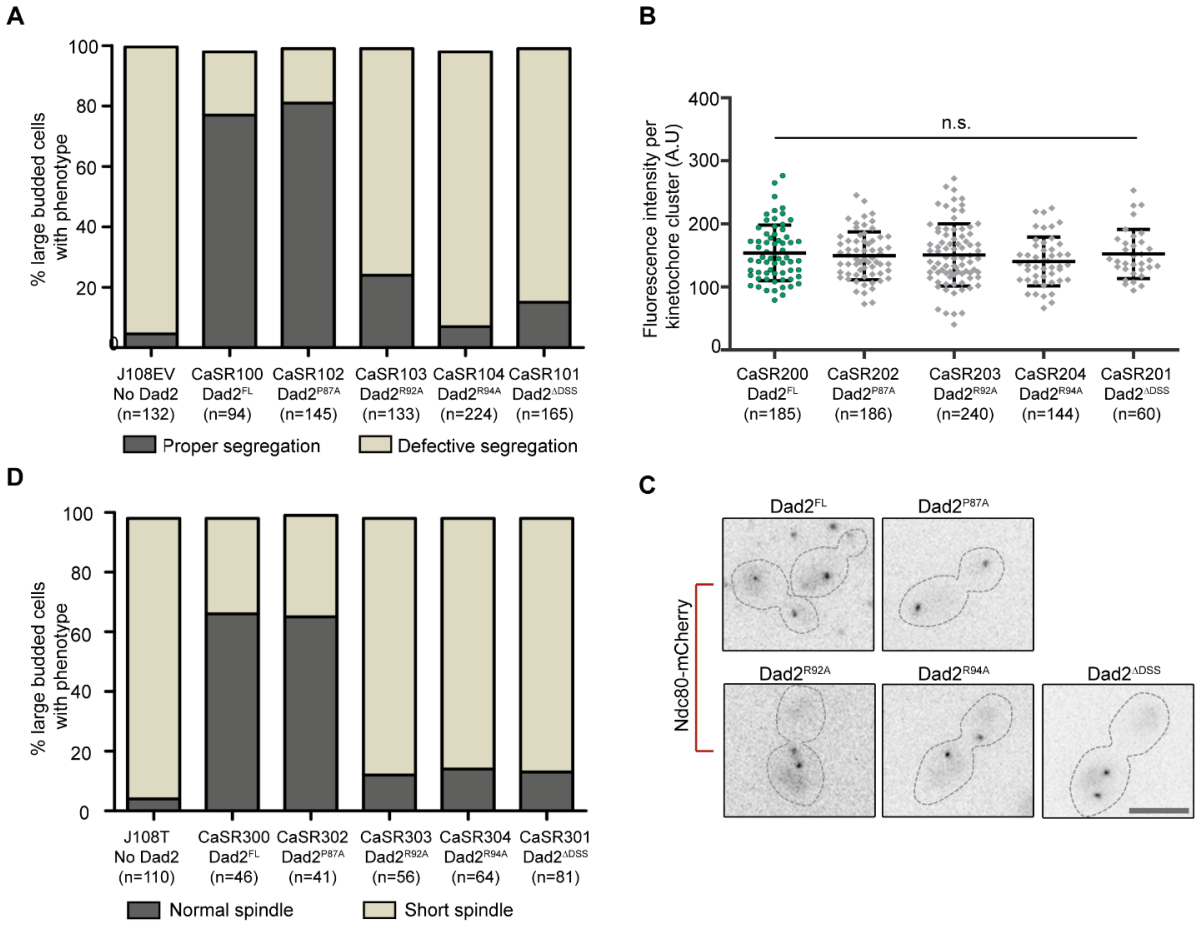


Figure 3-6. Mutations in R92 and R94 results in an increase in cells with nuclear missegregation and short spindle structure.
(A) Kinetochore integrity in Dad2 point mutants CaSR202 (Dad2^{P87A}, Ndc80-mCherry), CaSR203 (Dad2^{R92A}, Ndc80-mCherry), CaSR204 (Dad2^{R94A}, Ndc80-mCherry) in comparison with CaSR200 (Dad2^{FL}, Ndc80-mCherry) and CaSR201 (Dad2$^{\Delta DSS}$, Ndc80-mCherry), measured by the localization of Ndc80-mCherry after growth in YPDU at 30˚C for 8 h. Fluorescence intensity per kinetochore cluster in each strain is plotted, 'n.s' indicates no significant difference. (B) Representative images of Ndc80 localization in each of these strains is shown. Scale bar, 5µm. (C) The indicated strains ectopically expressing various versions of Dad2 were harvested after 8 h growth in YPDU at 30˚C. The nuclear mass was visualized under a fluorescence microscope after staining with DAPI. The percentage of cells showing defective segregation in each case is plotted (*y*-axis). (D) The indicated strains ectopically expressing various versions of Dad2 along with mCherry-tagged Tub1 were harvested after growth in YPDU at 30˚C for 8 h. Cells were washed, examined under a confocal laser scanning microscope, and the percentage of cells showing the short spindle structure in each strain is plotted (*y*-axis).

We then analyzed the spindle structure by localizing Tub1-mCherry in CaSR302 (*dad2/PCK1pr-DAD2, RPS1/RPS1(DAD2pr-DAD2^{P87A}-TAP-NAT), TUB1/TUB1-mCherry*), CaSR303 (*dad2/PCK1pr-DAD2, RPS1/RPS1(DAD2pr-DAD2^{R92A}-TAP-NAT), TUB1/TUB1-mCherry*), and CaSR304 (*dad2/PCK1pr-DAD2, RPS1/RPS1(DAD2pr-DAD2^{R94A}-TAP-NAT), TUB1/TUB1-mCherry*) strains that expressed the above point mutants. Similar to the trend observed for nuclear segregation, wildtype-like anaphase spindle was detected in the strains CaSR300 (expressing Dad2^{FL}, Tub1-mCherry) and CaSR302 (expressing Dad2^{P87A}, Tub1-mCherry) respectively. On the other hand, the arginine mutants CaSR103 (expressing Dad2^{R92A}, Tub1-mCherry) and CaSR104 (expressing Dad2^{R94A}, Tub1-mCherry) showed short spindle phenotype similar to that of the DSS deletion strain (CaS301) in more than 84% of large budded cells (Figure 3-6D). The above mutation analysis strongly suggests that the conserved positive charges in the form of R92 and R94 mediate DSS function, which we propose to be the microtubule-binding domain of Dad2 in *C. albicans*. To investigate if CaDad2 can indeed bind to microtubules, we resorted to an *in vitro* biochemical approach.

CaDad2 binds to microtubules in a DSS-dependent manner *in vitro*

To quantitatively estimate the ability of Dad2 to bind microtubules, we made a series of constructs in the pET28b expression vector system such that the following versions of Dad2 can be expressed as a 6xHis fusion protein in *E. coli* and subsequently purified by Ni-NTA based affinity chromatography. The constructs include Dad2^{FL}, Dad2^{P87A}, Dad2^{R92A}, Dad2^{R94A}, Dad2^{R92AR94A}, and Dad2$^{\Delta DSS}$. The expression and affinity purification conditions were optimized for Dad2^{FL} and used for other mutants of Dad2 (described in the methods section).

We estimated the dissociation constant for the interaction of CaDad2 and taxol-stabilized microtubules by a microtubule co-sedimentation assay. For a fixed concentration of Dad2^{FL} (50µM), we titrated varying concentrations of taxol-stabilized microtubules (0- 30 µM in 5 µM increments). The presence of Dad2 in the pellet fraction provided the first evidence that CaDad2^{FL} can bind to microtubules *in vitro*. Based on the amount of Dad2 in the pellet fraction that co-sedimented with microtubules as compared to the amount of Dad2 in the supernatant (free, unbound Dad2), we estimated the dissociation constant (K_d) to be 4.9 µM for Dad2^{FL} (Figure 3-7). The K_d value for Dad2^{P87A} and Dad2^{FL} were similar, hinting at no loss in binding when the conserved proline was mutated. On the other hand, a gradual reduction in the binding efficiency was seen when the arginines were replaced with alanines independently or as a double point mutant. Mutant Dad2 proteins- Dad2^{R92A}, Dad2^{R94A}, and Dad2^{R92AR94A} showed three to four-fold less efficient binding as compared to full-length protein with K_d values 14.4 µM, 16.8 µM, and 19.1 µM respectively. The deletion of DSS exhibited the maximum reduction in

binding (K_d=24 µM), a five-fold reduction in binding as compared to $Dad2^{FL}$. We demonstrate, for the first time, the microtubule-binding ability of Dad2, with its binding efficiency governed by the conserved arginines. This strongly suggests that the DSS forms a microtubule-binding domain of Dad2. The microtubule co-sedimentation assay was performed by Ms. Sujaya Thannimangalath from Prof. Tapas Manna's laboratory (IISER, Thiruvananthapuram) as a part of a collaboration on this part of the work.

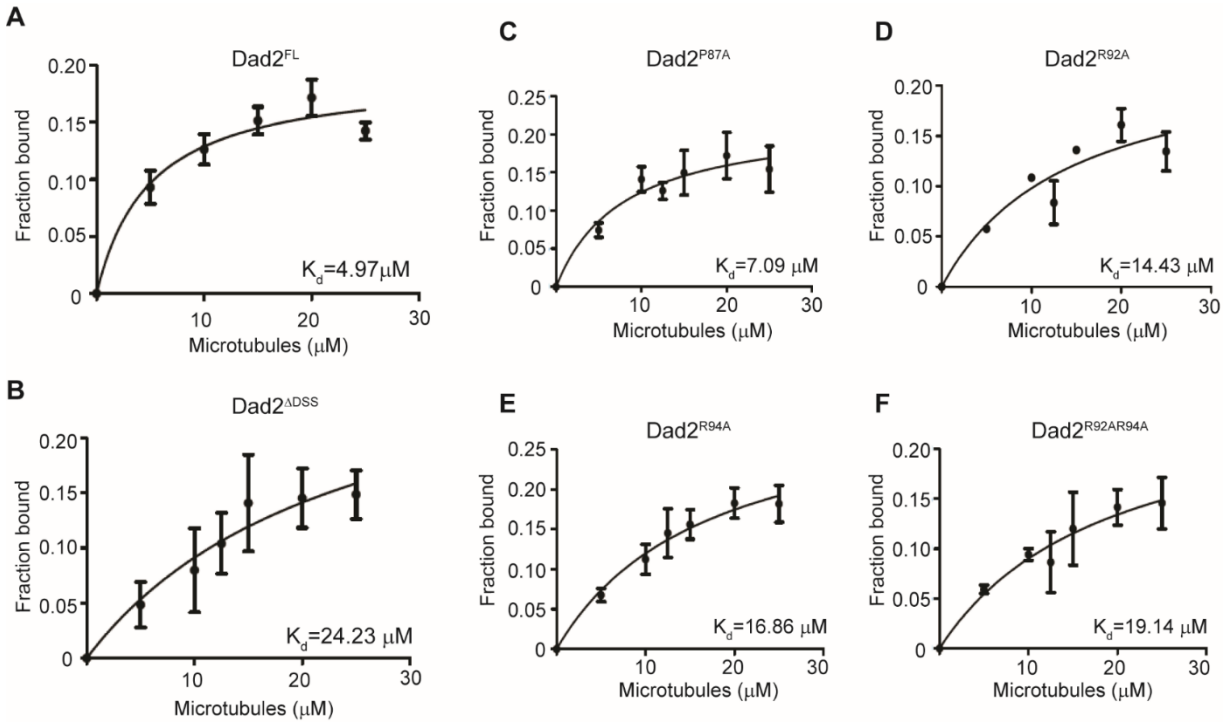


Figure 3-7. CaDad2 binds taxol-stabilized microtubules *in vitro*, and the binding is mediated by DSS.
(A-F) The dissociation constant of the interactions between various versions of CaDad2 purified *in vitro* and microtubules were estimated by co-sedimentation assay. Graphs indicate the fraction of Dad2 bound to microtubules (*y*-axis, fraction bound) estimated from coomassie stained SDS-PAGE gels loaded with sedimented microtubules and the associated Dad2 along with the unbound Dad2. The *x*-axis indicates varying concentrations of microtubules tested in the assay.

DSS function is conserved across point and regional centromeres

Previous *in vitro* studies on the interaction network of ScDam1 complex- both within the complex and with the microtubules, did not detect any contact between Dad2 and microtubules (LEGAL *et al.* 2016; JENNI AND HARRISON 2018). However, the DSS region we identified and showed to mediate microtubule binding in CaDad2 is conserved in ScDad2 as well. This prompted us to investigate if the function of DSS is also conserved in *S. cerevisiae* that contains point centromeres.

To study the role of DSS in *Sc*Dad2 function, we resorted to a previously reported temperature-sensitive (ts) mutant of Dad2 wherein the genomic allele of Dad2 was deleted and a ts allele was present in a *CEN* plasmid pCJ055 (CJY077- *dad2::KanMX6 his3Δ200 leu2Δ1::pCJ055 (Dad2^{TS})*) (JANKE *et al.* 2002). In this strain, we independently transformed four pRS313 (*CEN/ARS/HIS)*-based plasmids in which the various versions of Dad2 was cloned with a GFP tag at the C-terminus. The transformed constructs include pScDad2^{FL}, pScDad2^{R126A} (equivalent to CaDad2^{R92A}), pScDad2^{R128A} (equivalent to CaDad2^{R94A}), and pScDad2$^{\Delta DSS}$ giving rise to strains ScSR01 (*dad2::KanMX6 his3Δ200::pScDad2^{FL} leu2Δ1::pCJ055 (Dad2^{TS})*), ScSR02 (*dad2::KanMX6 his3Δ200::pScDad2^{R126A} leu2Δ1::pCJ055 (Dad2^{TS})*), ScSR03 (*dad2::KanMX6 his3Δ200::pScDad2^{R128A} leu2Δ1::pCJ055 (Dad2^{TS})*), and ScSR04 (*dad2::KanMX6 his3Δ200::pScDad2$^{\Delta DSS}$ leu2Δ1::pCJ055 (Dad2^{TS})*). In all of these strains, the expression of Dad2-GFP is driven by the native promoter of ScDad2. An empty vector pRS313G that contains only GFP cloned in a pRS313 plasmid served as the negative control (strain ScEV). Once these strains were obtained (See materials and methods), they were subsequently propagated in CM-Leu-His media, and the functional significance of each of these residues was assayed for at 37˚C (Figure 3-8A and 3-8B).

We first tested if any of the Dad2 constructs we generated could grow at 37˚C by spotting cells at different dilutions. Growth at 26 ˚C in the same media was used as the loading control for spotting (Figure 3-8C). As expected, the empty vector carrying control strain (ScEV) remained sensitive to 37˚C, and the strain ScSR01 expressing Dad2^{FL} showed complementation of growth at this temperature. We observed that the expression of Dad2^{R128A} in ScSR03 cells also supported growth at 37˚C. However, the strains ScSR02 (expressing Dad2^{R126A}) and ScSR04 (expressing Dad2$^{\Delta DSS}$) failed to grow at 37˚C, suggesting that DSS, more specifically R126, is essential for Dad2 function in *S. cerevisiae*.

To understand the basis of the lack of complementation of growth at 37˚C in ScSR02 and ScSR04 strains, we analyzed the localization of Dad2-GFP in strains ScSR01-04 at 37˚C (Figure 3-8D). In the strains ScSR01 (expressing Dad2^{FL}-GFP) and ScSR03 (expressing Dad2^{R128A}-GFP), a punctate localization typical of clustered kinetochores in yeasts was observed for Dad2. We could detect a similar localization of Dad2^{R126A} in ScSR02 cells as well; however, the proportion of cells showing this localization was reduced. This can be ascribed to the loss of the *CEN* plasmid harboring this allele, which is also evident from the growth observed in the permissive temperature (28˚C). However, no punctate localization of Dad2$^{\Delta DSS}$ could be detected at 37˚C. This was striking as the deletion of DSS did not hamper kinetochore integrity in *C. albicans*. Probing the whole-cell extracts from ScSR01 and ScSR04 cells expressing Dad2^{FL} and Dad2$^{\Delta DSS}$ with anti-GFP antibodies confirmed that these proteins were indeed expressed (Figure 3-8E). This suggested that DSS plays a conserved but more critical role in Dad2 function in *S. cerevisiae* as compared to *C. albicans*. This difference can be attributed to the difference in

the localization dynamics of the Dam1 complex in these two organisms with different types of centromeres.

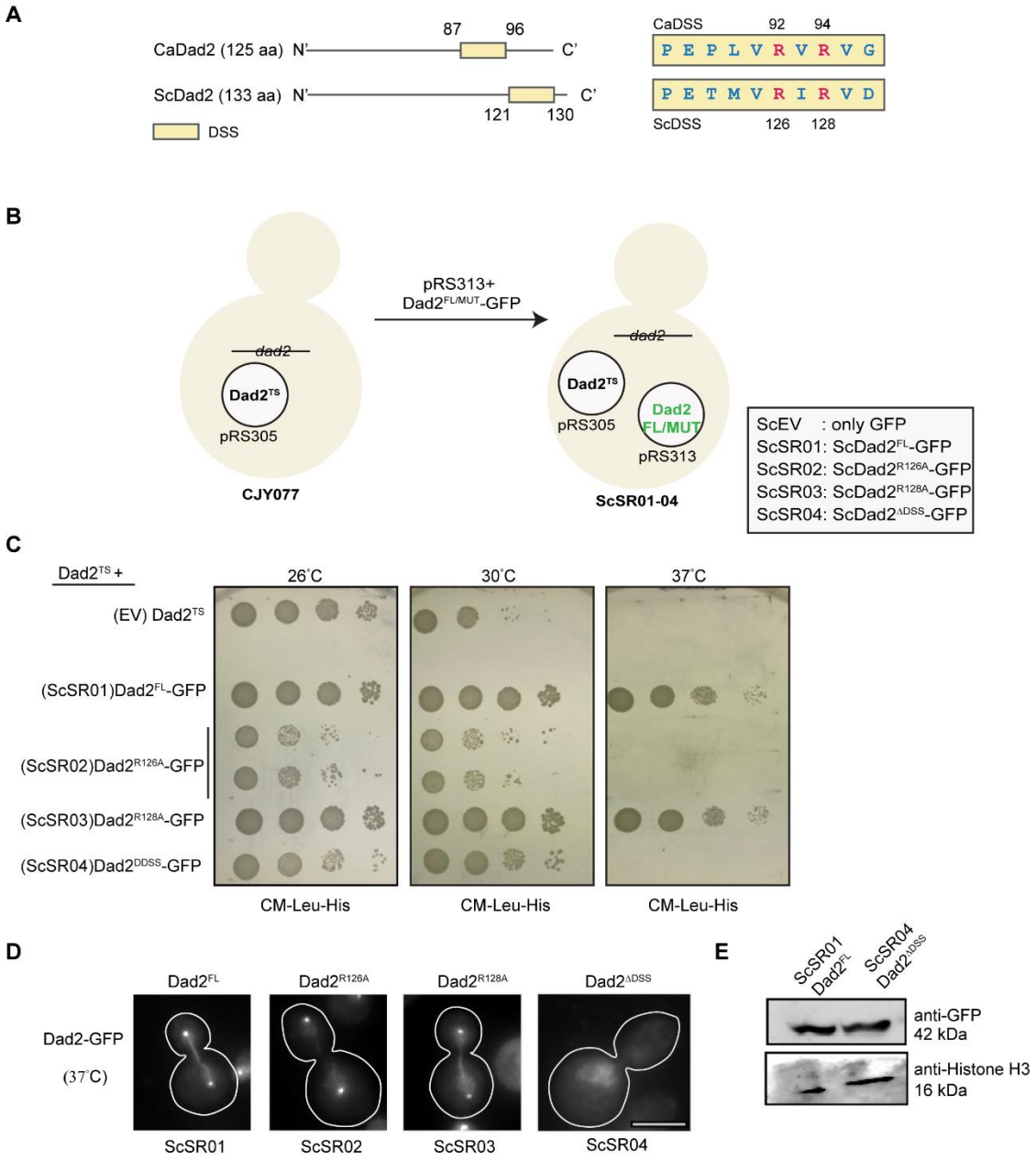


Figure 3-8. DSS is required for kinetochore localization of Dad2 in *S. cerevisiae*.
(A) Line diagrams of CaDad2 and ScDad2 indicating the position of DSS in the respective sequences are shown. The amino acid sequence of DSS depicting the arginine residues (red) conserved in both species is shown in the right panel. (B) Schematic of the strains developed to assay for DSS function in *S. cerevisiae*. Wild-type Dad2 or Dad2 with mutations in the DSS (121-130 aa in ScDad2) as

mentioned in the box were cloned into pRS313 plasmid along with a GFP tag at C-terminus and were used to transform parent strain CJY077 to yield strains ScSR01-04. ScEV corresponds to vector control strain carrying empty plasmid without Dad2. (C) Single colonies of the indicated strains were grown on CM-Leu-His media at 26˚C for 14 h, serially diluted ten-fold and spotted (10^5 to 10^2) on three CM-Leu-His plates incubated at 26˚C, 30˚C, and 37˚C respectively. Plates were photographed after 48 h incubation. (D) Microscopic images showing the localization of different versions of Dad2 expressed respectively in ScSR01-ScSR04 cells grown at 37˚C for 4 h. Scale bar, 5μm. (E) Expression of $Dad2^{FL}$ and $Dad2^{\Delta DSS}$ in whole-cell extracts prepared from ScSR01 and ScSR04 cells grown at 26˚C in CM-Leu-His media probed with anti-GFP and anti-Histone H3 antibodies.

Discussion

One of the major challenges for the kinetochore ensemble is to remain associated with the depolymerizing plus ends of the microtubules. This is achieved by a cumulative effect of several points of contact by the Ndc80 complex and the Dam1 complex with the microtubules (WEI *et al.* 2005; WESTERMANN *et al.* 2005; CHEESEMAN *et al.* 2006; WESTERMANN *et al.* 2006; CIFERRI *et al.* 2008). Among the Dam1 complex subunits, two proteins Dam1 and Duo1, have been reported to directly interact with microtubules in *S. cerevisiae* (LEGAL et al. 2016). Consistently, the *dam1-19* temperature-sensitive mutant show short spindle phenotype, and truncation at the C-terminus of Dam1 or Duo1 compromised the microtubule-binding of Dam1 complex carrying these mutations (WESTERMANN *et al.* 2005; LEGAL *et al.* 2016).

In this part of the thesis work, we identified a conserved 10-amino acid long Dad2 signature sequence (DSS) present in otherwise poorly conserved Dad2 family of proteins. Functional analysis of this region in CaDad2 revealed that the DSS is a previously unknown microtubule-binding domain of Dad2 in *C. albicans*. Previous studies on the conditional mutants of the Dam1C subunits, including Dad2 in *C. albicans* revealed that depletion of these proteins results in the disintegration of the kinetochore ensemble, followed by proteasomal degradation of $CENPA^{Cse4}$ (THAKUR AND SANYAL 2012). This also resulted in the activation of SAC, resulting in cell cycle arrest at the large budded stage with the unsegregated nuclear mass stuck at the bud neck (THAKUR AND SANYAL 2011). However, the deletion of DSS or point mutations in the conserved residues of DSS did not show any compromise in the kinetochore integrity. A modest increase in the proportion of large budded cells was observed. The occurrence of short spindle structures in these cells, along with defective nuclear segregation, suggested that the kinetochore-microtubule interactions were affected upon perturbing the DSS. The presence of cells at all stages of budding after depletion of endogenous Dad2 for 8 h suggests that the modest increase in the cells with large budded cells could be due to a transient delay/arrest, providing the additional time required for the DSS mutants to establish proper kinetochore-microtubule attachments. Analysis of DSS mutants in a SAC deficient genetic background, such as *mad2* mutant, will provide further insights into this aspect.

In vitro assays using purified full-length or DSS-mutated versions of Dad2 protein revealed that the DSS and the resident arginines are critical for Dad2-microtubule interactions (Figure 3-9). This is also suggestive of the electrostatic nature of these interactions with microtubules, as observed for other microtubule-binding proteins (ABAD *et al.* 2014). The presence of conserved arginine residues in the DSS is reminiscent of the microtubule-binding domains identified in the functionally analogous Ska1 complex subunits (SCHMIDT *et al.* 2012; ABAD *et al.* 2014; ABAD *et al.* 2016). The microtubule-binding domain of Ska1 ($Ska1^{133-215}$) is enriched with positively charged amino acids lysine and arginine that form discrete clusters making multiple contacts with microtubules. The DSS, with a predicted pI >9, can form a similar

positively charged surface on Dad2. However, no interaction between ScDad2 and microtubules were identified in the previous chemical crosslinking based assays (ZELTER *et al.* 2015; LEGAL *et al.* 2016; JENNI AND HARRISON 2018). This could partly be due to the nature of the chemical crosslinker used in the assay (EDC: 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride) that does not crosslink arginine residues. Indeed, interactions between arginines present in Ska1 and microtubules were detected using EDC after the arginines were mutated to lysine, which is a compatible substrate of EDC (ABAD *et al.* 2014).

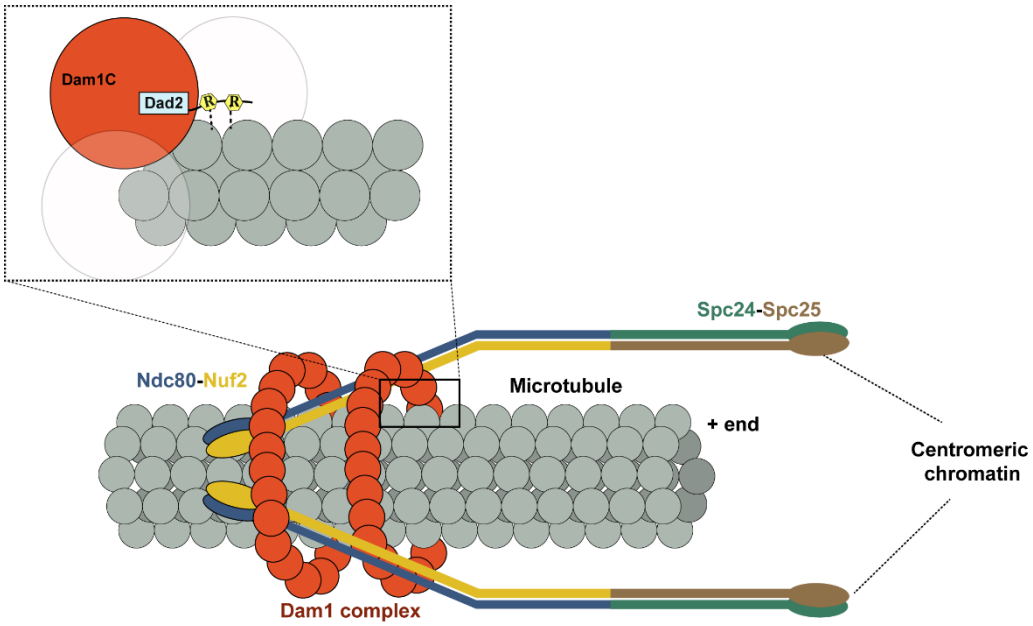


Figure 3-9 Schematic representation of Dad2-microtubule interaction mediated by DSS.
The microtubule binding protein complexes of the outer kinetochore- Ndc80 complex and the Dam1 complex are schematically represented. Red spheres represent the Dam1 complex monomers encircling the microtubules. Inset: The interaction between Dad2 and microtubules identified in this study is schematically represented in the inset. The blue box represents the N terminal ordered domain of Dad2 and the line emanating from the box represents the disordered C-terminus of Dad2 in which the DSS motif is present. The two arginine residues analyzed in this study is proposed to interact with microtubules (indicated in dashed lines).

