Supplementary information for

Centromere-mediated chromosome break drives karyotype evolution in closely related *Malassezia* species

Sundar Ram Sankaranarayanan[1], Giuseppe Ianiri[2], Md. Hashim Reza[1], Bhagya C. Thimmappa[1,#], Promit Ganguly[1], Marco A. Coelho[2], Sheng Sun[2], Rahul Siddharthan[3], Christian Tellgren-Roth[4], Thomas L Dawson Jr.[5,6], Joseph Heitman[2,*] and Kaustuv Sanyal[1,*]

Corresponding authors

Joseph Heitman (heitm001@duke.edu)

Kaustuv Sanyal (sanyal@jncasr.ac.in)

This file includes

Supplementary materials and methods

Figures S1-S4

Tables S1-S8

Supplementary references

Supplementary Materials and Methods

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

The allele for N-terminal tagging of Mtw1 with GFP was prepared by gap repair (1) in the *Saccharomyces cerevisiae* BY4741 strain. Briefly, a 1.6 kb fragment consisting of upstream and promoter sequence of the *MTW1* gene and a 1.6 kb fragment having *MTW1* ORF along with downstream sequence was amplified from *M. sympodialis* genomic DNA. The GFP ORF (without the stop codon) and NAT were amplified from pVY7 and pAIM1 respectively (2, 3). *S. cerevisiae* was transformed with all four fragments and the linearized plasmid pGI3 (digested with KpnI and BamHI) and the epitope tagged allele was assembled in an ordered way by gap repair. Total DNA was isolated from *S. cerevisiae* and *E. coli* DH5α strain was transformed. 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 (2, 4).

Microscopic imaging of live cells and processing

The GFP-Mtw1 strain was inoculated to 1% v/v from a saturated starter culture grown in mDixon media. Upon growth for 6 h, these cells were pelleted at 4,000 rpm and washed thrice 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 prior imaging. The images were acquired at room temperature using laser scanning inverted confocal microscope LSM 880-Airyscan (ZEISS, Plan Apochromat 63x, NA oil 1.4) equipped with highly sensitive photo-detectors. 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/ ImageJ. All the images were processed post acquisition with minimal adjustments to levels and linear contrast until the signals were highlighted.

Preparation of *M. sympodialis* spheroplasts

Cells grown on mDixon’s media were washed with water by centrifugation at 4000 rpm for 5 minutes. Cells were resuspended in 10 mL of 5% (v/v) 2-mercaptoethanol solution in water, 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 haziarnum* and Zymolyase-20T were added at 20 mg/mL and 100 µg/mL respectively. The spheroplasting suspension was incubated at

30°C/ 65 rpm for 6 to 8 h. Digestion was checked by microscopic observation. Spheroplasts were washed with ice cold PBS and used as per the experimental design. (Adapted from (5)).

Indirect Immunofluorescence

The GFP-Mtw1 strain was inoculated to 1% (v/v) from a saturated starter culture grown in mDixon media. After growth for 6 h, the cells were fixed by addition of formaldehyde to a final concentration of 3.7% for 1 h. Post fixing, the cells were washed with water and taken for preparation of spheroplasts (described above). Spheroplasts were washed with ice cold 1x-PBS and resuspended in ice cold 1x-PBS 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. Cell suspension was added to each well (15-20 µL) and allowed to stand at room temperature for 5 min. Cell suspension was aspirated and the slides were washed with water once 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 1x-PBS) 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, the slide was washed 8 times with 1x-PBS giving 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 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 inverted fluorescence microscope (ZEISSAxio 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.

Chromatin immunoprecipitation and sequencing

The protocol used for ChIP was adapted from the protocol used for *C. neoformans* (6). Logarithmically grown cells were fixed using formaldehyde at a final concentration of 1% for 30 min (for Mtw1 ChIP) and 15 min (for histone H3 ChIP) respectively. The reaction was quenched by the addition of glycine to a final concentration of 0.135 M. Cells were pelleted and processed for spheroplasting as described above. Spheroplasts were washed once sequentially using 10 mL of the following ice-cold buffers: 1x PBS, Buffer-I (0.25% Triton X-100, 10 mM EDTA, 0.5 mM EGTA, 10 mM Na-HEPES pH=6.5), Buffer-II (200 mM NaCl, 1mM EDTA, 0.5 mM EGTA, 10 mM Na-

HEPES pH=6.5). The pellet after final wash was resuspended in 1 mL lysis buffer (50 mM HEPES pH=7.4, 1% Triton X-100, 140 mM NaCl, 0.1% Na-deoxycholate, 1 mM EDTA) for every $1.5x10^9$ cells. Protease inhibitor cocktail was added to 1x final concentration. The resuspended spheroplasts were subjected to sonication using a Bioruptor (Diagenode) using 30 s ON/OFF pulse at high intensity mode with intermittent incubation in ice to obtain chromatin fragments of range 100-300 bp. Lysate was cleared after sonication by centrifugation at 13,000 rpm for 10 min at 4ºC. Input DNA fraction was separated at this step ($1/10^{th}$ volume of lysate) and processed for de-crosslinking by addition of 400 µL elution buffer (0.1 M $NaHCO_3$, 1% SDS) per 100 µL lysate (processing for de-crosslinking mentioned below). The remaining lysate was split equally and processed as IP and control samples. 20 µL GFP trap and blocked agarose beads respectively were used for IP and control. In the case of histone H3 ChIP, 10 µL anti H3 antibodies were used per IP 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 (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH=8.0, 150 mM NaCl), twice with 1 mL high salt wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH=8.0, 500 mM NaCl), once with 1 mL LiCl wash buffer (0.25 M LiCl, 1% NP-40, 1% Na-deoxycholate, 1 mM EDTA, 10 mM Tris pH=8.0) and twice with 1 mL 1xTE (10 mM Tris pH=8.0, 1 mM EDTA). Samples were rotated in a rotaspin for 5 min at room temperature for every wash (15 min in case of Histone H3 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. Samples were decrosslinked by addition of 20 µL 5 M NaCl and incubation at 65ºC for 6 h. Following this, samples were deproteinized by addition of 10 µL 0.5 M EDTA, 20 µL 1 M Tris pH6.8, 2 µL Proteinase K (20 mg/L) and incubation at 45ºC for 2 h. After incubation, samples were treated with 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 (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 followed by washing the pellet once with 70% ethanol. Air dried pellets were then resuspended in 30 µL sterile MilliQ water with 10 µg/mL RNAse. Samples were analysed by qPCR using the primers mentioned in the key resources table.

Analysis of sequencing data

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 systems) and scaled for representation purpose. RNA-sequencing data (E-MTAB-4589) from a previous study was downloaded 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.

Western Blotting

Protein lysates for western blot were prepared by the TCA method. One mL overnight grown cultures 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 desired volume of lysis buffer (0