Comparative analyses of DSS function between *C. albicans* and *S. cerevisiae* revealed a striking difference for the essentiality of this conserved region. While the kinetochore integrity is unaffected in *C. albicans*, the deletion of DSS resulted in the loss of kinetochore localization of Dad2 in *S. cerevisiae*. This variation stems from the fact that the interaction with mitotic spindle is essential for loading the Dam1C to kinetochores in *S. cerevisiae* (LI *et al.* 2002), unlike the case in *C. albicans* (THAKUR AND SANYAL 2011). Experimental perturbation by treatment with nocodazole, a microtubule poison, resulted in the loss of enrichment of Dam1C subunits- Dad2 and Ask1 at the centromeres (JANKE *et al.* 2002; LI *et*

al. 2002) in *S. cerevisiae*. Our results suggest that the interaction between Dad2 and microtubules mediated by DSS could be essential for loading the Dam1C in *S. cerevisiae*. Further studies on the localization of other kinetochore subunits will shed light on the role of DSS in Dad2 function in this species.

Analysis of protein disorder predicts that the N-terminus of Dad2 (1-85 aa) shows a propensity to form an ordered structure while the C-terminus is disordered. The presence of DSS is the disordered region of Dad2 makes it poised for interaction with other proteins, including microtubules. Several examples of protein-protein interactions within the kinetochore complex involve flexible and disordered regions of proteins (DIMITROVA *et al.* 2016; JENNI *et al.* 2017). Within the ScDam1C, interactions with the Ndc80C has been mapped to the disordered C-terminus of Ask1, Dam1, and Spc34 (JENNI AND HARRISON 2018). The known microtubule-binding regions of the Dam1 complex are also mapped to the flexible ends of Dam1 and Duo1 in *S. cerevisiae*. Flexible protein termini are favorable in creating interactions that are dynamic and requires adaptation to the varying configurations, like that between the Dam1 complex and microtubules at different stages of mitosis.

All known microtubule-binding proteins of the kinetochore, such as Dam1, Duo1, and Ndc80, are phospho-regulated by Ipl1 (CHEESEMAN *et al.* 2002; LAMPERT *et al.* 2010; TIEN *et al.* 2010). The microtubule-binding domain of Ska1 contains Aurora B phosphorylation sites consistent with the findings in yeast (CHAN *et al.* 2012b; ABAD *et al.* 2016). Intriguingly, no Ipl1 consensus sites are detected in DSS or Dad2 of both the species analyzed in this study. This alludes to several possibilities regarding the physiological relevance of the observed Dad2-microtubule interaction. The low K_d reported for Dad2-microtubule association could be attributed to the use of an isolated protein as compared to the entire Dam1C. Probably this motif forms an additional low-affinity contact point with the microtubules creating an additive effect when taken along with other known interactions. In such a scenario, the phosphorylation of other subunits like Dam1 and Duo1 by Ipl1 can significantly reduce the binding affinity towards microtubules, satisfying the SAC response. Alternatively, it should also be noted that Dad2 was found to localize to the spindle mid-zone in *C. albicans* (THAKUR AND SANYAL 2011). Regulation of microtubule dynamics at the spindle mid-zone is a critical determinant of spindle length (KHMELINSKII AND SCHIEBEL 2008). It is reasonable to assume that DSS is essential for the interaction of Dad2 with the microtubules at the spindle midzone, outside the context of the kinetochore cluster.

Although we still lack a mechanistic understanding of DSS function, it is evident that the region is required for high fidelity chromosome segregation. The fungus specific nature of the Dam1 complex, along with the conservation of DSS residues in several animal and plant pathogenic fungi, makes it an attractive target for the development of potent antifungals.

Chapter 4

Materials and Methods

Buffers and solutions used

All of the buffers and solutions used in this study were prepared in autoclaved double distilled water with analytical grade reagents. Solutions were stored at room temperature unless stated otherwise

1. TE Buffer (10x)- 100mM Tris (pH=8.0 or 7.5 as required for the assay), 10mM EDTA
2. LiOAc/TE solution- 0.1 M LiOAc (pH=7.5), 1x-TE (pH=7.5)
3. PEG solution- 0.1 M LiOAc (pH=7.5), 1xTE (pH=7.5), 42 %(w/v) PEG3350
4. Phosphate buffered saline (10x)- 80 g NaCl, 2 g KCl, 14.4 g Na_2HPO_4, 2.4 g KH_2PO_4 for 1L water (pH=7.4)
5. Genomic DNA extraction buffer for C. albicans- 2% TritonX-100, 1% SDS, 100 mM NaCl, 10 mM Tris-Cl (pH=8.0), 1 mM EDTA
6. Genomic DNA extraction buffer for Malassezia species- 100 mM Tris-Cl (pH=8.0), 1% (w/v) CTAB, 0.7M NaCl, 10 mM EDTA, 1% (v/v) 2-mercaptoethanol*, 0.3 mg/mL Proteinase-K*. ('*' added fresh)
7. Cell lysis buffer- 0.1 N NaOH, 1% SDS
8. SDS sample loading buffer (5x)- 30% Glycerol, 10% SDS, 250mM Tris-Cl (pH=6.8), 0.02% Bromophenol blue, 5% 2-mercaptoethanol*. ('*' added fresh)
9. SDS-PAGE tank buffer (10x)- 30 g Tris base, 144 g glycine, 10 g SDS for 1L buffer.

Buffers for chromatin immunoprecipitation (Buffers 10- 17)

10. Spheroplasting buffer for Malassezia- 40 mM Citric acid, 120 mM Na_2HPO_4 and 1.2 M Sorbitol supplemented with 20 mg/mL lysing enzymes from Trichoderma harzianum (Sigma), 0.2 µg/mL Chitosanase (Sigma), and 100 µg/mL Zymolyase-20T (MP Biomedicals)
11. Buffer I- 0.25% Triton X-100, 10 mM EDTA, 0.5 mM EGTA, 10 mM Na-HEPES pH=6.5
12. Buffer II- 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 10 mM Na-HEPES pH=6.5
13. Lysis buffer- 50 mM HEPES pH=7.4, 1% Triton X-100, 140 mM NaCl, 0.1% Na-deoxycholate, 1 mM EDTA supplemented with 1xPIC (Protease Inhibitor Cocktail); always freshly prepared
14. Low salt wash buffer- 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH=8.0, 150 mM NaCl (storage at 4˚C)

15. High salt wash buffer- 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH=8.0, 500 mM NaCl (storage at 4˚C)
16. LiCl wash buffer- 0.25 M LiCl, 1% NP-40, 1% Na-deoxycholate, 1 mM EDTA, 10 mM Tris pH=8.0 (storage at 4˚C)
17. Elution buffer- 0.1 M NaHCO3, 1% SDS (always freshly prepared)
18. CloNAT Nourseothricin- 100 mg/mL solution in autoclaved double distilled water
19. Ampicillin- 100 mg/mL solution in autoclaved double distilled water
20. Kanamycin- 30 mg/mL solution in autoclaved double distilled water
21. DAPI- 100 ng/mL in 70% glycerol
22. Hoechst 33342- 1 mg/mL solution in water from a 10 mg/mL stock

Antibodies and affinity beads

Resource	Description	Source	Catalog no.	Additional information
Antibody	anti-GFP (mouse monoclonal)	Roche	11814460 001	IF (1:100) WB (1:3000)
Antibody	anti-PSTAIRE (mouse monoclonal)	Abcam	Cat. no. 10345	WB (1:5000)
Antibody	anti-mouse IgG HRP (goat polyclonal)	Abcam	Cat. no. ab97023	WB (1:10000)
Antibody	anti-rabbit IgG HRP (goat polyclonal)	Bangalore Genei	HP-06	WB (1:10000)
Antibody	anti-H3 (rabbit polyclonal)	Abcam	Cat. no. ab1791	ChIP (5 µL per 500 µL IP fraction)
Antibody	anti-H4 (rabbit polyclonal)	Abcam	Cat. no. ab10158	ChIP (5 µL per 500 µL IP fraction)
Antibody	anti-FLAG (M2) (Mouse monoclonal)	Sigma	Cat. no. F1804	WB (1: 2500)
Antibody	anti-protein A (rabbit polyclonal)	Sigma	Cat. no. P9424	WB (1:5000)

Antibody	anti-mouse IgG AlexaFluor488 (goat polyclonal)	Invitrogen	A-11001	IF (1:500)
Affinity beads	GFP-trap beads	ChromoTek	Cat. no. gta-20	ChIP (20 µL per 500 µL fraction)
Affinity beads	Blocked agarose beads	ChromoTek	Cat. no. bab-20	ChIP (20 µL per 500 µL fraction)
Affinity beads	Protein-A sepharose beads	Sigma	Cat. no. P3391	ChIP (20 µL per 500 µL fraction)
Affinity beads	M2 anti-FLAG affinity gel	Sigma	Cat. no. A2220	ChIP (20 µL per 500 µL fraction)

Media and growth conditions

Malassezia strains were grown on modified Dixon's media (malt extract 36 g/L, desiccated Ox-bile 20 g/L, tween40 10 mL/L, peptone 6 g/L, glycerol 2 mL/L, oleic acid 2.89 mL/L). *M. sympodialis, M. furfur* strains were grown at 30°C. Cultures of *M. globosa* and *M. slooffiae* were grown at 32°C.

All *C. albicans* strains in this study, derived from J108 carrying Dad2 under the control of *PCK1* promoter, were maintained in YPSU media (succinate, disodium salt- 10 g/L, yeast extract- 10 g/L, peptone 20g/L, uridine- 0.1 mg/mL). To assay for the role of reintegrated Dad2 from RPS10 locus, YPDU media was used (dextrose- 20 g/L, yeast extract- 10 g/L, peptone 20g/L, uridine- 0.1 mg/mL). All *C. albicans* strains were grown at 30˚C.

The *S. cerevisiae* strain CJY077 carrying a temperature-sensitive allele of Dad2 was maintained in YPAD (dextrose- 20 g/L, yeast extract- 10 g/L, peptone 20g/L, adenine- 0.1 mg/mL). Once transformed with pRS313G based plasmids, the transformants were maintained in complete media lacking leucine and histidine. Routine growth of all these strains was performed at 28˚C. Functional complementation was performed at 37˚C.

Construction of strains related to the study in *Malassezia* species

Construction of the *M. sympodialis* strain expressing GFP-Mtw1

The allele for N-terminal tagging of Mtw1 with GFP was prepared by gap repair in the Saccharomyces cerevisiae BY4741 strain (ECKERT-BOULET *et al.* 2012). Briefly, a 1.6 kb fragment consisting of the upstream and promoter sequence of the MTW1 gene and a 1.6 kb fragment having the MTW1 ORF (Protein ID: SHO76526) along with the downstream sequence was amplified from *M. sympodialis* genomic DNA. The *GFP* ORF (without the stop codon) and *NAT* were amplified from plasmids pVY7 and pAIM1 respectively (YADAV *et al.* 2018b). *S. cerevisiae* was transformed with all four fragments and the linearized plasmid pGI3 (digested with *Kpn*I and *Bam*HI). The epitope-tagged allele was assembled in an ordered way by gap repair. Total DNA was isolated from *S. cerevisiae* transformants, and *E. coli* DH5α strain was transformed with dilutions of the genomic DNA to recover the assembled construct. The pGFP-Mtw1 construct was screened by restriction digestion and further confirmed by sequencing. The pGFP-Mtw1 construct was used to transform *M. sympodialis* strain ATCC42132 by *Agrobacterium tumefaciens*-mediated transconjugation (IANIRI *et al.* 2016; IANIRI *et al.* 2017a).

Construction of the *M. furfur* strain expressing CENP-A FLAG

The allele for C-terminal tagging of CENP-A with a 3xFLAG epitope tag was prepared by gap repair in the *S. cerevisiae* BY4741 strain (ECKERT-BOULET *et al.* 2012). Briefly, a 1 kb fragment consisting of the upstream and promoter sequence of the CENP-A gene of *M. furfur* including the ORF (CENP-A ORF coordinates in Chr1: 1,453,468- 1,453,921) and 1 kb fragment containing the sequence downstream of CENP-A ORF were amplified from *M. furfur* genomic DNA. The 3xFLAG tag was introduced in the reverse primer annealing to the CENP-A ORF. The *NAT* marker was amplified from plasmid pAIM1 as above. *S. cerevisiae* was transformed with all three fragments and plasmid pGI3 (digested with *Kpn*I and *Bam*HI). The epitope-tagged allele was assembled in an ordered way by gap repair. Total DNA was isolated from *S. cerevisiae* transformants, and *E. coli* DH5α strain was transformed with dilutions of the genomic DNA to recover the assembled construct. The resulting pMF1 construct was screened by restriction digestion and further confirmed by sequencing. The pMF1 construct was used to transform *M. furfur* strain CBS14141 by *Agrobacterium tumefaciens*-mediated transconjugation (IANIRI *et al.* 2016; IANIRI *et al.* 2017a) to obtain the epitope-tagged strain MF001.

Construction of strains related to the study in *C. albicans*

Construction of vector to integrate Dad2 at the *RPS1* locus

To generate the integration construct, the sequence corresponding to RPS1 was amplified using primer pair AD11/AD12 from Cip10 plasmid and cloned into pBS-NAT vector in *Not*I site to generate the plasmid pBS-RP10-NAT. The sequences corresponding to Dad2pr-Dad2-Protein-A were amplified as a *Sal*I-*Cla*I fragment from the previously reported strain J118 using primer pair AD05/AD06. This fragment was cloned into the same sites in the vector pBS-RP10-NAT to generate pDad2^{FL}-TAP.

Construction of Dad2^{MUT} reintegration vectors

Deletion of DSS and mutations in the DSS were introduced by overlap PCR using pDad2FL-TAP as the template. Primers AD05 and AD06 were used with internal primers carrying either deletion or mutation in the DSS as follows. The Dad2$^{\Delta DSS}$ allele was generated by overlap PCR with the fragments amplified using primer pairs AD05/AD13 and AD14/AD06. The Dad2^{P87A} allele was generated by overlap PCR with the fragments amplified using primers AD05/P87AF and P87AR/AD06. The Dad2^{R92A} allele was generated by overlap PCR with the fragments amplified using primers AD05/R92AF and R92AR/AD06. The Dad2^{R94A} allele was generated by overlap PCR with the fragments amplified using primers AD05/R94AF and R94AR/AD06. Each of these fragments was subsequently digested and cloned as *Sal*I fragment into pBS-RP10-NAT to generate pDad2$^{\Delta DSS}$ –TAP, pDad2^{P87A} –TAP, pDad2^{R92A} –TAP, and pDad2^{R94A} –TAP respectively. Each of the introduced mutations was confirmed by Sanger sequencing using primer NV34.

Construction of *C. albicans* strains expressing Dad2$^{FL/MUT}$

The Dad2 conditional mutant strain J108 has one copy of Dad2 replaced by *HIS1* marker and the expression of the other allele driven by a *PCK1* promoter (*URA3* marker). Each of the above plasmids viz. pDad2^{FL}-TAP, pDad2$^{\Delta DSS}$ –TAP, pDad2^{P87A} –TAP, pDad2^{R92A} –TAP, and pDad2^{R94A} –TAP were linearized with *Stu*I and were used to transform J108 giving rise to strains CaS100 (expressing protein-A tagged Dad2^{FL} from RPS10 locus), CaS101 (expressing protein-A Dad2$^{\Delta DSS}$ from RPS10 locus), CaS102 (expressing protein-A Dad2^{P87A} from RPS10 locus), CaS103 (expressing protein-A Dad2^{R92A} from RPS10 locus), and CaS104 (expressing protein-A Dad2^{R94A} from RPS10 locus) respectively. The parent plasmid pBS-RP10-NAT was also digested with *Stu*I and used to transform J108 to obtain the control strain J108C. The transformants were selected on YPSU plates supplemented with 200 µg/mL NAT. The expression of Dad2$^{FL/MUT}$ in these strains was confirmed by immunoblotting using anti-protein-A antibodies.

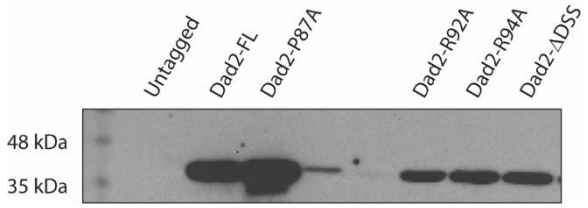


Figure 4-1. Expression of full-length and mutant Dad2 from RPS1 locus.
Whole-cell lysates from strains J108C, CaS100-104 were separated by SDS-PAGE using 12% gels. The proteins were transferred onto nitrocellulose membranes and probed with anti- Protein A antibodies for immunoblot analysis.

Construction of Dad2 mutant strains expressing Ndc80-mCherry

A previously published construct pNdc80-RFP-ARG4 to tag Ndc80 with mCherry at C-terminus was used (VARSHNEY AND SANYAL 2019). This plasmid was linearized with *Xho*I and was used to transform strains J108C, CaSR100, CaSR101, CaSR102, CaSR103, and CaSR104 to generate the strains J108N, CaSR200, CaSR201, CaSR202, CaSR203, and CaSR204 respectively. The transformants were selected on complete media with succinate as carbon source lacking arginine. Transformants were screened for the punctate localization of Ndc80-mCherry by fluorescence microscopy.

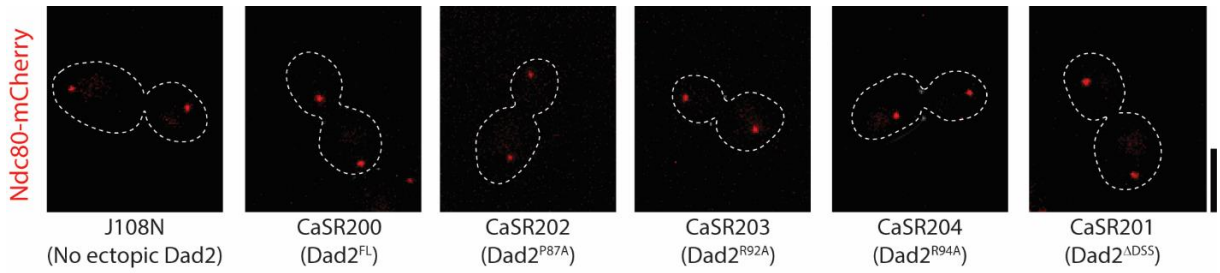


Figure 4-2. Localization of Ndc80-mCherry in Dad2 mutants.
Transformants obtained with Ndc80-mCherry-Arg4 cassette were purified for single colonies on selection plates. Single colonies of strains J108N and CaSR200-204 were grown to log phase in YPSU, harvested, washed, and observed under a fluorescence microscope for clustered kinetochore-like signals. Scale bar, 5µm.

Construction of Dad2 mutant strains expressing Tub1-mCherry

The plasmid pTub1-mCherry-HIS1 was digested with *Spe*I and *Eco*RV, and the 1.6 kb fragment containing sequences corresponding to the 3' end of the *TUB1* in frame with mCherry was isolated. The fragment was purified and ligated into the same sites in pFA-TAP-ARG322. The resulting plasmid

pTUB1-mCherry-ARG4 was linearized by *Xba*I and used to transform strains J108C, CaSR100, CaSR101, CaSR102, CaSR103, and CaSR104 to generate the strains J108T, CaSR300, CaSR301, CaSR302, CaSR303, and CaSR304 respectively. The transformants were selected on complete media with succinate as carbon source lacking arginine. Transformants were screened by fluorescence microscopy for proper spindle localization.

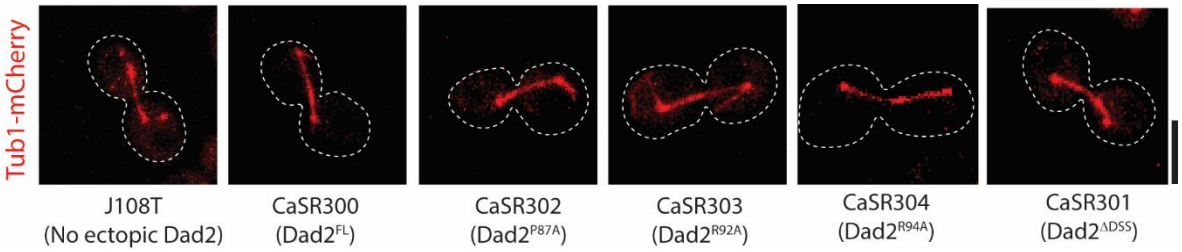


Figure 4-3. Localization of Tub1-mCherry in Dad2 mutants.
Transformants obtained with Tub1-mCherry-Arg4 cassette were purified for single colonies on selection plates. Single colonies of strains J108T, CaSR300-304 were grown to log phase in YPSU, harvested, washed, and observed under a fluorescence microscope for anaphase-like long spindle in large budded cells. Scale bar, 5µm.

Construction of vectors for heterologous expression of Dad2$^{FL/MUT}$

The Dad2 coding sequence (*C. albicans* ORF19.3551) was amplified by primer pair AD01/AD02 and cloned into *Nco*I and *Hind*III sites of pET28b (Novagen). The resulting plasmid pET28b-Dad2^{FL} expresses Dad2 as a C-terminally 6xHis tagged fusion protein from the T7 promoter. The same primer pair was used to amplify mutant versions of Dad2 viz. Dad2$^{\Delta DSS}$, Dad2^{P87A}, Dad2^{R92A}, and Dad2^{R94A} respectively from templates pDad2$^{\Delta DSS}$ –TAP, pDad2^{P87A} –TAP, pDad2^{R92A} –TAP, and pDad2^{R94A} –TAP. Each of these amplicons was cloned into *Nco*I and *Hind*III sites of pET28b to generate plasmids pET28b-Dad2$^{\Delta DSS}$, pET28b-Dad2^{P87A}, pET28b-Dad2^{R92A}, and pET28b-Dad2^{R94A} respectively. The mutations were confirmed by Sanger sequencing using the T7 promoter primer.

Construction of strains related to the study in *S. cerevisiae*

Construction of vectors to express Dad2$^{FL/MUT}$-GFP in pRS313

To express GFP tagged ScDad2$^{FL/MUT}$ from pRS313 (CEN/ARS/HIS1 plasmid), sequences coding for GFP along with terminator sequences of *ScACT1* was amplified using primer pair ScGFP-F/ScGFP-R and cloned as a *BamH*I-*Cla*I fragment into pRS313 resulting in pRS313G. This plasmid was subsequently used to clone Dad2 with/without mutations. For cloning ScDad2^{FL}, a fragment containing *ScDAD2pr-DAD2 ORF* was amplified from the genomic DNA of BY4741 strain with primers SR289/SR290 and cloned in frame with GFP as a *Sac*II-*BamH*I fragment into pRS313G. The resulting plasmid was named

pScDad2^{FL}. Since the DSS in *Sc*Dad2 is located near the stop codon, any desired mutation was incorporated in the reverse primer used to amplify Dad2. In this way, ScDad2^{R126A}, ScDad2^{R128A}, ScDad2$^{\Delta DSS}$ were amplified with primer pairs SR289/Sc126A-R, SR289/Sc128A-R, and SR289/ScΔDSS-R respectively, and cloned into the *Sac*II-*BamH*I sites of pRS313G. The resulting plasmids pScDad2^{R126A}, pScDad2^{R128A}, and pScDad2 $^{\Delta DSS}$ along with pScDad2^{FL} were confirmed by Sanger sequencing using SR289.

Construction of strains ScSR01-04

A previously reported conditional mutant of ScDad2 (CJY077: *MAT***a** *Δdad2::KanMX6 ura3-52 lys2-801 ade2-101 trp1Δ63 leu2Δ1::pCJ055(DAD2^{TS}, LEU2) his3Δ200*) was used as the parent strain. Each of the plasmids pScDad2^{FL} pScDad2^{R126A}, pScDad2^{R128A}, and pScDad2 $^{\Delta DSS}$ was used to transform CJY077 strain by the standard lithium acetate method to generate strains ScSR01, ScSR02, ScSR03, and ScSR04 respectively. The parent plasmid pRS313G was also used to transform CJY077 to generate the vector control strain (ScEV). Transformants were selected on CM-Leu-His media upon incubation at 28˚C. Selected colonies from the transformation plate were streaked again on CM-Leu-His media for single colonies and used for subsequent experiments.

C. albicans transformation

Transformation of *C. albicans* was performed by the lithium acetate method as described previously (Sanyal et al., 2004). A 50 ml culture of *C. albicans* cells was grown in YPSU medium till OD_{600}=0.8-1.0 and harvested by centrifugation for 5 min at 4000 rpm. The cells were washed once each with 20 mL of autoclaved water and LiOAc/TE solution. The cell suspension was centrifuged at 4000 rpm for 5 min. The cells were finally resuspended in three times the pellet volume in LiOAc/TE solution. Aliquots of 100 µL were made as required, including a No DNA control. Meanwhile, an aliquot of sheared Salmon sperm DNA (10 mg/mL) was boiled at 95˚C for 5 min and snap-chilled immediately. To all the tubes with cell suspension, 10 µL of sheared Salmon sperm DNA was added and thoroughly mixed, followed by the addition of 10-30 µL of the DNA to be transformed only to the test transformation tube. Subsequently, 700 µL of PEG solution was added to each tube, vortexed briefly and incubated at 30˚C for 12-14 h. Post incubation, the samples were subjected to heat shock at 42˚C for 45 min. Cells were washed, recovered in the desired media, and plated on an appropriate selection plate. Transformants obtained on the selection plates were streaked again on the secondary plates to obtain single colonies.

S. cerevisiae transformation

Freshly grown CJY077 cells from a YPAD plate were scraped using a sterile toothpick. Cells equivalent to a volume of 100 µL were used to two transformation experiments. The scraped cells were

resuspended in 1mL of sterile distilled water, vortexed and pelleted by centrifugation at 13,000 rpm for 5 s. The pellet was resuspended in 1 mL of 100 mM LiOAc solution and incubated at 30˚C for 5 min. The cell suspension was split equally into two tubes and cells were pelleted by centrifugation at 13,000 rpm for 5 s. To the cell pellet, 240 µL of PEG (50% w/v), 36 µL LiOAc (1M), 10 µL salmon sperm DNA (10 mg/mL), and 64 µL sterile water were sequentially added and vortexed to make a uniform suspension. For 350 µL of the above suspension, 5 µL plasmid DNA containing ~100 ng DNA was added and vortexed thoroughly. The transformation mix was subjected to heat-shock for 10 min at 42˚C. The cells were harvested by centrifugation at 13,000 rpm for 10 s, resuspended in 100 µL sterile water, and 30 µL of the suspension was plated on to CM-Leu-His plates. The plates were incubated at 28˚C.

E. coli competent cells preparation and transformation

Competent cells of DH5α and BL21(DE3) were prepared by the PEG method (Chung et al 1989). Cells from an overnight grown culture were seeded at 1% into flasks containing LB media and grown at 37˚C. Cells were chilled on ice when the OD_{600} reached 0.3-0.5. Cells were harvested by centrifugation at 2500 rpm for 5 min at 4˚C. Cell pellet from 50 mL culture was gently resuspended in 2 mL of ice-cold TSS buffer (2xLB, 10% (w/v) PEG3350, 100 mM $MgCl_2$, 5% (v/v) DMSO). Aliquots of 100 µL were made and snap-frozen using liquid N_2. For transformation, the desired amount of plasmid or ligation product (in not more than 10 µL volume) was added per tube after thawing on ice. Post addition, the tubed were tap mixed, incubated on ice for 30 min. Cells were recovered in LB media at 37˚C for 40 min. Cells were then harvested, resuspended, and plated on LB plates containing the appropriate antibiotic.

Genomic DNA preparation

For *Malassezia*, 10 mL of cells grown overnight were harvested and washed twice with autoclaved water. Glass beads equivalent to the pellet volume was added to the tube, followed by the addition of 750 µL of extraction buffer (see Buffers and Solutions used for composition). The cells were lysed by vortexing for 5 min, after which the suspension was incubated at 65˚C for 30 min. After cooling the tubes to room temperature, an equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) was added, vortexed briefly and centrifuged at 13,000 for 10 min. The aqueous phase was recovered, and the DNA was precipitated with 100% ethanol at -20˚C for 1 h. The tubes were centrifuged at 14,000 rpm at 4˚C, and the pellet was washed with ice-cold 70% ethanol. The pellet was air-dried and resuspended in 30 µL of 1xTE.

For *C. albicans* cells, 3-5 ml culture of grown overnight was harvested, washed with autoclaved water, and resuspended in 0.2 mL extraction buffer. Cells were lysed by vortexing with 0.3 g of acid-washed glass beads and 0.2 mL of phenol: chloroform: isoamyl alcohol (25:24:1). After centrifugation at

14,000 rpm for 10 min, the supernatant was collected, and DNA was precipitated with 100% ethanol at -20˚C for 1 h. The tubes were centrifuged at 14,000 rpm at 4˚C, and the pellet was washed with ice-cold 70% ethanol. The pellet was air-dried and resuspended in 30 µL of 1xTE.

Preparation of whole-cell lysates and Immunoblotting

Protein lysates for western blot were prepared by the TCA method. One mL overnight grown cultures (for *Malassezia*) or 3 OD_{600} equivalent cells (for *C. albicans* and *S. cerevisiae*) were harvested, washed, and resuspended in 400 µL of 12.5% ice-cold TCA solution. The suspension was vortexed briefly and stored at – 20°C for 4 to 6 h. The suspension was thawed on ice, pelleted at 14,000 rpm for 10 min, and washed twice with 350 µL of 80% Acetone (ice-cold). The washed pellets were air-dried completely and resuspended in the desired volume of lysis buffer (0.1 N NaOH+1% SDS). Samples were separated in 12% polyacrylamide gels and transferred onto nitrocellulose membranes and probed with appropriate antibodies (see the list of antibodies and affinity beads for dilutions). The blots were developed using Chemiluminescence Ultra substrate (BioRad) and imaged using the VersaDoc imaging system (BioRad).

Preparation of spheroplasts

Malassezia cells were grown on mDixon's medium and washed with water by centrifugation at 4000 rpm for 5 min. Cells were resuspended in 10 mL of 5% (v/v) 2-mercaptoethanol solution in water and incubated at 30°C/ 150 rpm for 45 min. The cells were pelleted, washed, and resuspended in 3 mL spheroplasting buffer (40 mM Citric acid, 120 mM Na_2HPO_4, and 1.2 M Sorbitol) for every $1.5x10^9$ cells. Cell clumps were dissociated by mild sonication for 30 s using the medium intensity setting in a Bioruptor (Diagenode). Lysing enzymes from *Trichoderma harzianum* (Sigma), Chitosanase (Sigma), and Zymolyase-20T (MP Biomedicals) was added at 20 mg/mL, 0.2 µg/mL, and 100 µg/mL, respectively. The spheroplasting suspension was incubated at 30°C/ 65 rpm for 6 to 8 h. The suspension was examined under a microscope to estimate the proportion of spheroplasts. Spheroplasts were washed with ice-cold 1xPBS and used as per the experimental design (adapted from (BOEKHOUT 2003)).

Chromatin immunoprecipitation

The ChIP protocol was adapted from the one implemented for *C. neoformans* (YADAV *et al.* 2018b). Logarithmically grown cells were fixed with formaldehyde at a final concentration of 1% for 30 min (for Mtw1 ChIP) and 15 min (for CENP-A, histone H3, and histone H4 ChIP) respectively. The reaction was quenched by the addition of glycine to 0.135 M final concentration. The cells were pelleted and processed for spheroplasting as described above. Spheroplasts were washed once sequentially using 10 mL of the following ice-cold buffers: 1xPBS, Buffer-I, and Buffer-II. The pellet after the final wash

was resuspended in 1 mL lysis buffer for every $1.5x10^9$ cells. A protease inhibitor cocktail (Roche) was added to 1x final concentration.

The resuspended spheroplasts were sonicated with a Bioruptor (Diagenode) using 30 s ON/OFF pulse at high-intensity mode with intermittent incubation on ice to obtain chromatin fragments in size range of 100-300 bp. The lysate was cleared after sonication by centrifugation at 13,000 rpm for 10 min at 4ºC. The input DNA fraction was separated at this step ($1/10^{th}$ volume of lysate) and processed for de-crosslinking by the addition of 400 µL elution buffer per 100 µL lysate (processing for de-crosslinking mentioned below). The remaining lysate was split equally and processed as IP and control samples. For GFP-Mtw1 ChIP, 20 µL GFP-trap beads and blocked agarose beads respectively were used for IP and control. For CENP-A-3xFLAG ChIP, 20 µL anti-FLAG affinity gel and blocked agarose beads respectively were used for IP and control. In the case of histone H3 or histone H4 ChIP, 5 µL of antibodies were used per IP fraction along with 20 µL Protein-A sepharose beads. Samples were rotated for 6 h at 4ºC. Post incubation, samples were sequentially washed as follows: twice with 1 mL low salt wash buffer, twice with 1 mL high salt wash buffer, once with 1 mL LiCl wash buffer and twice with 1 mL 1xTE. Samples were rotated in a rotaspin for 5 min at room temperature for every wash (15 min in case of histone H3 or histone H4 ChIP). After washes, DNA was eluted from the beads twice using 250 µL elution buffer. The samples for elution were incubated at 65ºC for 5 min, rotated for 15 min, and then collected by centrifugation.

Crosslinks were reversed by the addition of 20 µL 5 M NaCl to each tube and incubation at 65ºC for 6 h. The samples were then deproteinized by the addition of 10 µL 0.5 M EDTA, 20 µL 1 M Tris pH=6.8, 2 µL Proteinase K (20 mg/L), and incubation at 45ºC for 2 h. After incubation, samples were treated with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1) mix, and the aqueous phase was extracted by centrifugation. DNA was precipitated by addition of $1/10^{th}$ volume of 3 M Na-acetate, 1 µL glycogen (Sigma, 20 mg/mL), 1 mL absolute ethanol and incubation at -20ºC for at least 12 h. Finally, the samples were harvested by centrifugation at 13,000 rpm for 45 min at 4ºC followed by washing the pellet once with ice-cold 70% ethanol. Air-dried pellets were then resuspended in 20 µL sterile MilliQ water with 10 µg/mL RNAse. Samples were either processed for library preparation for ChIP-sequencing or analyzed by qPCR with the primers mentioned in the list of primers used in this study.

Data analysis

GFP-Mtw1 ChIP sequencing was performed at the Clevergene Biocorp. Pvt. Ltd., Bengaluru, India. A total of 46,704,720 and 63,524,912 150 bp paired-end reads were obtained for IP and Input samples, respectively. The reads were mapped to *M. sympodialis* ATCC42132 genome using Geneious 9.0 (http://www.geneious.com/) with default conditions. Each read was allowed to map

only once randomly anywhere in the genome. The alignments were exported to BAM files, sorted, and visualized using Integrative Genomics Viewer (IGV, Broad Institute). The images from IGV were imported into Adobe Photoshop (Adobe) and scaled for representation purposes. RNA-sequencing data (E-MTAB-4589) from a previous study was downloaded from the ArrayExpress website, sorted, and visualized using IGV. GC-content was calculated using Geneious 9.0 with a sliding window size of 250 bp. The data was exported as wig files and further visualized using IGV.

Synteny Analysis

For gene synteny conservation across the centromeres, the analysis was performed by BLAST as follows. The genomes for *M. restricta*, *M. nana*, and *M. dermatis* were downloaded from the NCBI genomes portal. The PacBio assembled genomes of *M. globosa* and *M. slooffiae* were used for synteny analysis. Synteny analysis was done in the context of ORFs flanking the centromeres of *M. sympodialis*. The protein sequences of these ORFs served as the query in BLAST analysis against the genome of other species. The local database for each genome was set up in the Geneious software for this analysis. Additionally, synteny analyses between *M. globosa* and *M. sympodialis* were conducted with megablast (word size: 28) and plotted together with GC content (calculated as the deviation from the genomic mean, in non-overlapping 1 kb windows), using Circos (v0.69-6) (KRZYWINSKI *et al.* 2009). Additional whole-genome alignments were conducted with Satsuma (GRABHERR *et al.* 2010), with default parameters. The linear synteny comparisons shown in Figures 2-9 and 2-10 were generated with the Python application EasyFig (SULLIVAN *et al.* 2011).

Species phylogeny

To reconstruct the phylogenetic relationship among the 9 *Malassezia* species selected, orthologs were identified using the bidirectional best-hit (BDBH), COGtriangles (v2.1), and OrthoMCL (v1.4) algorithms implemented in the GET_HOMOLOGUES software package (CONTRERAS-MOREIRA AND VINUESA 2013). The proteome of *M. sympodialis* ATCC42132 was used as a reference. A phylogeny was inferred from a set of 738 protein sequences as follows. Individual proteins were aligned using MAFFT v7.310 (L-INS-i strategy), and poorly aligned regions were trimmed with TrimAl (-gappyout). The resulting alignments were concatenated to obtain a final supermatrix consisting of a total of 441,200 amino acid sites (159,270 parsimony-informative). This sequence was input to IQ-TREE v1.6.5 (NGUYEN *et al.* 2015), and a maximum likelihood phylogeny was estimated using the LG+F+R4 amino acid model of substitution. Branch support values were obtained from 10,000 replicates of both ultrafast bootstrap approximation (UFBoot) and the nonparametric variant of the approximate likelihood ratio test (SH-aLRT) implemented in IQ-TREE. The best likelihood tree was graphically visualized with iTOL v4.3.3 (LETUNIC AND BORK 2019).

Timepoint assay for DSS function

To assay for DSS function, all the *C. albicans* strains were grown overnight in YPSU media, after which they were washed with 5 mL autoclaved water. The cells were then reinoculated to YPDU media to 0.2 OD_{600}. Cultures in YPDU were harvested at desired timepoints, after which they were harvested, washed, and processed as per the experiment desired. For DAPI staining, cells were fixed by adding formaldehyde to a final concentration of 1% and incubated for 1 h. Post fixing, cells were washed and resuspended in 100 µL water to which 300 µL ethanol was added. Samples were either incubated at room temperature for 1 h or stored at 4°C overnight. Cells were then harvested, rehydrated with water, and stained with DAPI before imaging.

Fluorescence microscopy of live cells- sample preparation and analysis

The GFP-Mtw1 strain was inoculated to 1% v/v from a saturated starter culture grown in mDixon medium. After growth for 6 h at 30°C/ 150 rpm, these cells were pelleted at 4,000 rpm and washed 3 times with 1x phosphate-buffered saline (PBS), and the cell suspension was placed on a clean glass slide. A coverslip was placed on the spot and sealed before imaging. *C. albicans* cells harvested at any desired time point were also processed the same way. The images were acquired at room temperature using a laser scanning inverted confocal microscope LSM 880-Airyscan (ZEISS, Plan Apochromat 63x, NA oil 1.4) equipped with highly sensitive photodetectors. The filters used were GFP/FITC 488 excitation and GFP/FITC 500/550 band pass, long pass for emission. Z-stack images were taken at every 0.3 µm and processed using ZEISS Zen software or ImageJ. All the images were processed post-acquisition with minimal adjustments to levels and linear contrast until the signals were highlighted.

Indirect Immunofluorescence

The GFP-Mtw1 strain was inoculated to 1% (v/v) from a saturated starter culture grown in mDixon medium. After 6 h growth, the cells were fixed by the addition of formaldehyde to a final concentration of 3.7% for 1 h. The cells were then washed with water and taken for preparation of spheroplasts (described above). Spheroplasts were washed with ice-cold 1xPBS and resuspended in ice-cold 1xPBS to a cell density suitable for microscopy. Slides for microscopy were washed with water and coated with poly L-Lysine (15 µL of 10 mg/mL solution per well) for 5 min at room temperature. The solution was aspirated and washed once with water. The cell suspension was added to each well (15-20 µL) and allowed to stand at room temperature for 5 min. The cell suspension was aspirated, and the slides were washed once with water to remove unbound cells. The slides were fixed in ice-cold methanol for 6 min followed by treatment with ice-cold acetone for 30 s. Post fixing, blocking solution (2% non-fat skim milk in 1xPBS) was added to each well, incubated at room temperature for 30 min. After this, the

blocking solution was aspirated, and primary antibodies were added (mouse anti-GFP antibodies [Sigma] at 1:100 dilution). After incubation for 1 h at room temperature, each slide was washed 8 times with 1xPBS giving a 2 min incubation for every wash. Secondary antibody solution (goat anti-mouse AlexaFluor488 [Invitrogen] at 1:500 dilution) was added to each well and incubated for 1 h in the dark at room temperature. Post-incubation, slides were washed as described above. Mounting medium (DAPI at 100 ng/mL in 70% glycerol) was added, incubated for 5 min, and aspirated out. Slides were sealed with a clean coverslip and proceeded for imaging. The images were acquired at room temperature using an inverted fluorescence microscope (ZEISS Axio Observer, Plan Apochromat 100x, NA oil 1.4). Z- stack images were taken at every 0.3 µm and processed using ZEISS Zen software/ ImageJ.

Expression and purification of Dad2 from *E. coli*

The pET28b-based constructs containing various versions of Dad2 were used to transform BL21 (DE3) (Stratagene). For protein expression, these transformants were grown in Luria Bertani (LB) broth (Himedia) at 37˚C/180 rpm until the mid-log phase ($O.D._{600}$= 0.5-0.6). The protein expression was induced by the addition of isopropyl β-D-1-628 thiogalactopyranoside (IPTG) to 1 mM concentration, and the cultures were then grown at 16°C/180 rpm for 15 h. Cells were harvested by centrifugation, and the pellets were frozen at −80°C. The cell pellet was resuspended in the lysis buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 5 mM Imidazole) with a protease inhibitor tablet (Roche), lysed with 1 mg/mL lysozyme, and sonicated in an ice bath (VC-600, Sonics & Materials Inc., USA) to rupture the cells. The lysate was centrifuged at 10,000 g for 40 min at 4 °C. The clarified cell lysate was applied to Ni-NTA resin (Qiagen, USA) packed in a Poly-Prep® Chromatography Column (BIORAD, USA). The resin was washed with the wash buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 50 mM Imidazole). Elution of the target protein was carried out in elution buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, and 250 mM Imidazole). The protein extract was concentrated and subjected to gel filtration pre-packed column Sephadex 75 -120 mL (GE Healthcare). Appropriate fractions were pooled and concentrated, resulting in a 10–15 mg/mL protein solution, estimated to be >96 % pure and homogeneous by SDS-PAGE. Amicon Ultra concentrators with 3 or 10 kDa molecular weight cut-off (Millipore) were used for concentration. Protein samples were snap-frozen in liquid nitrogen for long term storage at -80 °C.

Microtubule cosedimentation assays

Tubulin (Cytoskeleton Inc.) was polymerized in BRB80 (80 mM PIPES, 1 mM EGTA, and 1 mM $MgCl_2$) buffer containing 1 mM GTP and 20 µM Taxol at 37 °C for 15 min. Varying concentrations of the polymerized microtubules (10- 50 µM) were incubated with 50 µM purified $Dad2^{FL}$ or $Dad2^{mut}$ for 10 min at room temperature and then sedimented after passing through a 15 % sucrose cushion followed by centrifugation at 30,000xg for 45 min at 37°C. The separated pellet fractions were dissolved and

loaded onto SDS-PAGE (12%). The intensities of Coomassie Brilliant Blue stained-bands of the proteins in SDS-PAGE were measured by Quantity One software to estimate the binding coefficient. The background intensities of proteins were estimated by pelleting them through sucrose cushion in the absence of microtubules. The corrected band intensities of proteins relative to the background were obtained by subtracting background intensities from the sample data.

List of primers utilized in this study

Primers used in *Malassezia* work		
Primers to generate epitope tagging alleles in *M. sympodialis*		
Msy Mtw1 N-P1	GCGCGCCTAGGCCTCTGCAGGTCGACTCTGACTGACCACGACGAGCTG	Primers to tag Mtw1 with GFP at N-terminus
Msy Mtw1 N-P2	CGCCCTTGCTCACCATCGAGGGGTGGAGGTACAATAG	
Msy Mtw1 N-P3	CTATTGTACCTCCACCCCTCGATGGTGAGCAAGGGCG	
Msy Mtw1 N-P4	GCGTCCGAGGTGGACATGTACAGCTCGTCCATGCC	
Msy Mtw1 N-P5	GGCATGGACGAGCTGTAcATGTCCACCTCGGACGC	
Msy Mtw1 N-P6	GAGGATCTGCACCGTGGCACATTGCGCGATGATG	
Msy Mtw1 N-P7	CATCATCGCGCAATGTGCCACGGTGCAGATCCTC	
Msy Mtw1 N-P8	TGATTACGAATTCTTAATTAAGATATCGAGCGTCCTCTCCTATGTCTGACC	
Primers to generate epitope tagging alleles in *M. furfur*		
MfCse4 P1	GCGCGCCTAGGCCTCTGCAGGTCGACTCTATGCAGCAACAGGCACACATG	Primers to tag CENP-A with 3xFLAG tag at C-terminus
MfCse4 P2	CTACTTGTCATCGTCATCCTTGTAGTCGATGTCATGATCTTTATAATCACCGTCATGGTCTTTGTAGTCCCGGATGTCGCCCCAATG	

Mf NAT-F	GACTACAAAGACCATGACGGTGATTATAAAGAT CATGACATCGACTACAAGGATGACGATGACAAG TAGTCCACGGTGCAGATCCTCG	
Mf NAT-R	GCTTTCATAGGAACATGCCCTGCGTCCTCTCCTA TGTCTG	
MfCse4 P3	CAGACATAGGAGAGGACGCAGGGCATGTTCCTA TGAAAGC	
MfCse4 P4	TGATTACGAATTCTTAATTAAGATATCGAGGAGG CGATCAACCGGCTTAG	
Primers for *M. sympodialis* centromeres		
MS1 F1	AAGAATTGATAACATTGTTGCAC	*MsyCEN1* primers
MS1 R1	TAGAATAAAATGTCGCGAAGG	
MS2 F1	CTGAAGAAAAGAAACAAATTCG	*MsyCEN2* primers
MS2 R1	TCGGAAATCCCGCAAAAG	
MS3 F1	CATATTCAGCCTCCACTAAG	*MsyCEN3* primers
MS3 R1	CCTCTATCGAGTGCTCTAC	
MS4 F1	CGATATGGATTGGACTTATAAGTC	*MsyCEN4* primers
MS4 R1	AAAAGCAATACGTAGACGG	
MS5 F1	AAATTACCGACCAGAATTG	*MsyCEN5* primers
MS5 R1	ATCTGTGTCCGCTCTCATC	
MS6 F1	TTTGACGCTTTATTTGTGTTTC	*MsyCEN6* primers
MS6 R1	CACATATGCACGAATAATAAAACG	
MS7 F1	GATACATATTCTTACACTAATACTATTCG	*MsyCEN7* primers
MS7 R1	GCATAGAGCTAATATCTGATATTC	
MS8 F1	GGAAGCATGAGATATTGG	*MsyCEN8* primers
MS8 R1	AAACAAAGTAAAATTCTAATCACG	
MS8 LF1	CTCCTCCGATACGATTCAC	*MsyCEN8* L1 primers
MS8 LR1	CAGCCATTATCTCCGACAC	

MS8 LF2	CTGGGTAGATTGAGAATGAG	*MsyCEN8* L2 primers
MS8 LR2	CATGTATGTTCAGTCCCATG	
MS8 RF1	ATGATCCAAAAGAAAGCATAC	*MsyCEN8* R1 primers
MS8 RR1	GAAGTATGTCTGGGTGAAGC	
MS C3	GAAGACGACAACGATACC	Control primers away from *MsyCEN1*
MS C4	TAGCGAGTGAATAGCGTC	
Primers for chromoblot analysis in *M. globosa*		
Maglo_CBS7966_v2_Chr3_216001_216700_Forward:	GATGAGCGACGGAAACAAGC	Probe for Chr3
Maglo_CBS7966_v2_Chr3_216001_216700_Reverse:	AACTTCGTCCCATTCGCCTT	
Maglo_CBS7966_v2_Chr4_150001_150700_Forward:	CATCGAGATTGCAACACAGC	Probe for Chr4
Maglo_CBS7966_v2_Chr4_150001_150700_Reverse:	TGAACACAGGCGCCATTGTA	
Maglo_CBS7966_v2_Chr5_213001_213600_Forward:	TGCAATGAAGTCCGGCATGA	Probe for Chr5
Maglo_CBS7966_v2_Chr5_213001_213600_Reverse:	AGGCACACGTTCATCTGGTT	
Maglo_CBS7966_v2_Chr6_461001_461600_Forward:	TGCTCACCCAAAAGACGACC	Probe for Chr6
Maglo_CBS7966_v2_Chr6_461001_461600_Reverse:	CGCGGACCTGGAACTGTATT	
Primers for *M. globosa* centromeres		
Mg 1F	GAATTGCAATAGTAAGCCGAAC	*MgCEN1*
Mg 1R	GAATTATTCAACCCTTTGTACATC	

Mg 2F	GCAAAAGTTCTGGTTAAAC	*MgCEN2*
Mg 2R	TTCGTTAAATTACTGTCATTAG	
Mg 3F	GCATGTACAATTCTCTAAAAC	*MgCEN3*
Mg 3R	CAAGTTATCTTAATCCGCAAG	
Mg 4F	CAGAAAATAATAGTGATTGATAC	*MgCEN4*
Mg 4R	ATTTAAGATACATACACAATGC	
Mg 6F	GGAAATCCTGCGAGAATC	*MgCEN6*
Mg 6R	GCTGAATTCATAGAATCATTGAG	
Mg 7F	GATGATCCCAGTAACAACTG	*MgCEN7*
Mg 7R	GGTAGAATTGAATTTGTGTTTATC	
Mg 8F	GACTAGCGAATAAATCAATTGAC	*MgCEN8*
Mg 8R	TTAACCGTACCGAAAAACC	
Mg 9F	GAAAATAGTGACTGGTGGAC	*MgCEN9*
Mg 9R	GATTCTATTGCTATATTGTGCTTC	
Mg 5F	CTAAAAATGAAATTTGGGATAAAAC	*MgCEN5*
Mg 5R	AAGCACGATAAAAATCATAGC	
Mg2 L2F	CGTACCTTGTCCAAGAGC	*MgCEN2* L2 primer pair
Mg2 L2R	AGATCCATAGGCTTTGAATGC	
Mg2 L1F	ACCTTCGATTCTGTGACAC	*MgCEN2* L1 primer pair
Mg2 L1R	TGTTACACACTTTGCTTCGG	
Mg2 R1F	AGGTCCTGACGATGTAATTG	*MgCEN2* R1 primer pair
Mg2 R1R	GTTGTTGATGTATGTCGTTCATG	
Mg2 R2F	AGCTATGCGATGTTGTTCTG	*MgCEN2* R2 primer pair
Mg2 R2R	CAGACGAGGAACTATTGTGAG	
Mg C5	GCATAACATACGAGGATGTGC	Primer for control locus
Mg C6	ATAGTGCCTGAATCTGCTG	

Primers for *M. slooffiae* centromeres		
Slo1 FP	CAAATGAGCACAAACGTTG	*MslCEN1*
Slo1 RP	GGTAATTTACATTTCTTGTG	
Slo2 FP	ACTCAATAATCCAATAGAACC	*MslCEN2*
Slo2 RP	GAGAAAACATAAATGGTAGG	
Slo3 FP	AAACCGATTATCAATTCTCAAATG	*MslCEN3*
Slo3 RP	GTATCTGATTTGAAAACCTTCG	
Slo4 FP	TTCACGTGTAGCTACTTG	*MslCEN4*
Slo4 RP	AAATACAACAAACAACTAAAACG	
Slo5 FP	GAGCTGTGCAAGGTTAG	*MslCEN5*
Slo5 RP	GCCAAACAACGATGACG	
Slo6 FP	AATAGATTTGACAACCTTTGC	*MslCEN6*
Slo6 RP	TGCACAATTGTAAGAAAGC	
Slo7 FP	AATGCCAGATGATAAACTAGCTG	*MslCEN7*
Slo7 RP	GACTTCTGGCATAACTATTGG	
Slo8 FP	TTATGCTATTGTTTGAATCCG	*MslCEN8*
Slo8 RP	CTATCATTAAACGGAGAATACTC	
Slo9 FP	GACCTAGCTGTGCTTTTAG	*MslCEN9*
Slo9 RP	TTTCAGCAGCTTATTAGGC	
Slo C1	ACGACAAGCGTGTAAGG	Primers for control locus
Slo C2	CCAACTTCTTCCTGCAG	
Slo1 L2F	GGAACGTGACGAGATCAC	*MslCEN1* L2 primers
Slo1 L2R	GTGTAGATCCGAAGTCATCAC	
Slo1 L1F	ATCTCTGCAAGCTTCGG	*MslCEN1* L1 primers
Slo1 L1R	AGTGGATGCTTCATCTTCTG	
Slo1 R1F	CGAATGACTTCCTCAATGC	

Slo1 R1R	TGCAACAGCAGAAGAGTC	*MslCEN1* R1 primers
Slo1 R2F	ATGCCGACCACAATCC	*MslCEN1* R2 primers
Slo1 R2R	ACTGTGCCTGTTTCGC	
Primers for *M. furfur* centromeres		
MF1 F1	GATAGCAAACATGATTAAAGTAATAAC	*MfCEN1*
MF1 R1	GACCAAAATAATATATATTAACAAATG	
MF2 F1	CAAAAGTGAAGAAGCAGG	*MfCEN2*
MF2 R1	CACATATAAGAAGTAGAAAAGAAAACTC	
MF3 F1	CATGTCTGGACCTCGG	*MfCEN3*
MF3 R1	CGTGGTGAGAACACAAC	
MF4 F1	CCTAAACTTATGAACTGTTTATTC	*MfCEN4*
MF4 R1	GTTAAGTATTCCATAATGCTC	
MF5 F1	CTTCTGCCATCGTTTCTC	*MfCEN5*
MF5 R1	CTTGATTGTTCCTTCGTAATTAAC	
MF6 F2	CATGTATGTTAAACGTCATAGTAC	*MfCEN6*
MF6 R2	CGATTTGATCTATAATAACATAC	
MF7 F1	GTGAAGCTATAATATTATAGAATGAG	*MfCEN7*
MF7 R1	CGTTTGAATCATTATAATACTG	
MF7 LF1	GAAAGCTTCATTCGGAGC	*MfCEN7* L1 primer pair
MF7 LR1	CGTCTTGGGAAGAGCAG	
MF7 LF2	GGCGGATCATCTTTTCG	*MfCEN7* L2 primer pair
MF7 LR2	GATTCTGATCGTCGGAGG	
MF7 RF1	GTGGCACTACTGGATCG	*MfCEN7* R1 primer pair
MF7 RR1	CGTGTACCGGTACATGTG	
MF7 RF2	CTGTACCGCTACCTGC	

MF7 RR2	GTACGAATCGAGATCAACTG	*MfCEN7* R2 primer pair
GI154	GTCGGAGAAGCAGTCAATGC	Primers for control locus
NAT sFP	GTGCGGAGAAGGCATTGTTC	
Primers used in *Dad2* work		
AD01	CATGCCATGGCCTCGAAAACAAATACTGC	Primers to amplify Dad2 ORF
AD02	CCCAAGCTTTTCCGTGGATTC	
AD05	ACGCGTCGACATTTTGACTAGT TCTCAAATGGTTC	Primers to amplify Dad2Pr-Dad2ORF
AD06	CCATCGATCGATATCAAGCTTCAGGTTG	
P87AF	AATAACAAAGAGGCTTTCGCTGAACCATTAGTA AGAGTGCGTGTT	Primers to introduce P87A mutation
P87AR	AACACGCACTCTTACTAATGGTTCAGCGAAAGCC TCTTTGTTATT	
R92AF	GAACCATTAGTAGCAGTGCGTGTTGGACAATCA	Primers to introduce R92A mutation
R92AR	TGATTGTCCAACACGCACTGCTACTAATGGTTC	
R94AF	GAACCATTAGTAAGAGTGGCTGTTGGACAATCA AAT	Primers to introduce R94A mutation
R94AR	ATTTGATTGTCCAACAGCCACTCTTACTAATGGT TC	
delDSS FP	GTCTTGATTTTCTTCATTTGATTGTCCGAAAGCCT CTTTGTTATTATC	Primers to delete DSS
delDSS RP.	GATAATAACAAAGAGGCTTTCGGACAATCAAAT GAAGAAAATCAAGAC	
SR289	TCCCGCGGCATTGCGGCAGGTAAAATATC	Primers to amplify scDad2 ORF with promoter
SR290	CCCAAGCTTTTCGTTACCATCTACCCTAATTCTG	

R126A-RP	GAGGGATCCTTCGTTACCATCTACCCTAATTGCG ACCATTGTTTCC	Reverse primer to introduce R126A mutation in ScDad2
R128A-RP	GAGGGATCCTTCGTTACCATCTACGGCAATTCTG ACCATTGTTTCC	Reverse primer to introduce R128A mutation in ScDad2
ScdelDSS-RP	GAGGGATCCTTCGTTACCATCCAAGGGTACCAG ATCTTC	Reverse primer to delete DSS in ScDad2
ScGFP-FP	CATGGATCCATGGTGAGCAAGGGCG	Primers to amplify GFP with CYC1 terminator from pYES2CT
ScGFP-RP	GCTGATCGATGCAAATTAAAGCCTTCGAGC	

List of strains utilized in this study

Strains generated/used for *Malassezia* work	
ATCC42132	*Wild-type Malassezia sympodialis*
MSY001	GFP-Mtw1-NAT in *M. sympodialis* ATCC42132
CBS7966	*Wild-type Malassezia globosa*
CBS7956	*Wild-type Malassezia slooffiae*
CBS14141	*Wild-type Malassezia furfur*
MF001	*CENPA-3xFLAG-NAT in M. furfur JLKP23*
BY4741	*Wild-type S. cerevisiae strain (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*)
Strains generated/used for Dad2 work	
***C. albicans* strains**	

J108	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3)* (THAKUR AND SANYAL 2011)
J108C	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1 (NAT)*
CaSR100	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1 (Dad2pr-DAD2^{FL}-TAP-NAT)*
CaSR101	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2$^{\Delta DSS}$-TAP-NAT)*
CaSR102	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2^{P87A}-TAP-NAT)*
CaSR103	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2^{R927A}-TAP -NAT)*
CaSR104	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2^{R947A}-TAP-NAT)*
J108N	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1 (NAT) NDC80/NDC80-mCherry (ARG4)*
CaSR200	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2^{FL}-TAP-NAT) NDC80/NDC80-mCherry (ARG4)*
CaSR201	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2$^{\Delta DSS}$-TAP-NAT) NDC80/NDC80-mCherry (ARG4)*
CaSR202	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2^{P87A}-TAP-NAT) NDC80/NDC80-mCherry (ARG4)*

CaSR203	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2^{R927A}-TAP-NAT) NDC80/NDC80-mCherry (ARG4)*
CaSR204	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2^{R947A}-TAP-NAT) NDC80/NDC80-mCherry (ARG4)*
J108T	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1 (NAT) TUB1/TUB1-mCherry (ARG4)*
CaSR300	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2^{FL}-TAP-NAT) TUB1/TUB1-mCherry (ARG4)*
CaSR301	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2$^{\Delta DSS}$-TAP-NAT) TUB1/TUB1-mCherry (ARG4)*
CaSR302	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2^{P87A}-TAP-NAT) TUB1/TUB1-mCherry (ARG4)*
CaSR303	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2^{R927A}-TAP-NAT) TUB1/TUB1-mCherry (ARG4)*
CaSR304	*Δura3::imm434/ Δura3::imm434 Δhis1::hisG/ Δhis1::hisG Δarg4::hisG/ Δarg4::hisG dad2::HIS1/PCK1*pr*DAD2(URA3) RPS1/RPS1(Dad2pr-DAD2^{R947A}-TAP-NAT) TUB1/TUB1-mCherry (ARG4)*
***S. cerevisiae* strains**	
CJY077	*MAT*a *Δdad2::KanMX6 ura3-52 lys2-801 ade2-101 trp1Δ63 leu2Δ1::pCJ055(DAD2^{TS}, LEU2) his3Δ200* (JANKE et al. 2002)

ScSR-EV	*MAT*a *Δdad2::KanMX6 ura3-52 lys2-801 ade2-101 trp1Δ63 leu2Δ1::pCJ055 his3Δ200::pRS313G (GFP, HIS3)*
ScSR01	*MAT*a *Δdad2::KanMX6 ura3-52 lys2-801 ade2-101 trp1Δ63 leu2Δ1::pCJ055 his3Δ200::pScDad2^{FL} (DAD2^{FL}GFP, HIS3)*
ScSR02	*MAT*a *Δdad2::KanMX6 ura3-52 lys2-801 ade2-101 trp1Δ63 leu2Δ1::pCJ055 his3Δ200:: pScDad2^{R126A} (DAD2^{R126A}GFP, HIS3)*
ScSR03	*MAT*a *Δdad2::KanMX6 ura3-52 lys2-801 ade2-101 trp1Δ63 leu2Δ1::pCJ055 his3Δ200:: pScDad2^{R128A} (DAD2^{R186A}GFP, HIS3)*
ScSR04	*MAT*a *Δdad2::KanMX6 ura3-52 lys2-801 ade2-101 trp1Δ63 leu2Δ1::pCJ055 his3Δ200:: pScDad2$^{\Delta DSS}$ (DAD2$^{\Delta DSS}$GFP, HIS3)*

References

Abad, M. A., B. Medina, A. Santamaria, J. Zou, C. Plasberg-Hill *et al.*, 2014 Structural basis for microtubule recognition by the human kinetochore Ska complex. Nat Commun 5**:** 2964.

Abad, M. A., J. Zou, B. Medina-Pritchard, E. A. Nigg, J. Rappsilber *et al.*, 2016 Ska3 Ensures Timely Mitotic Progression by Interacting Directly With Microtubules and Ska1 Microtubule Binding Domain. Sci Rep 6**:** 34042.

Abi-Rached, L., A. Gilles, T. Shiina, P. Pontarotti and H. Inoko, 2002 Evidence of en bloc duplication in vertebrate genomes. Nat Genet 31**:** 100-105.

Agarwal, M., G. Mehta and S. K. Ghosh, 2015 Role of Ctf3 and COMA subcomplexes in meiosis: Implication in maintaining Cse4 at the centromere and numeric spindle poles. Biochim Biophys Acta 1853**:** 671-684.

Akiyoshi, B., and K. Gull, 2014 Discovery of unconventional kinetochores in kinetoplastids. Cell 156**:** 1247-1258.

Akiyoshi, B., K. K. Sarangapani, A. F. Powers, C. R. Nelson, S. L. Reichow *et al.*, 2010 Tension directly stabilizes reconstituted kinetochore-microtubule attachments. Nature 468**:** 576-579.

Alkan, C., M. Ventura, N. Archidiacono, M. Rocchi, S. C. Sahinalp *et al.*, 2007 Organization and evolution of primate centromeric DNA from whole-genome shotgun sequence data. PLoS Comput Biol 3**:** 1807-1818.

Alushin, G. M., V. H. Ramey, S. Pasqualato, D. A. Ball, N. Grigorieff *et al.*, 2010 The Ndc80 kinetochore complex forms oligomeric arrays along microtubules. Nature 467**:** 805-810.

Amano, M., A. Suzuki, T. Hori, C. Backer, K. Okawa *et al.*, 2009 The CENP-S complex is essential for the stable assembly of outer kinetochore structure. J Cell Biol 186**:** 173-182.

Amaro, A. C., C. P. Samora, R. Holtackers, E. Wang, I. J. Kingston *et al.*, 2010 Molecular control of kinetochore-microtubule dynamics and chromosome oscillations. Nat Cell Biol 12**:** 319-329.

Amend, A., 2014 From dandruff to deep-sea vents: *Malassezia*-like fungi are ecologically hyper-diverse. PLoS Pathog 10**:** e1004277.

Anderson, M. Z., A. Saha, A. Haseeb and R. J. Bennett, 2017 A chromosome 4 trisomy contributes to increased fluconazole resistance in a clinical isolate of Candida albicans. Microbiology 163**:** 856-865.

Arenz, B. E., B. W. Held, J. A. Jurgens, R. L. Farrell and R. A. Blanchette, 2006 Fungal diversity in soils and historic wood from the Ross Sea Region of Antarctica. Soil Biology and Biochemistry 38**:** 3057-3064.

Asakura, K., S. Iwaguchi, M. Homma, T. Sukai, K. Higashide *et al.*, 1991 Electrophoretic karyotypes of clinically isolated yeasts of Candida albicans and C. glabrata. J Gen Microbiol 137**:** 2531-2538.

Ashbee, H. R., and E. G. Evans, 2002 Immunology of diseases associated with Malassezia species. Clin Microbiol Rev 15**:** 21-57.

Aykut, B., S. Pushalkar, R. Chen, Q. Li, R. Abengozar *et al.*, 2019 The fungal mycobiome promotes pancreatic oncogenesis via activation of MBL. Nature 574**:** 264-267.

Bachewich, C., D. Y. Thomas and M. Whiteway, 2003 Depletion of a polo-like kinase in Candida albicans activates cyclase-dependent hyphal-like growth. Mol Biol Cell 14**:** 2163-2180.

Bailey, A. O., T. Panchenko, K. M. Sathyan, J. J. Petkowski, P. J. Pai *et al.*, 2013 Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. Proc Natl Acad Sci U S A 110**:** 11827-11832.

Bailey, J. A., Z. Gu, R. A. Clark, K. Reinert, R. V. Samonte *et al.*, 2002 Recent segmental duplications in the human genome. Science 297**:** 1003-1007.

Baillon, H., 1889 Traité de botanique médicale cryptogami que.

Baker, R. E., and K. Rogers, 2006 Phylogenetic analysis of fungal centromere H3 proteins. Genetics 174**:** 1481-1492.

Baker, R. J., and J. W. Bickham, 1980 Karyotypic evolution in bats: evidence of extensive and conservative chromosomal evolution in closely related taxa. Systematic Biology 29**:** 239-253.

Bakkeren, G., and J. W. Kronstad, 1994 Linkage of mating-type loci distinguishes bipolar from tetrapolar mating in basidiomycetous smut fungi. Proc Natl Acad Sci U S A 91**:** 7085-7089.

Bancroft, J., P. Auckland, C. P. Samora and A. D. McAinsh, 2015 Chromosome congression is promoted by CENP-Q- and CENP-E-dependent pathways. J Cell Sci 128**:** 171-184.

Banerjee, M., D. S. Thompson, A. Lazzell, P. L. Carlisle, C. Pierce *et al.*, 2008 UME6, a novel filament-specific regulator of Candida albicans hyphal extension and virulence. Mol Biol Cell 19**:** 1354-1365.

Barfatani, M., R. J. Munn and O. A. Schjeide, 1964 An Ultrastructure Study of Pityrosporum Orbiculare. J Invest Dermatol 43**:** 231-233.

Barnhart, M. C., P. H. Kuich, M. E. Stellfox, J. A. Ward, E. A. Bassett *et al.*, 2011 HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. J Cell Biol 194**:** 229-243.

Barra, V., and D. Fachinetti, 2018 The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. Nat Commun 9**:** 4340.

Barraclough, T. G., C. W. Birky, Jr. and A. Burt, 2003 Diversification in sexual and asexual organisms. Evolution 57**:** 2166-2172.

Battaglia, E., 2003 Centromere, kinetochore, kinochore, kinetosome, kinosome, kinetomere, kinomere, kinetocentre, kinocentre: history, etymology and intepretation. Caryologia 56**:** 1-21.

Baum, M., V. K. Ngan and L. Clarke, 1994 The centromeric K-type repeat and the central core are together sufficient to establish a functional *Schizosaccharomyces pombe* centromere. Mol Biol Cell 5**:** 747-761.

Baum, M., K. Sanyal, P. K. Mishra, N. Thaler and J. Carbon, 2006 Formation of functional centromeric chromatin is specified epigenetically in Candida albicans. Proc Natl Acad Sci U S A 103**:** 14877-14882.

Begerow, D., M. Stoll and R. Bauer, 2006 A phylogenetic hypothesis of Ustilaginomycotina based on multiple gene analyses and morphological data. Mycologia 98**:** 906-916.

Bennett, R. J., 2015 The parasexual lifestyle of Candida albicans. Curr Opin Microbiol 28**:** 10-17.

Bennett, R. J., and A. D. Johnson, 2003 Completion of a parasexual cycle in Candida albicans by induced chromosome loss in tetraploid strains. EMBO J 22**:** 2505-2515.

Bensasson, D., J. Dicks, J. M. Ludwig, C. J. Bond, A. Elliston *et al.*, 2019 Diverse Lineages of Candida albicans Live on Old Oaks. Genetics 211**:** 277-288.

Bensasson, D., M. Zarowiecki, A. Burt and V. Koufopanou, 2008 Rapid evolution of yeast centromeres in the absence of drive. Genetics 178**:** 2161-2167.

Bergmann, J. H., J. N. Jakubsche, N. M. Martins, A. Kagansky, M. Nakano *et al.*, 2012 Epigenetic engineering: histone H3K9 acetylation is compatible with kinetochore structure and function. J Cell Sci 125**:** 411-421.

Berman, J., 2006 Morphogenesis and cell cycle progression in Candida albicans. Curr Opin Microbiol 9**:** 595-601.

Berman, J., 2012 Candida albicans. Curr Biol 22**:** R620-622.

Berman, J., 2016 Ploidy plasticity: a rapid and reversible strategy for adaptation to stress. FEMS Yeast Res 16.

Bernad, R., P. Sanchez, T. Rivera, M. Rodriguez-Corsino, E. Boyarchuk *et al.*, 2011 Xenopus HJURP and condensin II are required for CENP-A assembly. J Cell Biol 192**:** 569-582.

Biggins, S., 2013 The composition, functions, and regulation of the budding yeast kinetochore. Genetics 194**:** 817-846.

Biggins, S., and A. W. Murray, 2001 The budding yeast protein kinase Ipl1/Aurora allows the absence of tension to activate the spindle checkpoint. Genes Dev 15**:** 3118-3129.

Biscotti, M. A., E. Olmo and J. S. Heslop-Harrison, 2015 Repetitive DNA in eukaryotic genomes. Chromosome Res 23**:** 415-420.

Black, B. E., D. R. Foltz, S. Chakravarthy, K. Luger, V. L. Woods, Jr. *et al.*, 2004 Structural determinants for generating centromeric chromatin. Nature 430**:** 578-582.

Black, B. E., L. E. Jansen, P. S. Maddox, D. R. Foltz, A. B. Desai *et al.*, 2007 Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. Mol Cell 25**:** 309-322.

Blower, M. D., 2016 Centromeric Transcription Regulates Aurora-B Localization and Activation. Cell Rep 15**:** 1624-1633.

Blower, M. D., B. A. Sullivan and G. H. Karpen, 2002 Conserved organization of centromeric chromatin in flies and humans. Dev Cell 2**:** 319-330.

Bobkov, G. O. M., N. Gilbert and P. Heun, 2018 Centromere transcription allows CENP-A to transit from chromatin association to stable incorporation. J Cell Biol 217**:** 1957-1972.

Boekhout, T., and R. W. Bosboom, 1994 Karyotyping of *Malassezia* yeasts - taxonomic and epidemiologic implications. Systematic and Applied Microbiology 17**:** 146-153.

Boekhout, T., E. Guého-Kellermann, P. Mayser and A. Velegraki, 2010 *Malassezia and the skin: science and clinical practice*. Springer Science & Business Media.

Boekhout, T., M. Kamp and E. Gueho, 1998 Molecular typing of *Malassezia* species with PFGE and RAPD. Med Mycol 36**:** 365-372.

Boekhout, T. a. N., E.S.,, 2003 *Pulsed Field Gel Electrophoresis (PFGE) of Yeasts*.

Bougnoux, M. E., C. Pujol, D. Diogo, C. Bouchier, D. R. Soll *et al.*, 2008 Mating is rare within as well as between clades of the human pathogen Candida albicans. Fungal Genet Biol 45**:** 221-231.

Bowers, J. E., B. A. Chapman, J. Rong and A. H. Paterson, 2003 Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422**:** 433-438.

Braun, B. R., M. van Het Hoog, C. d'Enfert, M. Martchenko, J. Dungan *et al.*, 2005 A human-curated annotation of the Candida albicans genome. PLoS Genet 1**:** 36-57.

Brown, C. J., K. M. Todd and R. F. Rosenzweig, 1998 Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. Mol Biol Evol 15**:** 931-942.

Brown, G. D., D. W. Denning, N. A. Gow, S. M. Levitz, M. G. Netea *et al.*, 2012 Hidden killers: human fungal infections. Sci Transl Med 4**:** 165rv113.

Brown, K. S., Jr., B. Von Schoultz and E. Suomalainen, 2004 Chromosome evolution in Neotropical Danainae and Ithomiinae (Lepidoptera). Hereditas 141**:** 216-236.

Burrack, L. S., S. E. Applen and J. Berman, 2011 The requirement for the Dam1 complex is dependent upon the number of kinetochore proteins and microtubules. Curr Biol 21**:** 889-896.

Burrack, L. S., H. F. Hutton, K. J. Matter, S. A. Clancey, I. Liachko *et al.*, 2016 Neocentromeres Provide Chromosome Segregation Accuracy and Centromere Clustering to Multiple Loci along a Candida albicans Chromosome. PLoS Genet 12**:** e1006317.

Butler, G., M. D. Rasmussen, M. F. Lin, M. A. Santos, S. Sakthikumar *et al.*, 2009 Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature 459**:** 657-662.

Buttrick, G. J., and J. B. Millar, 2011 Ringing the changes: emerging roles for DASH at the kinetochore-microtubule Interface. Chromosome Res 19**:** 393-407.

Byrd, A. L., Y. Belkaid and J. A. Segre, 2018 The human skin microbiome. Nat Rev Microbiol 16**:** 143-155.

Byrne, K. P., and K. H. Wolfe, 2005 The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. Genome Res 15**:** 1456-1461.

Cabanes, F. J., 2014 Malassezia yeasts: how many species infect humans and animals? PLoS Pathog 10**:** e1003892.

Cabanes, F. J., S. D. Coutinho, L. Puig, M. R. Bragulat and G. Castella, 2016 New lipid-dependent Malassezia species from parrots. Rev Iberoam Micol 33**:** 92-99.

Cabanes, F. J., B. Theelen, G. Castella and T. Boekhout, 2007 Two new lipid-dependent Malassezia species from domestic animals. FEMS Yeast Res 7**:** 1064-1076.

Cabanes, F. J., S. Vega and G. Castella, 2011 Malassezia cuniculi sp. nov., a novel yeast species isolated from rabbit skin. Med Mycol 49**:** 40-48.

Cai, M., and R. W. Davis, 1990 Yeast centromere binding protein CBF1, of the helix-loop-helix protein family, is required for chromosome stability and methionine prototrophy. Cell 61**:** 437-446.

Cai, M. J., and R. W. Davis, 1989 Purification of a yeast centromere-binding protein that is able to distinguish single base-pair mutations in its recognition site. Mol Cell Biol 9**:** 2544-2550.

Caldas, G. V., and J. G. DeLuca, 2014 KNL1: bringing order to the kinetochore. Chromosoma 123**:** 169-181.

Camahort, R., B. Li, L. Florens, S. K. Swanson, M. P. Washburn *et al.*, 2007 Scm3 is essential to recruit the histone h3 variant cse4 to centromeres and to maintain a functional kinetochore. Mol Cell 26**:** 853-865.

Carlsten, J. O., Z. Szilagyi, B. Liu, M. D. Lopez, E. Szaszi *et al.*, 2012 Mediator promotes CENP-a incorporation at fission yeast centromeres. Mol Cell Biol 32**:** 4035-4043.

Carlton, J. M., S. V. Angiuoli, B. B. Suh, T. W. Kooij, M. Pertea *et al.*, 2002 Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii. Nature 419**:** 512-519.

Carmena, M., M. Wheelock, H. Funabiki and W. C. Earnshaw, 2012 The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. Nat Rev Mol Cell Biol 13**:** 789-803.

Carroll, C. W., K. J. Milks and A. F. Straight, 2010 Dual recognition of CENP-A nucleosomes is required for centromere assembly. J Cell Biol 189**:** 1143-1155.

Carroll, C. W., M. C. Silva, K. M. Godek, L. E. Jansen and A. F. Straight, 2009 Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. Nat Cell Biol 11**:** 896-902.

Casselton, L. A., and N. S. Olesnicky, 1998 Molecular genetics of mating recognition in basidiomycete fungi. Microbiol Mol Biol Rev 62**:** 55-70.

Celis, A. M., A. M. Vos, S. Triana, C. A. Medina, N. Escobar *et al.*, 2017 Highly efficient transformation system for *Malassezia furfur* and *Malassezia pachydermatis* using *Agrobacterium tumefaciens*-mediated transformation. J Microbiol Methods 134**:** 1-6.

Chan, F. L., O. J. Marshall, R. Saffery, B. W. Kim, E. Earle *et al.*, 2012a Active transcription and essential role of RNA polymerase II at the centromere during mitosis. Proc Natl Acad Sci U S A 109**:** 1979-1984.

Chan, F. L., and L. H. Wong, 2012 Transcription in the maintenance of centromere chromatin identity. Nucleic Acids Res 40**:** 11178-11188.

Chan, Y. W., A. A. Jeyaprakash, E. A. Nigg and A. Santamaria, 2012b Aurora B controls kinetochore-microtubule attachments by inhibiting Ska complex-KMN network interaction. J Cell Biol 196**:** 563-571.

Chatterjee, G., S. R. Sankaranarayanan, K. Guin, Y. Thattikota, S. Padmanabhan *et al.*, 2016 Repeat-associated fission yeast-like regional centromeres in the ascomycetous budding yeast *Candida tropicalis*. PLoS Genet 12**:** e1005839.

Cheeseman, I. M., 2014 The kinetochore. Cold Spring Harb Perspect Biol 6**:** a015826.

Cheeseman, I. M., S. Anderson, M. Jwa, E. M. Green, J. Kang *et al.*, 2002 Phospho-regulation of kinetochore-microtubule attachments by the Aurora kinase Ipl1p. Cell 111**:** 163-172.

Cheeseman, I. M., C. Brew, M. Wolyniak, A. Desai, S. Anderson *et al.*, 2001a Implication of a novel multiprotein Dam1p complex in outer kinetochore function. J Cell Biol 155**:** 1137-1145.

Cheeseman, I. M., J. S. Chappie, E. M. Wilson-Kubalek and A. Desai, 2006 The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. Cell 127**:** 983-997.

Cheeseman, I. M., and A. Desai, 2008 Molecular architecture of the kinetochore-microtubule interface. Nat Rev Mol Cell Biol 9**:** 33-46.

Cheeseman, I. M., M. Enquist-Newman, T. Muller-Reichert, D. G. Drubin and G. Barnes, 2001b Mitotic spindle integrity and kinetochore function linked by the Duo1p/Dam1p complex. J Cell Biol 152**:** 197-212.

Chen, C. C., S. Bowers, Z. Lipinszki, J. Palladino, S. Trusiak *et al.*, 2015 Establishment of Centromeric Chromatin by the CENP-A Assembly Factor CAL1 Requires FACT-Mediated Transcription. Dev Cell 34**:** 73-84.

Chen, C. C., M. L. Dechassa, E. Bettini, M. B. Ledoux, C. Belisario *et al.*, 2014 CAL1 is the Drosophila CENP-A assembly factor. J Cell Biol 204**:** 313-329.

Chen, E. S., S. Saitoh, M. Yanagida and K. Takahashi, 2003 A cell cycle-regulated GATA factor promotes centromeric localization of CENP-A in fission yeast. Mol Cell 11**:** 175-187.

Chen, Z., Y. Omori, S. Koren, T. Shirokiya, T. Kuroda *et al.*, 2019 De novo assembly of the goldfish (Carassius auratus) genome and the evolution of genes after whole-genome duplication. Sci Adv 5**:** eaav0547.

Chibana, H., S. Iwaguchi, M. Homma, A. Chindamporn, Y. Nakagawa *et al.*, 1994 Diversity of tandemly repetitive sequences due to short periodic repetitions in the chromosomes of Candida albicans. J Bacteriol 176**:** 3851-3858.

Chikashige, Y., N. Kinoshita, Y. Nakaseko, T. Matsumoto, S. Murakami *et al.*, 1989 Composite motifs and repeat symmetry in *S. pombe* centromeres: direct analysis by integration of NotI restriction sites. Cell 57**:** 739-751.

Chindamporn, A., Y. Nakagawa, I. Mizuguchi, H. Chibana, M. Doi *et al.*, 1998 Repetitive sequences (RPSs) in the chromosomes of Candida albicans are sandwiched between two novel stretches, HOK and RB2, common to each chromosome. Microbiology 144 (Pt 4)**:** 849-857.

Cho, U. S., and S. C. Harrison, 2011 Ndc10 is a platform for inner kinetochore assembly in budding yeast. Nat Struct Mol Biol 19**:** 48-55.

Cho, Y.-J., M. Park and W. H. Jung, 2019 Resequencing the Genome of *Malassezia restricta* Strain KCTC 27527. Microbiology Resource Announcements 8**:** e00213-00219.

Choi, E. S., A. Stralfors, A. G. Castillo, M. Durand-Dubief, K. Ekwall *et al.*, 2011 Identification of noncoding transcripts from within CENP-A chromatin at fission yeast centromeres. J Biol Chem 286**:** 23600-23607.

Choo, K. A., 1997 *The centromere*. Oxford University Press Oxford.

Chow, E. W., C. A. Morrow, J. T. Djordjevic, I. A. Wood and J. A. Fraser, 2012 Microevolution of Cryptococcus neoformans driven by massive tandem gene amplification. Mol Biol Evol 29**:** 1987-2000.

Chueh, A. C., E. L. Northrop, K. H. Brettingham-Moore, K. H. Choo and L. H. Wong, 2009 LINE retrotransposon RNA is an essential structural and functional epigenetic component of a core neocentromeric chromatin. PLoS Genet 5**:** e1000354.

Ciferri, C., J. De Luca, S. Monzani, K. J. Ferrari, D. Ristic *et al.*, 2005 Architecture of the human ndc80-hec1 complex, a critical constituent of the outer kinetochore. J Biol Chem 280**:** 29088-29095.

Ciferri, C., S. Pasqualato, E. Screpanti, G. Varetti, S. Santaguida *et al.*, 2008 Implications for kinetochore-microtubule attachment from the structure of an engineered Ndc80 complex. Cell 133**:** 427-439.

Clarke, L., and M. P. Baum, 1990 Functional analysis of a centromere from fission yeast: a role for centromere-specific repeated DNA sequences. Mol Cell Biol 10**:** 1863-1872.

Clarke, L., and J. Carbon, 1980 Isolation of a yeast centromere and construction of functional small circular chromosomes. Nature 287**:** 504-509.

Coelho, M. A., J. P. Sampaio and P. Goncalves, 2010 A deviation from the bipolar-tetrapolar mating paradigm in an early diverged basidiomycete. PLoS Genet 6.

Coelho, M. A., J. P. Sampaio and P. Goncalves, 2013 Living and thriving on the skin: Malassezia genomes tell the story. mBio 4**:** e00117-00113.

Coghlan, A., and K. H. Wolfe, 2002 Fourfold faster rate of genome rearrangement in nematodes than in Drosophila. Genome Res 12**:** 857-867.

Cohen, A. K., T. Y. Huh and C. W. Helleiner, 1973 Transcription of satellite DNA in mouse L-cells. Can J Biochem 51**:** 529-532.

Cohen, R. L., C. W. Espelin, P. De Wulf, P. K. Sorger, S. C. Harrison *et al.*, 2008 Structural and functional dissection of Mif2p, a conserved DNA-binding kinetochore protein. Mol Biol Cell 19**:** 4480-4491.

Comai, L., 2014 Genome elimination: translating basic research into a future tool for plant breeding. PLoS Biol 12**:** e1001876.

Conant, G. C., and K. H. Wolfe, 2007 Increased glycolytic flux as an outcome of whole-genome duplication in yeast. Mol Syst Biol 3**:** 129.

Contreras-Moreira, B., and P. Vinuesa, 2013 GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. Appl Environ Microbiol 79**:** 7696-7701.

Cook, L. G., 2000 Extraordinary and extensive karyotypic variation: a 48-fold range in chromosome number in the gall-inducing scale insect Apiomorpha (Hemiptera: Coccoidea: Eriococcidae). Genome 43**:** 255-263.

Corrochano, L. M., A. Kuo, M. Marcet-Houben, S. Polaino, A. Salamov *et al.*, 2016 Expansion of Signal Transduction Pathways in Fungi by Extensive Genome Duplication. Curr Biol 26**:** 1577-1584.

Darlington, C. D., 1937 Recent advances in cytology. Recent advances in cytology.

de Hoog, S., M. Monod, T. Dawson, T. Boekhout, P. Mayser *et al.*, 2017 Skin Fungi from Colonization to Infection. Microbiol Spectr 5.

De Wulf, P., A. D. McAinsh and P. K. Sorger, 2003 Hierarchical assembly of the budding yeast kinetochore from multiple subcomplexes. Genes Dev 17**:** 2902-2921.

Dehal, P., P. Predki, A. S. Olsen, A. Kobayashi, P. Folta *et al.*, 2001 Human chromosome 19 and related regions in mouse: conservative and lineage-specific evolution. Science 293**:** 104-111.

DeLuca, K. F., S. M. Lens and J. G. DeLuca, 2011 Temporal changes in Hec1 phosphorylation control kinetochore-microtubule attachment stability during mitosis. J Cell Sci 124**:** 622-634.

Dernburg, A. F., 2001 Here, there, and everywhere: kinetochore function on holocentric chromosomes. J Cell Biol 153**:** F33-38.

Dimitrova, Y. N., S. Jenni, R. Valverde, Y. Khin and S. C. Harrison, 2016 Structure of the MIND Complex Defines a Regulatory Focus for Yeast Kinetochore Assembly. Cell 167**:** 1014-1027 e1012.

Dorn, M., and K. Roehnert, 1977 Dimorphism of Pityrosporum orbiculare in a defined culture medium. J Invest Dermatol 69**:** 244-248.

Drinnenberg, I. A., D. deYoung, S. Henikoff and H. S. Malik, 2014 Recurrent loss of CenH3 is associated with independent transitions to holocentricity in insects. Elife 3.

Du, Y., C. N. Topp and R. K. Dawe, 2010 DNA binding of centromere protein C (CENPC) is stabilized by single-stranded RNA. PLoS Genet 6**:** e1000835.

Dumesic, P. A., C. M. Homer, J. J. Moresco, L. R. Pack, E. K. Shanle *et al.*, 2015 Product binding enforces the genomic specificity of a yeast polycomb repressive complex. Cell 160**:** 204-218.

Dumitru, R., D. H. Navarathna, C. P. Semighini, C. G. Elowsky, R. V. Dumitru *et al.*, 2007 In vivo and in vitro anaerobic mating in Candida albicans. Eukaryot Cell 6**:** 465-472.

Dunleavy, E. M., D. Roche, H. Tagami, N. Lacoste, D. Ray-Gallet *et al.*, 2009 HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell 137**:** 485-497.

Earnshaw, W. C., 2015 Discovering centromere proteins: from cold white hands to the A, B, C of CENPs. Nat Rev Mol Cell Biol 16**:** 443-449.

Earnshaw, W. C., and B. R. Migeon, 1985 Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome. Chromosoma 92**:** 290-296.

Earnshaw, W. C., and N. Rothfield, 1985 Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. Chromosoma 91**:** 313-321.

Eckert-Boulet, N., M. L. Pedersen, B. O. Krogh and M. Lisby, 2012 Optimization of ordered plasmid assembly by gap repair in *Saccharomyces cerevisiae*. Yeast 29**:** 323-334.

Egozcue, J., 1969 Primates-Comparative mammalian cytogenetics, pp. 357-389. Springer.

Eichler, E. E., 2001 Recent duplication, domain accretion and the dynamic mutation of the human genome. Trends Genet 17**:** 661-669.

Eichler, E. E., and D. Sankoff, 2003 Structural dynamics of eukaryotic chromosome evolution. Science 301**:** 793-797.

Eichstedt, 1846 Pilzbildung in der *Pityriasis versicolor*. Froriep Neue Notiz Natur Heilk

39:270.

Enquist-Newman, M., I. M. Cheeseman, D. Van Goor, D. G. Drubin, P. B. Meluh *et al.*, 2001 Dad1p, third component of the Duo1p/Dam1p complex involved in kinetochore function and mitotic spindle integrity. Mol Biol Cell 12**:** 2601-2613.

Euskirchen, G. M., 2002 Nnf1p, Dsn1p, Mtw1p, and Nsl1p: a new group of proteins important for chromosome segregation in *Saccharomyces cerevisiae*. Eukaryot Cell 1**:** 229-240.

Eymery, A., M. Callanan and C. Vourc'h, 2009 The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. Int J Dev Biol 53**:** 259-268.

Fachinetti, D., H. D. Folco, Y. Nechemia-Arbely, L. P. Valente, K. Nguyen *et al.*, 2013 A two-step mechanism for epigenetic specification of centromere identity and function. Nat Cell Biol 15**:** 1056-1066.

Fernius, J., and A. L. Marston, 2009 Establishment of cohesion at the pericentromere by the Ctf19 kinetochore subcomplex and the replication fork-associated factor, Csm3. PLoS Genet 5**:** e1000629.

Fierro, F., and J. F. Martin, 1999 Molecular mechanisms of chromosomal rearrangement in fungi. Crit Rev Microbiol 25**:** 1-17.

Findley, K., and E. A. Grice, 2014 The skin microbiome: a focus on pathogens and their association with skin disease. PLoS Pathog 10**:** e1004436.

Findley, K., J. Oh, J. Yang, S. Conlan, C. Deming *et al.*, 2013 Topographic diversity of fungal and bacterial communities in human skin. Nature 498**:** 367-370.

Fleig, U., M. Sen-Gupta and J. H. Hegemann, 1996 Fission yeast mal2+ is required for chromosome segregation. Mol Cell Biol 16**:** 6169-6177.

Flemming, W., 1882 *Zellsubstanz, kern und zelltheilung*. Vogel.

Folco, H. D., A. L. Pidoux, T. Urano and R. C. Allshire, 2008 Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. Science 319**:** 94-97.

Foley, E. A., and T. M. Kapoor, 2013 Microtubule attachment and spindle assembly checkpoint signalling at the kinetochore. Nat Rev Mol Cell Biol 14**:** 25-37.

Foltz, D. R., L. E. Jansen, A. O. Bailey, J. R. Yates, 3rd, E. A. Bassett *et al.*, 2009 Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. Cell 137**:** 472-484.

Foltz, D. R., L. E. Jansen, B. E. Black, A. O. Bailey, J. R. Yates, 3rd *et al.*, 2006 The human CENP-A centromeric nucleosome-associated complex. Nat Cell Biol 8**:** 458-469.

Forche, A., D. Abbey, T. Pisithkul, M. A. Weinzierl, T. Ringstrom *et al.*, 2011 Stress alters rates and types of loss of heterozygosity in Candida albicans. mBio 2.

Forche, A., K. Alby, D. Schaefer, A. D. Johnson, J. Berman *et al.*, 2008 The parasexual cycle in Candida albicans provides an alternative pathway to meiosis for the formation of recombinant strains. PLoS Biol 6**:** e110.

Franck, A. D., A. F. Powers, D. R. Gestaut, T. Gonen, T. N. Davis *et al.*, 2007 Tension applied through the Dam1 complex promotes microtubule elongation providing a direct mechanism for length control in mitosis. Nat Cell Biol 9**:** 832-837.

Fredga, K., 1977 Chromosomal changes in vertebrate evolution. Proceedings of the Royal Society of London. Series B. Biological Sciences 199**:** 377-397.

Freire-Beneitez, V., R. J. Price and A. Buscaino, 2016 The chromatin of *Candida albicans* pericentromeres bears features of both euchromatin and heterochromatin. Front Microbiol 7**:** 759.

Friedman, S., and M. Freitag, 2017 Centrochromatin of Fungi. Prog Mol Subcell Biol 56**:** 85-109.

Fukagawa, T., 2017 Critical histone post-translational modifications for centromere function and propagation. Cell Cycle 16**:** 1259-1265.

Fukagawa, T., M. Nogami, M. Yoshikawa, M. Ikeno, T. Okazaki *et al.*, 2004 Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nat Cell Biol 6**:** 784-791.

Garcia-Silva, M. R., L. Sollelis, C. R. MacPherson, S. Stanojcic, N. Kuk *et al.*, 2017 Identification of the centromeres of Leishmania major: revealing the hidden pieces. EMBO Rep 18**:** 1968-1977.

Gardner, M. J., N. Hall, E. Fung, O. White, M. Berriman *et al.*, 2002 Genome sequence of the human malaria parasite Plasmodium falciparum. Nature 419**:** 498-511.

Gioti, A., B. Nystedt, W. Li, J. Xu, A. Andersson *et al.*, 2013 Genomic insights into the atopic eczema-associated skin commensal yeast Malassezia sympodialis. mBio 4**:** e00572-00512.

Gonen, S., B. Akiyoshi, M. G. Iadanza, D. Shi, N. Duggan *et al.*, 2012 The structure of purified kinetochores reveals multiple microtubule-attachment sites. Nat Struct Mol Biol 19**:** 925-929.

Gong, Z., Y. Wu, A. Koblizkova, G. A. Torres, K. Wang *et al.*, 2012 Repeatless and repeat-based centromeres in potato: implications for centromere evolution. Plant Cell 24**:** 3559-3574.

Goodwin, T. J., and R. T. Poulter, 2000 Multiple LTR-retrotransposon families in the asexual yeast Candida albicans. Genome Res 10**:** 174-191.

Gordon, J. L., K. P. Byrne and K. H. Wolfe, 2011 Mechanisms of chromosome number evolution in yeast. PLoS Genet 7**:** e1002190.

Goshima, G., T. Kiyomitsu, K. Yoda and M. Yanagida, 2003 Human centromere chromatin protein hMis12, essential for equal segregation, is independent of CENP-A loading pathway. J Cell Biol 160**:** 25-39.

Goshima, G., S. Saitoh and M. Yanagida, 1999 Proper metaphase spindle length is determined by centromere proteins Mis12 and Mis6 required for faithful chromosome segregation. Genes Dev 13**:** 1664-1677.

Gow, N. A. R., 2013 A developmental program for Candida commensalism. Nature Genetics 45**:** 967-968.

Grabherr, M. G., P. Russell, M. Meyer, E. Mauceli, J. Alfoldi *et al.*, 2010 Genome-wide synteny through highly sensitive sequence alignment: Satsuma. Bioinformatics 26**:** 1145-1151.

Graser, Y., M. Volovsek, J. Arrington, G. Schonian, W. Presber *et al.*, 1996 Molecular markers reveal that population structure of the human pathogen Candida albicans exhibits both clonality and recombination. Proc Natl Acad Sci U S A 93**:** 12473-12477.

Greenfeder, S. A., and C. S. Newlon, 1992 Replication forks pause at yeast centromeres. Mol Cell Biol 12**:** 4056-4066.

Gregan, J., S. Polakova, L. Zhang, I. M. Tolic-Norrelykke and D. Cimini, 2011 Merotelic kinetochore attachment: causes and effects. Trends Cell Biol 21**:** 374-381.

Gresham, D., M. M. Desai, C. M. Tucker, H. T. Jenq, D. A. Pai *et al.*, 2008 The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. PLoS Genet 4**:** e1000303.

Grishchuk, E. L., M. I. Molodtsov, F. I. Ataullakhanov and J. R. McIntosh, 2005 Force production by disassembling microtubules. Nature 438**:** 384-388.

Gu, X., Y. Wang and J. Gu, 2002 Age distribution of human gene families shows significant roles of both large- and small-scale duplications in vertebrate evolution. Nat Genet 31**:** 205-209.

Gueho, E., G. Midgley and J. Guillot, 1996 The genus Malassezia with description of four new species. Antonie Van Leeuwenhoek 69**:** 337-355.

Guerrero, A. A., M. C. Gamero, V. Trachana, A. Futterer, C. Pacios-Bras *et al.*, 2010a Centromere-localized breaks indicate the generation of DNA damage by the mitotic spindle. Proc Natl Acad Sci U S A 107**:** 4159-4164.

Guerrero, A. A., A. C. Martinez and K. H. van Wely, 2010b Merotelic attachments and non-homologous end joining are the basis of chromosomal instability. Cell Div 5**:** 13.

Guillot, J., and E. Gueho, 1995 The diversity of Malassezia yeasts confirmed by rRNA sequence and nuclear DNA comparisons. Antonie Van Leeuwenhoek 67**:** 297-314.

Guimaraes, G. J., Y. Dong, B. F. McEwen and J. G. Deluca, 2008 Kinetochore-microtubule attachment relies on the disordered N-terminal tail domain of Hec1. Curr Biol 18**:** 1778-1784.

Guldner, H. H., H. J. Lakomek and F. A. Bautz, 1984 Human anti-centromere sera recognise a 19.5 kD non-histone chromosomal protein from HeLa cells. Clin Exp Immunol 58**:** 13-20.

Gusa, A., and S. Jinks-Robertson, 2019 Mitotic Recombination and Adaptive Genomic Changes in Human Pathogenic Fungi. Genes (Basel) 10.

Guse, A., C. W. Carroll, B. Moree, C. J. Fuller and A. F. Straight, 2011 In vitro centromere and kinetochore assembly on defined chromatin templates. Nature 477**:** 354-358.

Haaf, T., A. G. Mater, J. Wienberg and D. C. Ward, 1995 Presence and abundance of CENP-B box sequences in great ape subsets of primate-specific alpha-satellite DNA. J Mol Evol 41**:** 487-491.

Haaf, T., and H. F. Willard, 1997 Chromosome-specific alpha-satellite DNA from the centromere of chimpanzee chromosome 4. Chromosoma 106**:** 226-232.

Hahnenberger, K. M., M. P. Baum, C. M. Polizzi, J. Carbon and L. Clarke, 1989 Construction of functional artificial minichromosomes in the fission yeast Schizosaccharomyces pombe. Proc Natl Acad Sci U S A 86**:** 577-581.

Hall, I. M., G. D. Shankaranarayana, K. Noma, N. Ayoub, A. Cohen *et al.*, 2002 Establishment and maintenance of a heterochromatin domain. Science 297**:** 2232-2237.

Harrington, J. J., G. Van Bokkelen, R. W. Mays, K. Gustashaw and H. F. Willard, 1997 Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. Nat Genet 15**:** 345-355.

Hauffe, H. C., and J. Piálek, 1997 Evolution of the chromosomal races of Mus musculus domesticus in the Rhaetian Alps: the roles of whole-arm reciprocal translocation and zonal raciation. Biological Journal of the Linnean Society 62**:** 255-278.

Hayashi, T., Y. Fujita, O. Iwasaki, Y. Adachi, K. Takahashi *et al.*, 2004 Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. Cell 118**:** 715-729.

Helgeson, L. A., A. Zelter, M. Riffle, M. J. MacCoss, C. L. Asbury *et al.*, 2018 Human Ska complex and Ndc80 complex interact to form a load-bearing assembly that strengthens kinetochore-microtubule attachments. Proc Natl Acad Sci U S A 115**:** 2740-2745.

Helmrich, A., M. Ballarino and L. Tora, 2011 Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. Mol Cell 44**:** 966-977.

Hemmerich, P., T. Stoyan, G. Wieland, M. Koch, J. Lechner *et al.*, 2000 Interaction of yeast kinetochore proteins with centromere-protein/transcription factor Cbf1. Proc Natl Acad Sci U S A 97**:** 12583-12588.

Henikoff, S., K. Ahmad and H. S. Malik, 2001 The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293**:** 1098-1102.

Hickman, M. A., C. Paulson, A. Dudley and J. Berman, 2015 Parasexual Ploidy Reduction Drives Population Heterogeneity Through Random and Transient Aneuploidy in Candida albicans. Genetics 200**:** 781-794.

Hill, A., and K. Bloom, 1987 Genetic manipulation of centromere function. Mol Cell Biol 7**:** 2397-2405.

Hindriksen, S., S. M. A. Lens and M. A. Hadders, 2017 The Ins and Outs of Aurora B Inner Centromere Localization. Front Cell Dev Biol 5**:** 112.

Hinshaw, S. M., and S. C. Harrison, 2013 An Iml3-Chl4 heterodimer links the core centromere to factors required for accurate chromosome segregation. Cell Rep 5**:** 29-36.

Hinshaw, S. M., and S. C. Harrison, 2018 Kinetochore Function from the Bottom Up. Trends Cell Biol 28**:** 22-33.

Hipp, A. L., 2007 Nonuniform processes of chromosome evolution in sedges (Carex: Cyperaceae). Evolution 61**:** 2175-2194.

Hirai, A., R. Kano, K. Makimura, E. R. Duarte, J. S. Hamdan *et al.*, 2004 Malassezia nana sp. nov., a novel lipid-dependent yeast species isolated from animals. Int J Syst Evol Microbiol 54**:** 623-627.

Hofmann, C., I. M. Cheeseman, B. L. Goode, K. L. McDonald, G. Barnes *et al.*, 1998 Saccharomyces cerevisiae Duo1p and Dam1p, novel proteins involved in mitotic spindle function. J Cell Biol 143**:** 1029-1040.

Holliday, R., 1984 The biological significance of meiosis. Symp Soc Exp Biol 38**:** 381-394.

Honnavar, P., A. Chakrabarti, S. Dogra, S. Handa and S. M. Rudramurthy, 2015 Phenotypic and molecular characterization of Malassezia japonica isolated from psoriasis vulgaris patients. J Med Microbiol 64**:** 232-236.

Honnavar, P., G. S. Prasad, A. Ghosh, S. Dogra, S. Handa *et al.*, 2016 Malassezia arunalokei sp. nov., a Novel Yeast Species Isolated from Seborrheic Dermatitis Patients and Healthy Individuals from India. J Clin Microbiol 54**:** 1826-1834.

Hori, T., M. Amano, A. Suzuki, C. B. Backer, J. P. Welburn *et al.*, 2008 CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. Cell 135**:** 1039-1052.

Hori, T., W. H. Shang, A. Toyoda, S. Misu, N. Monma *et al.*, 2014 Histone H4 Lys 20 monomethylation of the CENP-A nucleosome is essential for kinetochore assembly. Dev Cell 29**:** 740-749.

Hornung, P., M. Maier, G. M. Alushin, G. C. Lander, E. Nogales *et al.*, 2011 Molecular architecture and connectivity of the budding yeast Mtw1 kinetochore complex. J Mol Biol 405**:** 548-559.

Hornung, P., P. Troc, F. Malvezzi, M. Maier, Z. Demianova *et al.*, 2014 A cooperative mechanism drives budding yeast kinetochore assembly downstream of CENP-A. J Cell Biol 206**:** 509-524.

Howell, B. J., B. F. McEwen, J. C. Canman, D. B. Hoffman, E. M. Farrar *et al.*, 2001 Cytoplasmic dynein/dynactin drives kinetochore protein transport to the spindle poles and has a role in mitotic spindle checkpoint inactivation. J Cell Biol 155**:** 1159-1172.

Howman, E. V., K. J. Fowler, A. J. Newson, S. Redward, A. C. MacDonald *et al.*, 2000 Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. Proc Natl Acad Sci U S A 97**:** 1148-1153.

Hoyt, M. A., L. Totis and B. T. Roberts, 1991 S. cerevisiae genes required for cell cycle arrest in response to loss of microtubule function. Cell 66**:** 507-517.

Hsueh, Y. P., and J. Heitman, 2008 Orchestration of sexual reproduction and virulence by the fungal mating-type locus. Curr Opin Microbiol 11**:** 517-524.

Huis In 't Veld, P. J., S. Jeganathan, A. Petrovic, P. Singh, J. John *et al.*, 2016 Molecular basis of outer kinetochore assembly on CENP-T. Elife 5.

Hull, C. M., and A. D. Johnson, 1999 Identification of a mating type-like locus in the asexual pathogenic yeast Candida albicans. Science 285**:** 1271-1275.

Hull, R. M., C. Cruz, C. V. Jack and J. Houseley, 2017 Environmental change drives accelerated adaptation through stimulated copy number variation. PLoS Biol 15**:** e2001333.

Ianiri, G., S. Applen Clancey, S. C. Lee and J. Heitman, 2017a FKBP12-dependent inhibition of calcineurin mediates immunosuppressive antifungal drug action in *Malassezia*. mBio 8**:** e01752-01717.

Ianiri, G., A. F. Averette, J. M. Kingsbury, J. Heitman and A. Idnurm, 2016 Gene function analysis in the ubiquitous human commensal and pathogen *Malassezia* genus. mBio 7 , e01853-16.

Ianiri, G., K. J. Boyce and A. Idnurm, 2017b Isolation of conditional mutations in genes essential for viability of *Cryptococcus neoformans*. Curr Genet 63**:** 519-530.

Ianiri, G., M. A. Coelho, F. Ruchti, F. Sparber, T. J. McMahon *et al.*, 2020 Horizontal gene transfer in the human and skin commensal Malassezia: a bacterially-derived flavohemoglobin is required for NO resistance and host interaction. bioRxiv**:** 2020.2001.2028.923367.

Ianiri, G., G. Dagotto, S. Sun and J. Heitman, 2019 Advancing Functional Genetics Through Agrobacterium-Mediated Insertional Mutagenesis and CRISPR/Cas9 in the Commensal and Pathogenic Yeast Malassezia. Genetics 212**:** 1163-1179.

Ideue, T., Y. Cho, K. Nishimura and T. Tani, 2014 Involvement of satellite I noncoding RNA in regulation of chromosome segregation. Genes Cells 19**:** 528-538.

Ikeno, M., B. Grimes, T. Okazaki, M. Nakano, K. Saitoh *et al.*, 1998 Construction of YAC-based mammalian artificial chromosomes. Nat Biotechnol 16**:** 431-439.

International Chicken Genome Sequencing, C., 2004 Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432**:** 695-716.

Ishii, K., Y. Ogiyama, Y. Chikashige, S. Soejima, F. Masuda *et al.*, 2008 Heterochromatin integrity affects chromosome reorganization after centromere dysfunction. Science 321**:** 1088-1091.

Izuta, H., M. Ikeno, N. Suzuki, T. Tomonaga, N. Nozaki *et al.*, 2006 Comprehensive analysis of the ICEN (Interphase Centromere Complex) components enriched in the CENP-A chromatin of human cells. Genes Cells 11**:** 673-684.

Janbon, G., F. Sherman and E. Rustchenko, 1998 Monosomy of a specific chromosome determines L-sorbose utilization: a novel regulatory mechanism in Candida albicans. Proc Natl Acad Sci U S A 95**:** 5150-5155.

Janke, C., J. Ortiz, T. U. Tanaka, J. Lechner and E. Schiebel, 2002 Four new subunits of the Dam1-Duo1 complex reveal novel functions in sister kinetochore biorientation. EMBO J 21**:** 181-193.

Jdo, J., A. Baldini, D. C. Ward, S. T. Reeders and R. A. Wells, 1991 Origin of human chromosome 2: an ancestral telomere-telomere fusion. Proc Natl Acad Sci U S A 88**:** 9051-9055.

Jehn, B., R. Niedenthal and J. H. Hegemann, 1991 In vivo analysis of the Saccharomyces cerevisiae centromere CDEIII sequence: requirements for mitotic chromosome segregation. Mol Cell Biol 11**:** 5212-5221.

Jenni, S., Y. N. Dimitrova, R. Valverde, S. M. Hinshaw and S. C. Harrison, 2017 Molecular Structures of Yeast Kinetochore Subcomplexes and Their Roles in Chromosome Segregation. Cold Spring Harb Symp Quant Biol 82**:** 83-89.

Jenni, S., and S. C. Harrison, 2018 Structure of the DASH/Dam1 complex shows its role at the yeast kinetochore-microtubule interface. Science 360**:** 552-558.

Jeyaprakash, A. A., A. Santamaria, U. Jayachandran, Y. W. Chan, C. Benda *et al.*, 2012 Structural and functional organization of the Ska complex, a key component of the kinetochore-microtubule interface. Mol Cell 46**:** 274-286.

Jiao, Y., N. J. Wickett, S. Ayyampalayam, A. S. Chanderbali, L. Landherr *et al.*, 2011 Ancestral polyploidy in seed plants and angiosperms. Nature 473**:** 97-100.

Jo, J. H., C. Deming, E. A. Kennedy, S. Conlan, E. C. Polley *et al.*, 2016 Diverse Human Skin Fungal Communities in Children Converge in Adulthood. J Invest Dermatol 136**:** 2356-2363.

Joglekar, A. P., D. Bouck, K. Finley, X. Liu, Y. Wan *et al.*, 2008 Molecular architecture of the kinetochore-microtubule attachment site is conserved between point and regional centromeres. J Cell Biol 181**:** 587-594.

Jones, M. H., J. B. Bachant, A. R. Castillo, T. H. Giddings, Jr. and M. Winey, 1999 Yeast Dam1p is required to maintain spindle integrity during mitosis and interacts with the Mps1p kinase. Mol Biol Cell 10**:** 2377-2391.

Joseph, A., A. R. Mitchell and O. J. Miller, 1989 The organization of the mouse satellite DNA at centromeres. Exp Cell Res 183**:** 494-500.

Kabeche, L., H. D. Nguyen, R. Buisson and L. Zou, 2018 A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. Science 359**:** 108-114.

Kalitsis, P., and K. H. Choo, 2012 The evolutionary life cycle of the resilient centromere. Chromosoma 121**:** 327-340.

Kanellopoulou, C., S. A. Muljo, A. L. Kung, S. Ganesan, R. Drapkin *et al.*, 2005 Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev 19**:** 489-501.

Kapoor, S., L. Zhu, C. Froyd, T. Liu and L. N. Rusche, 2015 Regional centromeres in the yeast *Candida lusitaniae* lack pericentromeric heterochromatin. Proc Natl Acad Sci U S A 112**:** 12139-12144.

Kapoor, T. M., M. A. Lampson, P. Hergert, L. Cameron, D. Cimini *et al.*, 2006 Chromosomes can congress to the metaphase plate before biorientation. Science 311**:** 388-391.

Kato, H., J. Jiang, B. R. Zhou, M. Rozendaal, H. Feng *et al.*, 2013 A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. Science 340**:** 1110-1113.

Kellis, M., B. W. Birren and E. S. Lander, 2004 Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. Nature 428**:** 617-624.

Khmelinskii, A., and E. Schiebel, 2008 Assembling the spindle midzone in the right place at the right time. Cell Cycle 7**:** 283-286.

Kiermaier, E., S. Woehrer, Y. Peng, K. Mechtler and S. Westermann, 2009 A Dam1-based artificial kinetochore is sufficient to promote chromosome segregation in budding yeast. Nat Cell Biol 11**:** 1109-1115.

Killinger, K., M. Bohm, P. Steinbach, G. Hagemann, M. Bluggel *et al.*, 2020 Auto-inhibition of Mif2/CENP-C ensures centromere-dependent kinetochore assembly in budding yeast. EMBO J 39**:** e102938.

Kim, M., Y. J. Cho, M. Park, Y. Choi, S. Y. Hwang *et al.*, 2018 Genomic Tandem Quadruplication is Associated with Ketoconazole Resistance in Malassezia pachydermatis. J Microbiol Biotechnol 28**:** 1937-1945.

Kim, T. M., R. Xi, L. J. Luquette, R. W. Park, M. D. Johnson *et al.*, 2013 Functional genomic analysis of chromosomal aberrations in a compendium of 8000 cancer genomes. Genome Res 23**:** 217-227.

Kipling, D., A. R. Mitchell, H. Masumoto, H. E. Wilson, L. Nicol *et al.*, 1995 CENP-B binds a novel centromeric sequence in the Asian mouse Mus caroli. Mol Cell Biol 15**:** 4009-4020.

Kiyomitsu, T., H. Murakami and M. Yanagida, 2011 Protein interaction domain mapping of human kinetochore protein Blinkin reveals a consensus motif for binding of spindle assembly checkpoint proteins Bub1 and BubR1. Mol Cell Biol 31**:** 998-1011.

Kiyomitsu, T., C. Obuse and M. Yanagida, 2007 Human Blinkin/AF15q14 is required for chromosome alignment and the mitotic checkpoint through direct interaction with Bub1 and BubR1. Dev Cell 13**:** 663-676.

Klare, K., J. R. Weir, F. Basilico, T. Zimniak, L. Massimiliano *et al.*, 2015 CENP-C is a blueprint for constitutive centromere-associated network assembly within human kinetochores. J Cell Biol 210**:** 11-22.

Klosterman, S. J., K. V. Subbarao, S. Kang, P. Veronese, S. E. Gold *et al.*, 2011 Comparative genomics yields insights into niche adaptation of plant vascular wilt pathogens. PLoS Pathog 7**:** e1002137.

Kobayashi, N., Y. Suzuki, L. W. Schoenfeld, C. A. Muller, C. Nieduszynski *et al.*, 2015 Discovery of an unconventional centromere in budding yeast redefines evolution of point centromeres. Curr Biol 25**:** 2026-2033.

Koo, D. H., F. Han, J. A. Birchler and J. Jiang, 2011 Distinct DNA methylation patterns associated with active and inactive centromeres of the maize B chromosome. Genome Res 21**:** 908-914.

Kops, G. J., B. A. Weaver and D. W. Cleveland, 2005 On the road to cancer: aneuploidy and the mitotic checkpoint. Nat Rev Cancer 5**:** 773-785.

Krenn, V., and A. Musacchio, 2015 The Aurora B Kinase in Chromosome Bi-Orientation and Spindle Checkpoint Signaling. Front Oncol 5**:** 225.

Krzywinski, M., J. Schein, I. Birol, J. Connors, R. Gascoyne *et al.*, 2009 Circos: an information aesthetic for comparative genomics. Genome Res 19**:** 1639-1645.

Lacefield, S., D. T. Lau and A. W. Murray, 2009 Recruiting a microtubule-binding complex to DNA directs chromosome segregation in budding yeast. Nat Cell Biol 11**:** 1116-1120.

Lachke, S. A., S. R. Lockhart, K. J. Daniels and D. R. Soll, 2003 Skin facilitates Candida albicans mating. Infect Immun 71**:** 4970-4976.

Lachner, M., D. O'Carroll, S. Rea, K. Mechtler and T. Jenuwein, 2001 Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410**:** 116-120.

Ladurner, R., and A. F. Straight, 2016 MIS12/MIND Control at the Kinetochore. Cell 167**:** 889-891.

Lai, X., L. Cao, H. Tan, S. Fang, Y. Huang *et al.*, 2007 Fungal communities from methane hydrate-bearing deep-sea marine sediments in South China Sea. ISME J 1**:** 756-762.

Lampert, F., P. Hornung and S. Westermann, 2010 The Dam1 complex confers microtubule plus end-tracking activity to the Ndc80 kinetochore complex. J Cell Biol 189**:** 641-649.

Lampson, M. A., and I. M. Cheeseman, 2011 Sensing centromere tension: Aurora B and the regulation of kinetochore function. Trends Cell Biol 21**:** 133-140.

Lang, J., A. Barber and S. Biggins, 2018 An assay for de novo kinetochore assembly reveals a key role for the CENP-T pathway in budding yeast. Elife 7.

Le Calvez, T., G. Burgaud, S. Mahe, G. Barbier and P. Vandenkoornhuyse, 2009 Fungal diversity in deep-sea hydrothermal ecosystems. Appl Environ Microbiol 75**:** 6415-6421.

Lechner, J., and J. Carbon, 1991 A 240 kd multisubunit protein complex, CBF3, is a major component of the budding yeast centromere. Cell 64**:** 717-725.

Lee, H. R., W. Zhang, T. Langdon, W. Jin, H. Yan *et al.*, 2005 Chromatin immunoprecipitation cloning reveals rapid evolutionary patterns of centromeric DNA in Oryza species. Proc Natl Acad Sci U S A 102**:** 11793-11798.

Legal, T., J. Zou, A. Sochaj, J. Rappsilber and J. P. Welburn, 2016 Molecular architecture of the Dam1 complex-microtubule interaction. Open Biol 6.

Lephart, P. R., and P. T. Magee, 2006 Effect of the major repeat sequence on mitotic recombination in Candida albicans. Genetics 174**:** 1737-1744.

Letessier, A., G. A. Millot, S. Koundrioukoff, A. M. Lachages, N. Vogt *et al.*, 2011 Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. Nature 470**:** 120-123.

Letunic, I., and P. Bork, 2019 Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47**:** W256-W259.

Li, F., L. Sonbuchner, S. A. Kyes, C. Epp and K. W. Deitsch, 2008 Nuclear non-coding RNAs are transcribed from the centromeres of Plasmodium falciparum and are associated with centromeric chromatin. J Biol Chem 283**:** 5692-5698.

Li, R., and A. W. Murray, 1991 Feedback control of mitosis in budding yeast. Cell 66**:** 519-531.

Li, X., and R. B. Nicklas, 1995 Mitotic forces control a cell-cycle checkpoint. Nature 373**:** 630-632.

Li, Y., J. Bachant, A. A. Alcasabas, Y. Wang, J. Qin *et al.*, 2002 The mitotic spindle is required for loading of the DASH complex onto the kinetochore. Genes Dev 16**:** 183-197.

Li, Y., W. Yu, Y. Liang and X. Zhu, 2007 Kinetochore dynein generates a poleward pulling force to facilitate congression and full chromosome alignment. Cell Res 17**:** 701-712.

Lima-de-Faria, A., 1949 Genetics, origin and evolution of kinetochores. Hereditas 35**:** 422-444.

Limon, J. J., J. Tang, D. Li, A. J. Wolf, K. S. Michelsen *et al.*, 2019 *Malassezia* is associated with crohn's disease and exacerbates colitis in mouse models. Cell Host Microbe 25**:** 377-388 e376.

Ling, Y. H., and K. W. Y. Yuen, 2019 Point centromere activity requires an optimal level of centromeric noncoding RNA. Proc Natl Acad Sci U S A 116**:** 6270-6279.

Liu, D., M. Vleugel, C. B. Backer, T. Hori, T. Fukagawa *et al.*, 2010 Regulated targeting of protein phosphatase 1 to the outer kinetochore by KNL1 opposes Aurora B kinase. J Cell Biol 188**:** 809-820.

Liu, S. T., J. B. Rattner, S. A. Jablonski and T. J. Yen, 2006 Mapping the assembly pathways that specify formation of the trilaminar kinetochore plates in human cells. J Cell Biol 175**:** 41-53.

Liu, X., I. McLeod, S. Anderson, J. R. Yates, 3rd and X. He, 2005 Molecular analysis of kinetochore architecture in fission yeast. EMBO J 24**:** 2919-2930.

Logsdon, G. A., E. J. Barrey, E. A. Bassett, J. E. DeNizio, L. Y. Guo *et al.*, 2015 Both tails and the centromere targeting domain of CENP-A are required for centromere establishment. J Cell Biol 208**:** 521-531.

London, N., and S. Biggins, 2014 Signalling dynamics in the spindle checkpoint response. Nat Rev Mol Cell Biol 15**:** 736-747.

London, N., S. Ceto, J. A. Ranish and S. Biggins, 2012 Phosphoregulation of Spc105 by Mps1 and PP1 regulates Bub1 localization to kinetochores. Curr Biol 22**:** 900-906.

Lorch, J. M., J. M. Palmer, K. J. Vanderwolf, K. Z. Schmidt, M. L. Verant *et al.*, 2018 *Malassezia vespertilionis* sp. nov.: a new cold-tolerant species of yeast isolated from bats. Persoonia-Molecular Phylogeny and Evolution of Fungi 41**:** 56-70.

Louis, E. J., E. S. Naumova, A. Lee, G. Naumov and J. E. Haber, 1994 The chromosome end in yeast: its mosaic nature and influence on recombinational dynamics. Genetics 136**:** 789-802.

Luo, M. C., K. R. Deal, E. D. Akhunov, A. R. Akhunova, O. D. Anderson *et al.*, 2009 Genome comparisons reveal a dominant mechanism of chromosome number reduction in grasses and accelerated genome evolution in *Triticeae*. Proc Natl Acad Sci U S A 106**:** 15780-15785.

Lynch, M., and J. S. Conery, 2000 The evolutionary fate and consequences of duplicate genes. Science 290**:** 1151-1155.

Lysak, M. A., A. Berr, A. Pecinka, R. Schmidt, K. McBreen *et al.*, 2006 Mechanisms of chromosome number reduction in Arabidopsis thaliana and related Brassicaceae species. Proc Natl Acad Sci U S A 103**:** 5224-5229.

Ma, L. J., A. S. Ibrahim, C. Skory, M. G. Grabherr, G. Burger *et al.*, 2009 Genomic analysis of the basal lineage fungus *Rhizopus oryzae* reveals a whole-genome duplication. PLoS Genet 5**:** e1000549.

Magee, B. B., and P. T. Magee, 2000 Induction of mating in Candida albicans by construction of MTLa and MTLalpha strains. Science 289**:** 310-313.

Magidson, V., C. B. O'Connell, J. Loncarek, R. Paul, A. Mogilner *et al.*, 2011 The spatial arrangement of chromosomes during prometaphase facilitates spindle assembly. Cell 146**:** 555-567.

Malik, H. S., and S. Henikoff, 2001 Adaptive evolution of Cid, a centromere-specific histone in Drosophila. Genetics 157**:** 1293-1298.

Malik, H. S., and S. Henikoff, 2009 Major evolutionary transitions in centromere complexity. Cell 138**:** 1067-1082.

Maloney, K. A., L. L. Sullivan, J. E. Matheny, E. D. Strome, S. L. Merrett *et al.*, 2012 Functional epialleles at an endogenous human centromere. Proc Natl Acad Sci U S A 109**:** 13704-13709.

Malvezzi, F., G. Litos, A. Schleiffer, A. Heuck, K. Mechtler *et al.*, 2013 A structural basis for kinetochore recruitment of the Ndc80 complex via two distinct centromere receptors. EMBO J 32**:** 409-423.

Maskell, D. P., X. W. Hu and M. R. Singleton, 2010 Molecular architecture and assembly of the yeast kinetochore MIND complex. J Cell Biol 190**:** 823-834.

Massey, S. E., G. Moura, P. Beltrao, R. Almeida, J. R. Garey *et al.*, 2003 Comparative evolutionary genomics unveils the molecular mechanism of reassignment of the CTG codon in Candida spp. Genome Res 13**:** 544-557.

Masumoto, H., H. Masukata, Y. Muro, N. Nozaki and T. Okazaki, 1989 A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. J Cell Biol 109**:** 1963-1973.

Matsumoto, T., S. Murakami, O. Niwa and M. Yanagida, 1990 Construction and characterization of centric circular and acentric linear chromosomes in fission yeast. Current Genetics 18**:** 323-330.

Mattos, V. F., L. S. Carvalho, M. A. Carvalho and M. C. Schneider, 2018 Insights into the origin of the high variability of multivalent-meiotic associations in holocentric chromosomes of Tityus (Archaeotityus) scorpions. PLoS One 13**:** e0192070.

Maure, J. F., S. Komoto, Y. Oku, A. Mino, S. Pasqualato *et al.*, 2011 The Ndc80 loop region facilitates formation of kinetochore attachment to the dynamic microtubule plus end. Curr Biol 21**:** 207-213.

McElroy, K. E., R. D. Denton, J. Sharbrough, L. Bankers, M. Neiman *et al.*, 2017 Genome Expression Balance in a Triploid Trihybrid Vertebrate. Genome Biol Evol 9**:** 968-980.

McEwen, B. F., G. K. Chan, B. Zubrowski, M. S. Savoian, M. T. Sauer *et al.*, 2001 CENP-E is essential for reliable bioriented spindle attachment, but chromosome alignment can be achieved via redundant mechanisms in mammalian cells. Mol Biol Cell 12**:** 2776-2789.

McGrew, J., B. Diehl and M. Fitzgerald-Hayes, 1986 Single base-pair mutations in centromere element III cause aberrant chromosome segregation in Saccharomyces cerevisiae. Mol Cell Biol 6**:** 530-538.

McIntosh, J. R., E. L. Grishchuk, M. K. Morphew, A. K. Efremov, K. Zhudenkov *et al.*, 2008 Fibrils connect microtubule tips with kinetochores: a mechanism to couple tubulin dynamics to chromosome motion. Cell 135**:** 322-333.

McKinley, K. L., and I. M. Cheeseman, 2016 The molecular basis for centromere identity and function. Nat Rev Mol Cell Biol 17**:** 16-29.

McKinley, K. L., N. Sekulic, L. Y. Guo, T. Tsinman, B. E. Black *et al.*, 2015 The CENP-L-N Complex Forms a Critical Node in an Integrated Meshwork of Interactions at the Centromere-Kinetochore Interface. Mol Cell 60**:** 886-898.

McLysaght, A., K. Hokamp and K. H. Wolfe, 2002 Extensive genomic duplication during early chordate evolution. Nat Genet 31**:** 200-204.

Measday, V., D. W. Hailey, I. Pot, S. A. Givan, K. M. Hyland *et al.*, 2002 Ctf3p, the Mis6 budding yeast homolog, interacts with Mcm22p and Mcm16p at the yeast outer kinetochore. Genes Dev 16**:** 101-113.

Mellone, B. G., W. Zhang and G. H. Karpen, 2009 Frodos found: Behold the CENP-a "Ring" bearers. Cell 137**:** 409-412.

Melters, D. P., K. R. Bradnam, H. A. Young, N. Telis, M. R. May *et al.*, 2013 Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. Genome Biol 14**:** R10.

Melters, D. P., L. V. Paliulis, I. F. Korf and S. W. Chan, 2012 Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. Chromosome Res 20**:** 579-593.

Meluh, P. B., and D. Koshland, 1995 Evidence that the MIF2 gene of Saccharomyces cerevisiae encodes a centromere protein with homology to the mammalian centromere protein CENP-C. Mol Biol Cell 6**:** 793-807.

Meluh, P. B., and D. Koshland, 1997 Budding yeast centromere composition and assembly as revealed by in vivo cross-linking. Genes Dev 11**:** 3401-3412.

Mendiburo, M. J., J. Padeken, S. Fulop, A. Schepers and P. Heun, 2011 Drosophila CENH3 is sufficient for centromere formation. Science 334**:** 686-690.

Meraldi, P., A. D. McAinsh, E. Rheinbay and P. K. Sorger, 2006 Phylogenetic and structural analysis of centromeric DNA and kinetochore proteins. Genome Biol 7**:** R23.

Merson-Davies, L. A., and F. C. Odds, 1989 A morphology index for characterization of cell shape in Candida albicans. J Gen Microbiol 135**:** 3143-3152.

Miller, M. G., and A. D. Johnson, 2002 White-opaque switching in Candida albicans is controlled by mating-type locus homeodomain proteins and allows efficient mating. Cell 110**:** 293-302.

Miller, S. A., M. L. Johnson and P. T. Stukenberg, 2008 Kinetochore attachments require an interaction between unstructured tails on microtubules and Ndc80(Hec1). Curr Biol 18**:** 1785-1791.

Miranda, J. J., P. De Wulf, P. K. Sorger and S. C. Harrison, 2005 The yeast DASH complex forms closed rings on microtubules. Nat Struct Mol Biol 12**:** 138-143.

Mirchenko, L., and F. Uhlmann, 2010 Sli15(INCENP) dephosphorylation prevents mitotic checkpoint reengagement due to loss of tension at anaphase onset. Curr Biol 20**:** 1396-1401.

Mitra, S., J. Gomez-Raja, G. Larriba, D. D. Dubey and K. Sanyal, 2014 Rad51-Rad52 mediated maintenance of centromeric chromatin in *Candida albicans*. PLoS Genet 10**:** e1004344.

Moore, L. L., and M. B. Roth, 2001 HCP-4, a CENP-C-like protein in Caenorhabditis elegans, is required for resolution of sister centromeres. J Cell Biol 153**:** 1199-1208.

Morand, S. C., M. Bertignac, A. Iltis, I. Kolder, W. Pirovano *et al.*, 2019 Complete Genome Sequence of Malassezia restricta CBS 7877, an Opportunist Pathogen Involved in Dandruff and Seborrheic Dermatitis. Microbiol Resour Announc 8.

Moroi, Y., A. L. Hartman, P. K. Nakane and E. M. Tan, 1981 Distribution of kinetochore (centromere) antigen in mammalian cell nuclei. J Cell Biol 90**:** 254-259.

Moroi, Y., C. Peebles, M. J. Fritzler, J. Steigerwald and E. M. Tan, 1980 Autoantibody to centromere (kinetochore) in scleroderma sera. Proc Natl Acad Sci U S A 77**:** 1627-1631.

Motamedi, M. R., A. Verdel, S. U. Colmenares, S. A. Gerber, S. P. Gygi *et al.*, 2004 Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell 119**:** 789-802.

Murat, F., J. H. Xu, E. Tannier, M. Abrouk, N. Guilhot *et al.*, 2010 Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. Genome Res 20**:** 1545-1557.

Muro, Y., H. Masumoto, K. Yoda, N. Nozaki, M. Ohashi *et al.*, 1992 Centromere protein B assembles human centromeric alpha-satellite DNA at the 17-bp sequence, CENP-B box. J Cell Biol 116**:** 585-596.

Musacchio, A., and A. Desai, 2017 A molecular view of kinetochore assembly and function. Biology (Basel) 6**:** 5-54.

Mythreye, K., and K. S. Bloom, 2003 Differential kinetochore protein requirements for establishment versus propagation of centromere activity in Saccharomyces cerevisiae. J Cell Biol 160**:** 833-843.

Nagaki, K., Z. Cheng, S. Ouyang, P. B. Talbert, M. Kim *et al.*, 2004 Sequencing of a rice centromere uncovers active genes. Nat Genet 36**:** 138-145.

Nagaki, K., P. B. Talbert, C. X. Zhong, R. K. Dawe, S. Henikoff *et al.*, 2003 Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of Arabidopsis thaliana centromeres. Genetics 163**:** 1221-1225.

Nagata, R., H. Nagano, D. Ogishima, Y. Nakamura, M. Hiruma *et al.*, 2012 Transmission of the major skin microbiota, Malassezia, from mother to neonate. Pediatr Int 54**:** 350-355.

Nagpal, H., and T. Fukagawa, 2016 Kinetochore assembly and function through the cell cycle. Chromosoma 125**:** 645-659.

Nakano, M., S. Cardinale, V. N. Noskov, R. Gassmann, P. Vagnarelli *et al.*, 2008 Inactivation of a human kinetochore by specific targeting of chromatin modifiers. Dev Cell 14**:** 507-522.

Nakayashiki, H., N. Kadotani and S. Mayama, 2006 Evolution and diversification of RNA silencing proteins in fungi. J Mol Evol 63**:** 127-135.

Naranjo-Ortiz, M. A., and T. Gabaldon, 2019 Fungal evolution: diversity, taxonomy and phylogeny of the Fungi. Biol Rev Camb Philos Soc 94**:** 2101-2137.

Navarro-Mendoza, M. I., C. Perez-Arques, S. Panchal, F. E. Nicolas, S. J. Mondo *et al.*, 2019 Early diverging fungus *Mucor circinelloides* lacks centromeric histone CENP-A and displays a mosaic of point and regional centromeres. *Curr Biol* 22**:** 3791-3802.

Nekrasov, V. S., M. A. Smith, S. Peak-Chew and J. V. Kilmartin, 2003 Interactions between centromere complexes in Saccharomyces cerevisiae. Mol Biol Cell 14**:** 4931-4946.

Ness, F., and M. Aigle, 1995 RTM1: a member of a new family of telomeric repeated genes in yeast. Genetics 140**:** 945-956.

Newmeyer, D., and D. R. Galeazzi, 1977 The Instability of Neurospora Duplication Dp(IL-->IR)H4250 , and Its Genetic Control. Genetics 85**:** 461-487.

Ng, C. T., L. Deng, C. Chen, H. H. Lim, J. Shi *et al.*, 2019 Electron cryotomography analysis of Dam1C/DASH at the kinetochore-spindle interface in situ. J Cell Biol 218**:** 455-473.

Ng, R., and J. Carbon, 1987 Mutational and in vitro protein-binding studies on centromere DNA from Saccharomyces cerevisiae. Mol Cell Biol 7**:** 4522-4534.

Nguyen, L. T., H. A. Schmidt, A. von Haeseler and B. Q. Minh, 2015 IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32**:** 268-274.

Ni, M., M. Feretzaki, S. Sun, X. Wang and J. Heitman, 2011 Sex in fungi. Annu Rev Genet 45**:** 405-430.

Niikura, Y., R. Kitagawa, H. Ogi, R. Abdulle, V. Pagala *et al.*, 2015 CENP-A K124 Ubiquitylation Is Required for CENP-A Deposition at the Centromere. Dev Cell 32**:** 589-603.

Nishihashi, A., T. Haraguchi, Y. Hiraoka, T. Ikemura, V. Regnier *et al.*, 2002 CENP-I is essential for centromere function in vertebrate cells. Dev Cell 2**:** 463-476.

Nishino, T., K. Takeuchi, K. E. Gascoigne, A. Suzuki, T. Hori *et al.*, 2012 CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. Cell 148**:** 487-501.

Nobile, C. J., and A. D. Johnson, 2015 Candida albicans Biofilms and Human Disease. Annu Rev Microbiol 69**:** 71-92.

Noor, M. A., K. L. Grams, L. A. Bertucci and J. Reiland, 2001 Chromosomal inversions and the reproductive isolation of species. Proc Natl Acad Sci U S A 98**:** 12084-12088.

Oegema, K., A. Desai, S. Rybina, M. Kirkham and A. A. Hyman, 2001 Functional analysis of kinetochore assembly in Caenorhabditis elegans. J Cell Biol 153**:** 1209-1226.

Oh, J., A. L. Byrd, C. Deming, S. Conlan, N. C. S. Program *et al.*, 2014 Biogeography and individuality shape function in the human skin metagenome. Nature 514**:** 59-64.

Ohkuni, K., and K. Kitagawa, 2011 Endogenous transcription at the centromere facilitates centromere activity in budding yeast. Curr Biol 21**:** 1695-1703.

Ohkuni, K., and K. Kitagawa, 2012 Role of transcription at centromeres in budding yeast. Transcription 3**:** 193-197.

Ohno, S., 1970 *Evolution by gene duplication*. Springer Science & Business Media.

Ohzeki, J., J. H. Bergmann, N. Kouprina, V. N. Noskov, M. Nakano *et al.*, 2012 Breaking the HAC Barrier: histone H3K9 acetyl/methyl balance regulates CENP-A assembly. EMBO J 31**:** 2391-2402.

Ohzeki, J., M. Nakano, T. Okada and H. Masumoto, 2002 CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. J Cell Biol 159**:** 765-775.

Okada, M., K. Okawa, T. Isobe and T. Fukagawa, 2009 CENP-H-containing complex facilitates centromere deposition of CENP-A in cooperation with FACT and CHD1. Mol Biol Cell 20**:** 3986-3995.

Ortiz, J., O. Stemmann, S. Rank and J. Lechner, 1999 A putative protein complex consisting of Ctf19, Mcm21, and Okp1 represents a missing link in the budding yeast kinetochore. Genes Dev 13**:** 1140-1155.

Otto, S. P., and J. Whitton, 2000 Polyploid incidence and evolution. Annu Rev Genet 34**:** 401-437.

Ozeri-Galai, E., R. Lebofsky, A. Rahat, A. C. Bester, A. Bensimon *et al.*, 2011 Failure of origin activation in response to fork stalling leads to chromosomal instability at fragile sites. Mol Cell 43**:** 122-131.

Padmanabhan, S., J. Thakur, R. Siddharthan and K. Sanyal, 2008 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 105**:** 19797-19802.

Pagliuca, C., V. M. Draviam, E. Marco, P. K. Sorger and P. De Wulf, 2009 Roles for the conserved spc105p/kre28p complex in kinetochore-microtubule binding and the spindle assembly checkpoint. PLoS One 4**:** e7640.

Palmer, D. K., K. O'Day, H. L. Trong, H. Charbonneau and R. L. Margolis, 1991 Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proc Natl Acad Sci U S A 88**:** 3734-3738.

Palmer, D. K., K. O'Day, M. H. Wener, B. S. Andrews and R. L. Margolis, 1987 A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. J Cell Biol 104**:** 805-815.

Pande, K., C. Chen and S. M. Noble, 2013 Passage through the mammalian gut triggers a phenotypic switch that promotes Candida albicans commensalism. Nat Genet 45**:** 1088-1091.

Panja, G., 1927 The *Malassezia* of the Skin, their Cultivation, Morphology and Species, pp. in *Far Eastern Assoc. Trop. Med. Trans. Seventh Congress, British India, 1927.*

Pekgoz Altunkaya, G., F. Malvezzi, Z. Demianova, T. Zimniak, G. Litos *et al.*, 2016 CCAN Assembly Configures Composite Binding Interfaces to Promote Cross-Linking of Ndc80 Complexes at the Kinetochore. Curr Biol 26**:** 2370-2378.

Perea-Resa, C., and M. D. Blower, 2018 Centromere Biology: Transcription Goes on Stage. Mol Cell Biol 38.

Perpelescu, M., and T. Fukagawa, 2011 The ABCs of CENPs. Chromosoma 120**:** 425-446.

Pesenti, M. E., J. R. Weir and A. Musacchio, 2016 Progress in the structural and functional characterization of kinetochores. Curr Opin Struct Biol 37**:** 152-163.

Petrovic, A., J. Keller, Y. Liu, K. Overlack, J. John *et al.*, 2016 Structure of the MIS12 Complex and Molecular Basis of Its Interaction with CENP-C at Human Kinetochores. Cell 167**:** 1028-1040 e1015.

Pidoux, A. L., E. S. Choi, J. K. Abbott, X. Liu, A. Kagansky *et al.*, 2009 Fission yeast Scm3: A CENP-A receptor required for integrity of subkinetochore chromatin. Mol Cell 33**:** 299-311.

Pidoux, A. L., W. Richardson and R. C. Allshire, 2003 Sim4: a novel fission yeast kinetochore protein required for centromeric silencing and chromosome segregation. J Cell Biol 161**:** 295-307.

Pierard, J., and P. Dockx, 1972 The ultrastructure of tinea versicolor and Malassezia furfur. Int J Dermatol 11**:** 116-124.

Pinsky, B. A., C. Kung, K. M. Shokat and S. Biggins, 2006 The Ipl1-Aurora protein kinase activates the spindle checkpoint by creating unattached kinetochores. Nat Cell Biol 8**:** 78-83.

Pinsky, B. A., C. R. Nelson and S. Biggins, 2009 Protein phosphatase 1 regulates exit from the spindle checkpoint in budding yeast. Curr Biol 19**:** 1182-1187.

Pinsky, B. A., S. Y. Tatsutani, K. A. Collins and S. Biggins, 2003 An Mtw1 complex promotes kinetochore biorientation that is monitored by the Ipl1/Aurora protein kinase. Dev Cell 5**:** 735-745.

Plohl, M., N. Mestrovic and B. Mravinac, 2014 Centromere identity from the DNA point of view. Chromosoma 123**:** 313-325.

Porro, M. N., S. Passi, F. Caprilli and R. Mercantini, 1977 Induction of hyphae in cultures of Pityrosporum by cholesterol and cholesterol esters. J Invest Dermatol 69**:** 531-534.

Pot, I., V. Measday, B. Snydsman, G. Cagney, S. Fields *et al.*, 2003 Chl4p and iml3p are two new members of the budding yeast outer kinetochore. Mol Biol Cell 14**:** 460-476.

Powers, A. F., A. D. Franck, D. R. Gestaut, J. Cooper, B. Gracyzk *et al.*, 2009 The Ndc80 kinetochore complex forms load-bearing attachments to dynamic microtubule tips via biased diffusion. Cell 136**:** 865-875.

Prohic, A., T. Jovovic Sadikovic, M. Krupalija-Fazlic and S. Kuskunovic-Vlahovljak, 2016 Malassezia species in healthy skin and in dermatological conditions. Int J Dermatol 55**:** 494-504.

Quenet, D., and Y. Dalal, 2014 A long non-coding RNA is required for targeting centromeric protein A to the human centromere. Elife 3**:** e03254.

Rago, F., K. E. Gascoigne and I. M. Cheeseman, 2015 Distinct organization and regulation of the outer kinetochore KMN network downstream of CENP-C and CENP-T. Curr Biol 25**:** 671-677.

Ranson, H., C. Claudianos, F. Ortelli, C. Abgrall, J. Hemingway *et al.*, 2002 Evolution of supergene families associated with insecticide resistance. Science 298**:** 179-181.

Ravi, M., and S. W. Chan, 2010 Haploid plants produced by centromere-mediated genome elimination. Nature 464**:** 615-618.

Reinhart, B. J., and D. P. Bartel, 2002 Small RNAs correspond to centromere heterochromatic repeats. Science 297**:** 1831.

Rhind, N., Z. Chen, M. Yassour, D. A. Thompson, B. J. Haas *et al.*, 2011 Comparative functional genomics of the fission yeasts. Science 332**:** 930-936.

Ribeiro, S. A., P. Vagnarelli, Y. Dong, T. Hori, B. F. McEwen *et al.*, 2010 A super-resolution map of the vertebrate kinetochore. Proc Natl Acad Sci U S A 107**:** 10484-10489.

Rieder, C. L., 1982 The formation, structure, and composition of the mammalian kinetochore and kinetochore fiber. Int Rev Cytol 79**:** 1-58.

Rief, N., and M. Lobrich, 2002 Efficient rejoining of radiation-induced DNA double-strand breaks in centromeric DNA of human cells. J Biol Chem 277**:** 20572-20582.

Rieseberg, L. H., 2001 Chromosomal rearrangements and speciation. Trends Ecol Evol 16**:** 351-358.

Roach, K. C., B. D. Ross and H. S. Malik, 2012 Rapid evolution of centromeres and centromeric/kinetochore proteins. Rapidly Evolving Genes and Genetic Systems; Singh, RS, Xu, J., Kulathinal, RJ, Eds**:** 83-93.

Rocchi, M., R. Stanyon and N. Archidiacono, 2009 Evolutionary new centromeres in primates. Prog Mol Subcell Biol 48**:** 103-152.

Rosic, S., F. Kohler and S. Erhardt, 2014 Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. J Cell Biol 207**:** 335-349.

Roy, B., L. S. Burrack, M. A. Lone, J. Berman and K. Sanyal, 2011 CaMtw1, a member of the evolutionarily conserved Mis12 kinetochore protein family, is required for efficient inner kinetochore assembly in the pathogenic yeast *Candida albicans*. Mol Microbiol 80**:** 14-32.

Roy, B., and K. Sanyal, 2011 Diversity in requirement of genetic and epigenetic factors for centromere function in fungi. Eukaryot Cell 10**:** 1384-1395.

Rudd, M. K., M. G. Schueler and H. F. Willard, 2003 Sequence organization and functional annotation of human centromeres. Cold Spring Harb Symp Quant Biol 68**:** 141-149.

Sabouraud, R. J. A., 1902 *Maladies du cuir chevelu*. Masson.

Sadeghi, L., L. Siggens, J. P. Svensson and K. Ekwall, 2014 Centromeric histone H2B monoubiquitination promotes noncoding transcription and chromatin integrity. Nat Struct Mol Biol 21**:** 236-243.

Saitoh, H., J. Tomkiel, C. A. Cooke, H. Ratrie, 3rd, M. Maurer *et al.*, 1992 CENP-C, an autoantigen in scleroderma, is a component of the human inner kinetochore plate. Cell 70**:** 115-125.

Saitoh, S., K. Takahashi and M. Yanagida, 1997 Mis6, a fission yeast inner centromere protein, acts during G1/S and forms specialized chromatin required for equal segregation. Cell 90**:** 131-143.

Samejima, I., C. Spanos, L. Alves Fde, T. Hori, M. Perpelescu *et al.*, 2015 Whole-proteome genetic analysis of dependencies in assembly of a vertebrate kinetochore. J Cell Biol 211**:** 1141-1156.

Sandall, S., and A. Desai, 2007 When it comes to couple(r)s, do opposites attract? Nature Structural & Molecular Biology 14**:** 790-792.

Sanei, M., R. Pickering, K. Kumke, S. Nasuda and A. Houben, 2011 Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proc Natl Acad Sci U S A 108**:** 498-505.

Santaguida, S., and A. Musacchio, 2009 The life and miracles of kinetochores. EMBO J 28**:** 2511-2531.

Sanyal, K., M. Baum and J. Carbon, 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.

Sanyal, K., and J. Carbon, 2002 The CENP-A homolog CaCse4p in the pathogenic yeast Candida albicans is a centromere protein essential for chromosome transmission. Proc Natl Acad Sci U S A 99**:** 12969-12974.

Sato, H., F. Masuda, Y. Takayama, K. Takahashi and S. Saitoh, 2012 Epigenetic inactivation and subsequent heterochromatinization of a centromere stabilize dicentric chromosomes. Curr Biol 22**:** 658-667.

Schleiffer, A., M. Maier, G. Litos, F. Lampert, P. Hornung *et al.*, 2012 CENP-T proteins are conserved centromere receptors of the Ndc80 complex. Nat Cell Biol 14**:** 604-613.

Schmidt, J. C., H. Arthanari, A. Boeszoermenyi, N. M. Dashkevich, E. M. Wilson-Kubalek *et al.*, 2012 The kinetochore-bound Ska1 complex tracks depolymerizing microtubules and binds to curved protofilaments. Dev Cell 23**:** 968-980.

Schotanus, K., and J. Heitman, 2019 Centromere deletion in *Cryptococcus deuterogattii* leads to neocentromere formation and chromosome fusions. bioRxiv**:** 526962.

Schubert, I., 2018 What is behind "centromere repositioning"? Chromosoma 127**:** 229-234.

Schueler, M. G., W. Swanson, P. J. Thomas, N. C. S. Program and E. D. Green, 2010 Adaptive evolution of foundation kinetochore proteins in primates. Mol Biol Evol 27**:** 1585-1597.

Schwartz, M., E. Zlotorynski, M. Goldberg, E. Ozeri, A. Rahat *et al.*, 2005 Homologous recombination and nonhomologous end-joining repair pathways regulate fragile site stability. Genes Dev 19**:** 2715-2726.

Screpanti, E., A. De Antoni, G. M. Alushin, A. Petrovic, T. Melis *et al.*, 2011 Direct binding of Cenp-C to the Mis12 complex joins the inner and outer kinetochore. Curr Biol 21**:** 391-398.

Seidl, M. F., and B. P. Thomma, 2014 Sex or no sex: evolutionary adaptation occurs regardless. Bioessays 36**:** 335-345.

Seidl, M. F., G. Van den Ackerveken, F. Govers and B. Snel, 2012 Reconstruction of oomycete genome evolution identifies differences in evolutionary trajectories leading to present-day large gene families. Genome Biol Evol 4**:** 199-211.

Selmecki, A., A. Forche and J. Berman, 2006 Aneuploidy and isochromosome formation in drug-resistant Candida albicans. Science 313**:** 367-370.

Selmecki, A., A. Forche and J. Berman, 2010 Genomic plasticity of the human fungal pathogen Candida albicans. Eukaryot Cell 9**:** 991-1008.

Selmecki, A., M. Gerami-Nejad, C. Paulson, A. Forche and J. Berman, 2008 An isochromosome confers drug resistance in vivo by amplification of two genes, ERG11 and TAC1. Mol Microbiol 68**:** 624-641.

Selmecki, A. M., K. Dulmage, L. E. Cowen, J. B. Anderson and J. Berman, 2009 Acquisition of aneuploidy provides increased fitness during the evolution of antifungal drug resistance. PLoS Genet 5**:** e1000705.

Senczek, D., U. Siesenop and K. H. Bohm, 1999 Characterization of *Malassezia* species by means of phenotypic characteristics and detection of electrophoretic karyotypes by pulsed-field gel electrophoresis (PFGE). Mycoses 42**:** 409-414.

Seoighe, C., N. Federspiel, T. Jones, N. Hansen, V. Bivolarovic *et al.*, 2000 Prevalence of small inversions in yeast gene order evolution. Proc Natl Acad Sci U S A 97**:** 14433-14437.

Session, A. M., Y. Uno, T. Kwon, J. A. Chapman, A. Toyoda *et al.*, 2016 Genome evolution in the allotetraploid frog Xenopus laevis. Nature 538**:** 336-343.

Sexton, C. E., and J. A. Roper, 1984 Spontaneous duplications and transpositions of a large chromosome segment in Aspergillus nidulans. J Gen Microbiol 130**:** 583-595.

Shah, H., K. Rawat, H. Ashar, R. Patkar and J. Manjrekar, 2019 Dual role for fungal-specific outer kinetochore proteins during cell cycle and development in Magnaporthe oryzae. Journal of Cell Science 132**:** jcs224147.

Shang, W. H., T. Hori, N. M. Martins, A. Toyoda, S. Misu *et al.*, 2013 Chromosome engineering allows the efficient isolation of vertebrate neocentromeres. Dev Cell 24**:** 635-648.

Shang, W. H., T. Hori, F. G. Westhorpe, K. M. Godek, A. Toyoda *et al.*, 2016 Acetylation of histone H4 lysine 5 and 12 is required for CENP-A deposition into centromeres. Nat Commun 7**:** 13465.

Sharakhov, I. V., A. C. Serazin, O. G. Grushko, A. Dana, N. Lobo *et al.*, 2002 Inversions and gene order shuffling in Anopheles gambiae and A. funestus. Science 298**:** 182-185.

Shrestha, R. L., D. Conti, N. Tamura, D. Braun, R. A. Ramalingam *et al.*, 2017 Aurora-B kinase pathway controls the lateral to end-on conversion of kinetochore-microtubule attachments in human cells. Nat Commun 8**:** 150.

Shrestha, R. L., and V. M. Draviam, 2013 Lateral to end-on conversion of chromosome-microtubule attachment requires kinesins CENP-E and MCAK. Curr Biol 23**:** 1514-1526.

Sichova, J., A. Volenikova, V. Dinca, P. Nguyen, R. Vila *et al.*, 2015 Dynamic karyotype evolution and unique sex determination systems in Leptidea wood white butterflies. BMC Evol Biol 15**:** 89.

Siddharthan, R., 2008 PhyloGibbs-MP: module prediction and discriminative motif-finding by Gibbs sampling. PLoS Comput Biol 4**:** e1000156.

Siddharthan, R., E. D. Siggia and E. van Nimwegen, 2005 PhyloGibbs: a Gibbs sampling motif finder that incorporates phylogeny. PLoS Comput Biol 1**:** e67.

Simmons, R. B., and E. Gueho, 1990 A new species of Malassezia. Mycological research 94**:** 1146-1149.

Sionov, E., H. Lee, Y. C. Chang and K. J. Kwon-Chung, 2010 Cryptococcus neoformans overcomes stress of azole drugs by formation of disomy in specific multiple chromosomes. PLoS Pathog 6**:** e1000848.

Smith, J. G., M. S. Caddle, G. H. Bulboaca, J. G. Wohlgemuth, M. Baum *et al.*, 1995 Replication of centromere II of *Schizosaccharomyces pombe*. Mol Cell Biol 15**:** 5165-5172.

Smith, K. M., J. M. Galazka, P. A. Phatale, L. R. Connolly and M. Freitag, 2012 Centromeres of filamentous fungi. Chromosome Res 20**:** 635-656.

Smith, K. M., P. A. Phatale, C. M. Sullivan, K. R. Pomraning and M. Freitag, 2011 Heterochromatin is required for normal distribution of Neurospora crassa CenH3. Mol Cell Biol 31**:** 2528-2542.

Soll, D. R., 2002 Candida commensalism and virulence: the evolution of phenotypic plasticity. Acta Trop 81**:** 101-110.

Spanu, P. D., J. C. Abbott, J. Amselem, T. A. Burgis, D. M. Soanes *et al.*, 2010 Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. Science 330**:** 1543-1546.

Sparber, F., C. De Gregorio, S. Steckholzer, F. M. Ferreira, T. Dolowschiak *et al.*, 2019 The Skin Commensal Yeast Malassezia Triggers a Type 17 Response that Coordinates Anti-fungal Immunity and Exacerbates Skin Inflammation. Cell Host Microbe 25**:** 389-403 e386.

Sreekumar, L., P. Jaitly, Y. Chen, B. C. Thimmappa, A. Sanyal *et al.*, 2019a Cis- and Trans-chromosomal Interactions Define Pericentric Boundaries in the Absence of Conventional Heterochromatin. Genetics 212**:** 1121-1132.

Sreekumar, L., K. Kumari, A. Bakshi, N. Varshney, B. C. Thimmappa *et al.*, 2019b Orc4 spatiotemporally stabilizes centromeric chromatin. bioRxiv**:** 465880.

Sridhar, S., T. Hori, R. Nakagawa, T. Fukagawa and K. Sanyal, 2019 Identification of bridgin, an unconventional linker, connects the outer kinetochore to centromeric chromatin. bioRxiv**:** 816199.

Stajich, J. E., S. K. Wilke, D. Ahrén, C. H. Au, B. W. Birren *et al.*, 2010 Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom Coprinopsis cinerea (Coprinus cinereus). Proceedings of the National Academy of Sciences 107**:** 11889-11894.

Steiner, F. A., and S. Henikoff, 2014 Holocentromeres are dispersed point centromeres localized at transcription factor hotspots. Elife 3**:** e02025.

Steiner, F. A., and S. Henikoff, 2015 Diversity in the organization of centromeric chromatin. Curr Opin Genet Dev 31**:** 28-35.

Stellfox, M. E., A. O. Bailey and D. R. Foltz, 2013 Putting CENP-A in its place. Cell Mol Life Sci 70**:** 387-406.

Sternberg, T. H., and F. M. Keddie, 1961 Immunofluorescence studies in tinea versicolor. A preliminary report. Arch Dermatol 84**:** 999-1003.

Stimpson, K. M., I. Y. Song, A. Jauch, H. Holtgreve-Grez, K. E. Hayden *et al.*, 2010 Telomere disruption results in non-random formation of de novo dicentric chromosomes involving acrocentric human chromosomes. PLoS Genet 6, **:** e1001061.

Stoler, S., K. C. Keith, K. E. Curnick and M. Fitzgerald-Hayes, 1995 A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. Genes Dev 9**:** 573-586.

Stoler, S., K. Rogers, S. Weitze, L. Morey, M. Fitzgerald-Hayes *et al.*, 2007 Scm3, an essential Saccharomyces cerevisiae centromere protein required for G2/M progression and Cse4 localization. Proc Natl Acad Sci U S A 104**:** 10571-10576.

Sudbery, P., N. Gow and J. Berman, 2004 The distinct morphogenic states of Candida albicans. Trends Microbiol 12**:** 317-324.

Sugata, N., E. Munekata and K. Todokoro, 1999 Characterization of a novel kinetochore protein, CENP-H. J Biol Chem 274**:** 27343-27346.

Sugita, T., M. Tajima, M. Takashima, M. Amaya, M. Saito *et al.*, 2004 A new yeast, Malassezia yamatoensis, isolated from a patient with seborrheic dermatitis, and its distribution in patients and healthy subjects. Microbiol Immunol 48**:** 579-583.

Sugita, T., M. Takashima, M. Kodama, R. Tsuboi and A. Nishikawa, 2003a Description of a new yeast species, *Malassezia japonica*, and its detection in patients with atopic dermatitis and healthy subjects. J Clin Microbiol 41**:** 4695-4699.

Sugita, T., M. Takashima, M. Kodama, R. Tsuboi and A. Nishikawa, 2003b Description of a new yeast species, Malassezia japonica, and its detection in patients with atopic dermatitis and healthy subjects. J Clin Microbiol 41**:** 4695-4699.

Sugita, T., M. Takashima, T. Shinoda, H. Suto, T. Unno *et al.*, 2002 New yeast species, Malassezia dermatis, isolated from patients with atopic dermatitis. J Clin Microbiol 40**:** 1363-1367.

Sugiyama, T., H. Cam, A. Verdel, D. Moazed and S. I. Grewal, 2005 RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. Proc Natl Acad Sci U S A 102**:** 152-157.

Sullivan, B. A., and G. H. Karpen, 2004 Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat Struct Mol Biol 11**:** 1076-1083.

Sullivan, K. F., M. Hechenberger and K. Masri, 1994 Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. J Cell Biol 127**:** 581-592.

Sullivan, M. J., N. K. Petty and S. A. Beatson, 2011 Easyfig: a genome comparison visualizer. Bioinformatics 27**:** 1009-1010.

Sun, S., M. A. Coelho, M. David-Palma, S. J. Priest and J. Heitman, 2019 The Evolution of Sexual Reproduction and the Mating-Type Locus: Links to Pathogenesis of Cryptococcus Human Pathogenic Fungi. Annu Rev Genet 53**:** 417-444.

Sun, S., V. Yadav, R. B. Billmyre, C. A. Cuomo, M. Nowrousian *et al.*, 2017 Fungal genome and mating system transitions facilitated by chromosomal translocations involving intercentromeric recombination. PLoS Biol 15**:** e2002527.

Takahashi, K., E. S. Chen and M. Yanagida, 2000 Requirement of Mis6 centromere connector for localizing a CENP-A-like protein in fission yeast. Science 288**:** 2215-2219.

Takeuchi, K., T. Nishino, K. Mayanagi, N. Horikoshi, A. Osakabe *et al.*, 2014 The centromeric nucleosome-like CENP-T-W-S-X complex induces positive supercoils into DNA. Nucleic Acids Res 42**:** 1644-1655.

Talavera, G., V. A. Lukhtanov, L. Rieppel, N. E. Pierce and R. Vila, 2013 In the shadow of phylogenetic uncertainty: the recent diversification of Lysandra butterflies through chromosomal change. Mol Phylogenet Evol 69**:** 469-478.

Talbert, P. B., T. D. Bryson and S. Henikoff, 2004 Adaptive evolution of centromere proteins in plants and animals. J Biol 3**:** 18.

Tanaka, K., N. Mukae, H. Dewar, M. van Breugel, E. K. James *et al.*, 2005a Molecular mechanisms of kinetochore capture by spindle microtubules. Nature 434**:** 987-994.

Tanaka, M., and S. Imamura, 1979 Immunological studies on Pityrosporum genus and Malassezia furfur. J Invest Dermatol 73**:** 321-324.

Tanaka, T. U., 2010 Kinetochore-microtubule interactions: steps towards bi-orientation. EMBO J 29**:** 4070-4082.

Tanaka, T. U., N. Rachidi, C. Janke, G. Pereira, M. Galova *et al.*, 2002 Evidence that the Ipl1-Sli15 (Aurora kinase-INCENP) complex promotes chromosome bi-orientation by altering kinetochore-spindle pole connections. Cell 108**:** 317-329.

Tanaka, T. U., M. J. Stark and K. Tanaka, 2005b Kinetochore capture and bi-orientation on the mitotic spindle. Nat Rev Mol Cell Biol 6**:** 929-942.

Tao, L., H. Du, G. Guan, Y. Dai, C. J. Nobile *et al.*, 2014 Discovery of a "white-gray-opaque" tristable phenotypic switching system in candida albicans: roles of non-genetic diversity in host adaptation. PLoS Biol 12**:** e1001830.

Thakur, J., and S. Henikoff, 2016 CENPT bridges adjacent CENPA nucleosomes on young human alpha-satellite dimers. Genome Res 26**:** 1178-1187.

Thakur, J., and K. Sanyal, 2011 The essentiality of the fungus-specific Dam1 complex is correlated with a one-kinetochore-one-microtubule interaction present throughout the cell cycle, independent of the nature of a centromere. Eukaryot Cell 10**:** 1295-1305.

Thakur, J., and K. Sanyal, 2012 A coordinated interdependent protein circuitry stabilizes the kinetochore ensemble to protect CENP-A in the human pathogenic yeast Candida albicans. PLoS Genet 8**:** e1002661.

Thakur, J., and K. Sanyal, 2013 Efficient neocentromere formation is suppressed by gene conversion to maintain centromere function at native physical chromosomal loci in Candida albicans. Genome Res 23**:** 638-652.

Theelen, B., C. Cafarchia, G. Gaitanis, I. D. Bassukas, T. Boekhout *et al.*, 2018 *Malassezia* ecology, pathophysiology, and treatment. Med Mycol 56**:** S10-25.

Thon, M. R., H. Pan, S. Diener, J. Papalas, A. Taro *et al.*, 2006 The role of transposable element clusters in genome evolution and loss of synteny in the rice blast fungus Magnaporthe oryzae. Genome Biol 7**:** R16.

Tien, J. F., N. T. Umbreit, D. R. Gestaut, A. D. Franck, J. Cooper *et al.*, 2010 Cooperation of the Dam1 and Ndc80 kinetochore complexes enhances microtubule coupling and is regulated by aurora B. J Cell Biol 189**:** 713-723.

Todd, R. T., T. D. Wikoff, A. Forche and A. Selmecki, 2019 Genome plasticity in Candida albicans is driven by long repeat sequences. Elife 8.

Tolomeo, D., O. Capozzi, R. R. Stanyon, N. Archidiacono, P. D'Addabbo *et al.*, 2017 Epigenetic origin of evolutionary novel centromeres. Sci Rep 7**:** 41980.

Tromer, E. C., J. J. E. van Hooff, G. Kops and B. Snel, 2019 Mosaic origin of the eukaryotic kinetochore. Proc Natl Acad Sci U S A 116**:** 12873-12882.

Tyler, B. M., S. Tripathy, X. Zhang, P. Dehal, R. H. Jiang *et al.*, 2006 Phytophthora genome sequences uncover evolutionary origins and mechanisms of pathogenesis. Science 313**:** 1261-1266.

Ubinski, C. V., L. S. Carvalho and M. C. Schneider, 2018 Mechanisms of karyotype evolution in the Brazilian scorpions of the subfamily Centruroidinae (Buthidae). Genetica 146**:** 475-486.

Umbreit, N. T., D. R. Gestaut, J. F. Tien, B. S. Vollmar, T. Gonen *et al.*, 2012 The Ndc80 kinetochore complex directly modulates microtubule dynamics. Proc Natl Acad Sci U S A 109**:** 16113-16118.

van Hooff, J. J., E. Tromer, L. M. van Wijk, B. Snel and G. J. Kops, 2017a Evolutionary dynamics of the kinetochore network in eukaryotes as revealed by comparative genomics. EMBO Rep 18**:** 1559-1571.

van Hooff, J. J. E., B. Snel and G. Kops, 2017b Unique Phylogenetic Distributions of the Ska and Dam1 Complexes Support Functional Analogy and Suggest Multiple Parallel Displacements of Ska by Dam1. Genome Biol Evol 9**:** 1295-1303.

Van Hooser, A. A., Ouspenski, II, H. C. Gregson, D. A. Starr, T. J. Yen *et al.*, 2001 Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. J Cell Sci 114**:** 3529-3542.

Varma, D., S. Chandrasekaran, L. J. Sundin, K. T. Reidy, X. Wan *et al.*, 2012 Recruitment of the human Cdt1 replication licensing protein by the loop domain of Hec1 is required for stable kinetochore-microtubule attachment. Nat Cell Biol 14**:** 593-603.

Varshney, N., and K. Sanyal, 2019 Aurora kinase Ipl1 facilitates bilobed distribution of clustered kinetochores to ensure error-free chromosome segregation in Candida albicans. Mol Microbiol 112**:** 569-587.

Verdel, A., S. Jia, S. Gerber, T. Sugiyama, S. Gygi *et al.*, 2004 RNAi-mediated targeting of heterochromatin by the RITS complex. Science 303**:** 672-676.

Vershinina, A. O., and V. A. Lukhtanov, 2017 Evolutionary mechanisms of runaway chromosome number change in Agrodiaetus butterflies. Sci Rep 7**:** 8199.

Vision, T. J., D. G. Brown and S. D. Tanksley, 2000 The origins of genomic duplications in Arabidopsis. Science 290**:** 2114-2117.

Volpe, T., V. Schramke, G. L. Hamilton, S. A. White, G. Teng *et al.*, 2003 RNA interference is required for normal centromere function in fission yeast. Chromosome Res 11**:** 137-146.

Volpe, T. A., C. Kidner, I. M. Hall, G. Teng, S. I. Grewal *et al.*, 2002 Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297**:** 1833-1837.

Voullaire, L. E., H. R. Slater, V. Petrovic and K. H. Choo, 1993 A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? Am J Hum Genet 52**:** 1153-1163.

Wang, H. W., S. Long, C. Ciferri, S. Westermann, D. Drubin *et al.*, 2008a Architecture and flexibility of the yeast Ndc80 kinetochore complex. J Mol Biol 383**:** 894-903.

Wang, K., Y. Wu, W. Zhang, R. K. Dawe and J. Jiang, 2014a Maize centromeres expand and adopt a uniform size in the genetic background of oat. Genome Res 24**:** 107-116.

Wang, Q. M., B. Theelen, M. Groenewald, F. Y. Bai and T. Boekhout, 2014b Moniliellomycetes and Malasseziomycetes, two new classes in Ustilaginomycotina. Persoonia 33**:** 41-47.

Wang, S., W. Jin and K. Wang, 2019 Centromere histone H3- and phospholipase-mediated haploid induction in plants. Plant Methods 15**:** 42.

Wang, Z., C. Zang, J. A. Rosenfeld, D. E. Schones, A. Barski *et al.*, 2008b Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 40**:** 897-903.

Warburton, P. E., C. A. Cooke, S. Bourassa, O. Vafa, B. A. Sullivan *et al.*, 1997 Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. Curr Biol 7**:** 901-904.

Wei, R. R., J. Al-Bassam and S. C. Harrison, 2007 The Ndc80/HEC1 complex is a contact point for kinetochore-microtubule attachment. Nat Struct Mol Biol 14**:** 54-59.

Wei, R. R., J. R. Schnell, N. A. Larsen, P. K. Sorger, J. J. Chou *et al.*, 2006 Structure of a central component of the yeast kinetochore: the Spc24p/Spc25p globular domain. Structure 14**:** 1003-1009.

Wei, R. R., P. K. Sorger and S. C. Harrison, 2005 Molecular organization of the Ndc80 complex, an essential kinetochore component. Proc Natl Acad Sci U S A 102**:** 5363-5367.

Weidman, F., 1925 Exfoliative dermatitis in the Indian rhinoceros (Rhinoceros unicornis), with description of a new species: Pityrosporum pachydermatis, pp. 36-44. Zoological Society of Philadelphia, PA.

Weir, J. R., A. C. Faesen, K. Klare, A. Petrovic, F. Basilico *et al.*, 2016 Insights from biochemical reconstitution into the architecture of human kinetochores. Nature 537**:** 249-253.

Welburn, J. P., E. L. Grishchuk, C. B. Backer, E. M. Wilson-Kubalek, J. R. Yates, 3rd *et al.*, 2009 The human kinetochore Ska1 complex facilitates microtubule depolymerization-coupled motility. Dev Cell 16**:** 374-385.

Welburn, J. P., M. Vleugel, D. Liu, J. R. Yates, 3rd, M. A. Lampson *et al.*, 2010 Aurora B phosphorylates spatially distinct targets to differentially regulate the kinetochore-microtubule interface. Mol Cell 38**:** 383-392.

Westermann, S., A. Avila-Sakar, H. W. Wang, H. Niederstrasser, J. Wong *et al.*, 2005 Formation of a dynamic kinetochore- microtubule interface through assembly of the Dam1 ring complex. Mol Cell 17**:** 277-290.

Westermann, S., I. M. Cheeseman, S. Anderson, J. R. Yates, 3rd, D. G. Drubin *et al.*, 2003 Architecture of the budding yeast kinetochore reveals a conserved molecular core. J Cell Biol 163**:** 215-222.

Westermann, S., H. W. Wang, A. Avila-Sakar, D. G. Drubin, E. Nogales *et al.*, 2006 The Dam1 kinetochore ring complex moves processively on depolymerizing microtubule ends. Nature 440**:** 565-569.

White, M. J., 1978 Modes of speciation, pp.

White, M. J. D., 1969 CHROMOSOMAL REARRANGEMENTS AND SPECIATION IN ANIMALS. Annual Review of Genetics 3**:** 75-98.

Whiteway, M., and C. Bachewich, 2007 Morphogenesis in Candida albicans. Annu Rev Microbiol 61**:** 529-553.

Wiemann, P., C. M. Sieber, K. W. von Bargen, L. Studt, E. M. Niehaus *et al.*, 2013 Deciphering the cryptic genome: genome-wide analyses of the rice pathogen Fusarium fujikuroi reveal complex regulation of secondary metabolism and novel metabolites. PLoS Pathog 9**:** e1003475.

Williams, J. S., T. Hayashi, M. Yanagida and P. Russell, 2009 Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. Mol Cell 33**:** 287-298.

Wolfe, K. H., and D. C. Shields, 1997 Molecular evidence for an ancient duplication of the entire yeast genome. Nature 387**:** 708-713.

Wong, L. H., K. H. Brettingham-Moore, L. Chan, J. M. Quach, M. A. Anderson *et al.*, 2007 Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. Genome Res 17**:** 1146-1160.

Wrighton, K. H., 2019 Malassezia restricta plays CARDs in the gut. Nat Rev Microbiol 17**:** 266-267.

Wu, G., H. Zhao, C. Li, M. P. Rajapakse, W. C. Wong *et al.*, 2015 Genus-wide comparative genomics of *Malassezia* delineates its phylogeny, physiology, and niche adaptation on human skin. PLoS Genet 11**:** e1005614.

Xia, G., X. Luo, T. Habu, J. Rizo, T. Matsumoto *et al.*, 2004 Conformation-specific binding of p31(comet) antagonizes the function of Mad2 in the spindle checkpoint. EMBO J 23**:** 3133-3143.

Xu, J., C. W. Saunders, P. Hu, R. A. Grant, T. Boekhout *et al.*, 2007 Dandruff-associated Malassezia genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens. Proc Natl Acad Sci U S A 104**:** 18730-18735.

Yadav, V., L. Sreekumar, K. Guin and K. Sanyal, 2018a Five pillars of centromeric chromatin in fungal pathogens. PLoS Pathog 14**:** e1007150.

Yadav, V., S. Sun, R. B. Billmyre, B. C. Thimmappa, T. Shea *et al.*, 2018b RNAi is a critical determinant of centromere evolution in closely related fungi. Proc Natl Acad Sci U S A 115**:** 3108-3113.

Yamagishi, Y., C. H. Yang, Y. Tanno and Y. Watanabe, 2012 MPS1/Mph1 phosphorylates the kinetochore protein KNL1/Spc7 to recruit SAC components. Nat Cell Biol 14**:** 746-752.

Yang, L., D. H. Koo, D. Li, T. Zhang, J. Jiang *et al.*, 2014 Next-generation sequencing, FISH mapping and synteny-based modeling reveal mechanisms of decreasing dysploidy in Cucumis. Plant J 77**:** 16-30.

Yang, Z., U. S. Tulu, P. Wadsworth and C. L. Rieder, 2007 Kinetochore dynein is required for chromosome motion and congression independent of the spindle checkpoint. Curr Biol 17**:** 973-980.

Yi, C., W. Zhang, X. Dai, X. Li, Z. Gong *et al.*, 2013 Identification and diversity of functional centromere satellites in the wild rice species Oryza brachyantha. Chromosome Res 21**:** 725-737.

Yunis, J. J., and O. Prakash, 1982 The origin of man: a chromosomal pictorial legacy. Science 215**:** 1525-1530.

Zeitlin, S. G., R. D. Shelby and K. F. Sullivan, 2001 CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. J Cell Biol 155**:** 1147-1157.

Zelter, A., M. Bonomi, J. O. Kim, N. T. Umbreit, M. R. Hoopmann *et al.*, 2015 The molecular architecture of the Dam1 kinetochore complex is defined by cross-linking based structural modelling. Nat Commun 6**:** 8673.

Zhang, H., and C. H. Freudenreich, 2007 An AT-rich sequence in human common fragile site FRA16D causes fork stalling and chromosome breakage in *S. cerevisiae*. Mol Cell 27**:** 367-379.

Zhang, H., A. Koblizkova, K. Wang, Z. Gong, L. Oliveira *et al.*, 2014 Boom-Bust Turnovers of Megabase-Sized Centromeric DNA in Solanum Species: Rapid Evolution of DNA Sequences Associated with Centromeres. Plant Cell 26**:** 1436-1447.

Zhang, K., K. Mosch, W. Fischle and S. I. Grewal, 2008 Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. Nat Struct Mol Biol 15**:** 381-388.

Zhang, W., B. Friebe, B. S. Gill and J. Jiang, 2010 Centromere inactivation and epigenetic modifications of a plant chromosome with three functional centromeres. Chromosoma 119**:** 553-563.

Zhao, Y., M. Dominska, A. Petrova, H. Bagshaw, R. J. Kokoska *et al.*, 2017 Properties of Mitotic and Meiotic Recombination in the Tandemly-Repeated CUP1 Gene Cluster in the Yeast Saccharomyces cerevisiae. Genetics 206**:** 785-800.

Zhu, Y., P. G. Engstrom, C. Tellgren-Roth, C. D. Baudo, J. C. Kennell *et al.*, 2017 Proteogenomics produces comprehensive and highly accurate protein-coding gene annotation in a complete genome assembly of *Malassezia sympodialis*. Nucleic Acids Res 45**:** 2629-2643.

List of publications

1. Gautam Chatterjee, **Sundar Ram Sankaranarayanan**, Krishnendu Guin, Yogitha Thattikota, Sreedevi Padmanabhan, Rahul Siddharthan, and Kaustuv Sanyal. "Repeat-associated fission yeast-like regional centromeres in the ascomycetous budding yeast Candida tropicalis." *PLoS genetics* 12, no. 2 (2016).

2. Shreyas Sridhar, Arti Dumbrepatil, Lakshmi Sreekumar, **Sundar Ram Sankaranarayanan**, Krishnendu Guin, and Kaustuv Sanyal. "Centromere and Kinetochore: Essential Components for Chromosome Segregation." *Gene Regulation, Epigenetics and Hormone Signaling* 1 (2017).

3. Zhu, Yafeng, Pär G. Engström, Christian Tellgren-Roth, Charles D. Baudo, John C. Kennell, Sheng Sun, R. Blake Billmyre, Markus S. Schröder, Anna Andersson, Tina Holm, Benjamin Sigurgeirsson, Guangxi Wu, **Sundar Ram Sankaranarayanan**, Rahul Siddharthan, Kaustuv Sanyal, Joakim Lundeberg, Björn Nystedt, Teun Boekhout, Thomas L. Dawson, Jr., Joseph Heitman, Annika Scheynius, Janne Lehtiö. "Proteogenomics produces comprehensive and highly accurate protein-coding gene annotation in a complete genome assembly of *Malassezia sympodialis*." *Nucleic acids research* 45, no. 5 (2017): 2629-2643.

4. **Sundar Ram Sankaranarayanan**, Giuseppe Ianiri, Marco A. Coelho, Md Hashim Reza, Bhagya C. Thimmappa, Promit Ganguly, Rakesh Netha Vadnala, Sheng Sun, Rahul Siddharthan, Christian Tellgren-Roth, Thomas L Dawson Jr, Joseph Heitman, Kaustuv Sanyal "Loss of centromere function drives karyotype evolution in closely related *Malassezia* species." *eLife* 9 (2020): e53944.

PLOS | GENETICS

RESEARCH ARTICLE

Repeat-Associated Fission Yeast-Like Regional Centromeres in the Ascomycetous Budding Yeast *Candida tropicalis*

Gautam Chatterjee[1¤a], Sundar Ram Sankaranarayanan[1], Krishnendu Guin[1], Yogitha Thattikota[1¤b], Sreedevi Padmanabhan[1¤c], Rahul Siddharthan[2], Kaustuv Sanyal[1]*

1 Molecular Mycology Laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore, India, **2** The Institute of Mathematical Sciences, C.I.T. Campus, Taramani, Chennai, India

¤a Current address: Department of Agricultural Biotechnology, Faculty Centre of Integrated Rural Development and Management, Ramakrishna Mission Vivekananda University, Narendrapur, Kolkata, India
¤b Current address: Institute for Research in Immunology and Cancer, Université de Montréal, Station Centre-Ville, Montreal, Canada
¤c Current address: Molecular Biology Laboratory, Department of Biotechnology, Veer Bahadur Singh Purvanchal University, Jaunpur, India
* sanyal@jncasr.ac.in

Citation: Chatterjee G, Sankaranarayanan SR, Guin K, Thattikota Y, Padmanabhan S, Siddharthan R, et al. (2016) Repeat-Associated Fission Yeast-Like Regional Centromeres in the Ascomycetous Budding Yeast *Candida tropicalis*. PLoS Genet 12(2): e1005839. doi:10.1371/journal.pgen.1005839

Editor: Beth A. Sullivan, Duke University, UNITED STATES

Received: August 4, 2015

Accepted: January 11, 2016

Published: February 4, 2016

Data Availability Statement: The sequencing data generated by the ChIP-seq experiments in this study have been submitted to the NCBI under the accession number SUB432163. The sequencing data generated by Sanger sequencing of Scnt 1, Scnt 7 and Scnt 9 have been deposited to the NCBI with following accession numbers KJ398406, KJ425116 and KJ398405 respectively.

Funding: This work is supported by a grant from DBT (Grant number: BT/PR14840/BRB/10/880/2010), Govt. of India and intramural funding of JNCASR to KS. RS acknowledges the PRISM project

Abstract

The centromere, on which kinetochore proteins assemble, ensures precise chromosome segregation. Centromeres are largely specified by the histone H3 variant CENP-A (also known as Cse4 in yeasts). Structurally, centromere DNA sequences are highly diverse in nature. However, the evolutionary consequence of these structural diversities on *de novo* CENP-A chromatin formation remains elusive. Here, we report the identification of centromeres, as the binding sites of four evolutionarily conserved kinetochore proteins, in the human pathogenic budding yeast *Candida tropicalis*. Each of the seven centromeres comprises a 2 to 5 kb non-repetitive *mid* core flanked by 2 to 5 kb inverted repeats. The repeat-associated centromeres of *C. tropicalis* all share a high degree of sequence conservation with each other and are strikingly diverged from the unique and mostly non-repetitive centromeres of related *Candida* species—*Candida albicans*, *Candida dubliniensis*, and *Candida lusitaniae*. Using a plasmid-based assay, we further demonstrate that pericentric inverted repeats and the underlying DNA sequence provide a structural determinant in CENP-A recruitment in *C. tropicalis*, as opposed to epigenetically regulated CENP-A loading at centromeres in *C. albicans*. Thus, the centromere structure and its influence on *de novo* CENP-A recruitment has been significantly rewired in closely related *Candida* species. Strikingly, the centromere structural properties along with role of pericentric repeats in *de novo* CENP-A loading in *C. tropicalis* are more reminiscent to those of the distantly related fission yeast *Schizosaccharomyces pombe*. Taken together, we demonstrate, for the first time, fission yeast-like repeat-associated centromeres in an ascomycetous budding yeast.

9
Centromere and Kinetochore: Essential Components for Chromosome Segregation

Shreyas Sridhar, Arti Dumbrepatil, Lakshmi Sreekumar, Sundar Ram Sankaranarayanan, Krishnendu Guin, and Kaustuv Sanyal

Jawaharlal Nehru Centre for Advanced Scientific Research, Molecular Biology and Genetics Unit, Molecular Mycology Laboratory, Jakkur, Bangalore 560 064 India

9.1
Introduction

Perpetuation of life occurs by the fundamental property of cells to divide. A somatic cell undergoes a cell cycle that is comprised of essentially two periods: interphase and mitosis. Interphase can be further divided into G1, S, and G2. G1 and G2 constitute gap phases, involving cell growth that prepare cells for genome duplication in synthesis (S) phase and subsequent segregation in mitotic (M) phase, respectively. The mitotic cell cycle ensures equal division of the duplicated genetic content of the mother nucleus with the help of the kinetochore and centromere. The kinetochore is a proteinaceous structure that assembles on centromere (*CEN*) DNA. The centromere/kinetochore generally appears as a constricted region of a metaphase chromosome (Figure 9.1a). The kinetochore complex interacts with microtubules on one side and centromeric chromatin on the other (Figure 9.1a). In most metazoans, multiple microtubules bind to each kinetochore, with an exception of certain budding yeasts where only a single microtubule appears to be associated with each kinetochore [1–4].

Apart from these general features of mitosis, organism-specific variations also exist. Mitosis is broadly classified in two types: closed mitosis and open mitosis (Figure 9.1b). This distinction primarily refers to the permeability of the nuclear envelope (NE), a bilayered membrane which along with the nuclear pore complexes (NPCs) regulate the entry and exit of molecules to and from the nucleus. Closed mitosis is considered to be the more primitive form of eukaryotic cell division, whereas open mitosis seems to have appeared several times during evolution. Plants and animals share open mitosis predominantly, while most fungi employ closed mitosis and variations of it. During closed mitosis, the NE

Gene Regulation, Epigenetics, and Hormone Signaling, First Edition. Edited by Subhrangsu S. Mandal.

Published online 18 January 2017

Nucleic Acids Research, 2017, Vol. 45, No. 5 **2629–2643**
doi: 10.1093/nar/gkx006

Proteogenomics produces comprehensive and highly accurate protein-coding gene annotation in a complete genome assembly of *Malassezia sympodialis*

Yafeng Zhu[1,†], Pär G. Engström[2,†], Christian Tellgren-Roth[3], Charles D. Baudo[4], John C. Kennell[4], Sheng Sun[5], R. Blake Billmyre[5], Markus S. Schröder[6], Anna Andersson[7], Tina Holm[7], Benjamin Sigurgeirsson[8], Guangxi Wu[9], Sundar Ram Sankaranarayanan[10], Rahul Siddharthan[11], Kaustuv Sanyal[10], Joakim Lundeberg[8], Björn Nystedt[12], Teun Boekhout[13], Thomas L. Dawson, Jr.[14], Joseph Heitman[5], Annika Scheynius[15,*,‡] and Janne Lehtiö[1,*,‡]

[1]Science for Life Laboratory, Department of Oncology-Pathology, Karolinska Institutet, 17121 Solna, Sweden, [2]Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, 17121 Solna, Sweden, [3]National Genomics Infrastructure, Science for Life Laboratory, Department of Immunology, Genetics and Pathology, Uppsala University, 75108 Uppsala, Sweden, [4]Department of Biology, Saint Louis University, St. Louis, MO 63103, USA, [5]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA, [6]School of Biomedical and Biomolecular Science, Conway Institute, University College Dublin, Belfield, Dublin 4, Ireland, [7]Department of Medicine Solna, Translational Immunology Unit, Karolinska Institutet and University Hospital, 17177 Stockholm, Sweden, [8]Science for Life Laboratory, School of Biotechnology, Royal Institute of Technology, 17121 Solna, Sweden, [9]Computational and Systems Biology, Genome Institute of Singapore, Agency for Science, Technology and Research (A*STAR), 138672, Singapore, [10]Molecular Mycology Laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560 064, India, [11]The Institute of Mathematical Sciences/HBNI, Taramani, Chennai 600 113, India, [12]Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, 75123 Uppsala, Sweden, [13]CBS-Fungal Biodiversity Centre, Utrecht, 3508, The Netherlands and Institute for Biodiversity and ecosystem Dynamics (IBED), University of Amsterdam, 1012 WX Amsterdam, The Netherlands, [14]Institute of Medical Biology, Agency for Science, Technology and Research (A*STAR), 138648, Singapore and [15]Science for Life Laboratory, Department of Clinical Science and Education, Karolinska Institutet, and Sachs' Children and Youth Hospital, Södersjukhuset, SE-118 83 Stockholm, Sweden

Received August 22, 2016; Revised December 23, 2016; Editorial Decision December 31, 2016; Accepted January 16, 2017

ABSTRACT

Complete and accurate genome assembly and annotation is a crucial foundation for comparative and functional genomics. Despite this, few complete eukaryotic genomes are available, and genome annotation remains a major challenge. Here, we present a complete genome assembly of the skin commensal yeast *Malassezia sympodialis* and demonstrate how proteogenomics can substantially improve gene annotation. Through long-read DNA sequencing, we obtained a gap-free genome assembly for *M. sympodialis* (ATCC 42132), comprising eight nuclear and one mitochondrial chromosome. We also sequenced and assembled four *M. sympodialis* clinical isolates, and showed their value for understanding *Malassezia* reproduction by confirming four alternative allele combinations at the two mating-type loci. Importantly, we demonstrated how proteomics data could be readily integrated with transcriptomics data in standard

*To whom correspondence should be addressed. Tel: +46 8 52481416; Email: janne.lehtio@ki.se
Correspondence may also be addressed to Annika Scheynius. Tel: +46 70 6057927; Email: annika.scheynius@ki.se
†These authors contributed equally to this work as the first authors.
‡These authors contributed equally to this work as the last authors.

eLIFE elifesciences.org

RESEARCH ARTICLE

Loss of centromere function drives karyotype evolution in closely related *Malassezia* species

Sundar Ram Sankaranarayanan[1], Giuseppe Ianiri[2†], Marco A Coelho[2], Md Hashim Reza[1], Bhagya C Thimmappa[1‡], Promit Ganguly[1], Rakesh Netha Vadnala[3], Sheng Sun[2], Rahul Siddharthan[3], Christian Tellgren-Roth[4], Thomas L Dawson Jnr[5,6], Joseph Heitman[2]*, Kaustuv Sanyal[1]*

[1]Molecular Mycology Laboratory, Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bengaluru, India; [2]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, United States; [3]The Institute of Mathematical Sciences/HBNI, Chennai, India; [4]National Genomics Infrastructure, Science for Life Laboratory, Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden; [5]Skin Research Institute Singapore, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore; [6]Department of Drug Discovery, Medical University of South Carolina, School of Pharmacy, Charleston, United States

***For correspondence:** heitm001@duke.edu (JH); sanyal@jncasr.ac.in (KS)

Present address: †Department of Agricultural, Environmental and Food Sciences, University of Molise, Campobasso, Italy; ‡Department of Biochemistry, Robert-Cedergren Centre for Bioinformatics and Genomics, University of Montreal, Montreal, Canada

Competing interests: The authors declare that no competing interests exist.

Funding: See page 26

Received: 26 November 2019
Accepted: 20 January 2020
Published: 20 January 2020

Reviewing editor: Wolf-Dietrich Heyer, University of California, Davis, United States

Abstract Genomic rearrangements associated with speciation often result in variation in chromosome number among closely related species. *Malassezia* species show variable karyotypes ranging between six and nine chromosomes. Here, we experimentally identified all eight centromeres in *M. sympodialis* as 3–5-kb long kinetochore-bound regions that span an AT-rich core and are depleted of the canonical histone H3. Centromeres of similar sequence features were identified as CENP-A-rich regions in *Malassezia furfur*, which has seven chromosomes, and histone H3 depleted regions in *Malassezia slooffiae* and *Malassezia globosa* with nine chromosomes each. Analysis of synteny conservation across centromeres with newly generated chromosome-level genome assemblies suggests two distinct mechanisms of chromosome number reduction from an inferred nine-chromosome ancestral state: (a) chromosome breakage followed by loss of centromere DNA and (b) centromere inactivation accompanied by changes in DNA sequence following chromosome–chromosome fusion. We propose that AT-rich centromeres drive karyotype diversity in the *Malassezia* species complex through breakage and inactivation.

Introduction

Centromeres are the genomic loci on which the kinetochore, a multi-subunit complex, assembles to facilitate high-fidelity chromosome segregation. The centromere-specific histone H3 variant CENP-A is the epigenetic hallmark of centromeres, as it replaces canonical histone H3 in the nucleosomes to make specialized centromeric chromatin that acts as the foundation to recruit other kinetochore proteins. A remarkable diversity in the organization of centromere DNA sequences has been observed to accomplish this conserved role (***Roy and Sanyal, 2011***; ***Yadav et al., 2018b***).

The smallest known centromeres are the point centromeres present in budding yeasts of the family *Saccharomycetaceae* that span <200 bp in length (***Clarke and Carbon, 1980***; ***Gordon et al., 2011***; ***Kobayashi et al., 2015***). These centromeres are organized into conserved DNA elements I, II, and III that are recognized by a cognate kinetochore protein complex called the CBF3 complex,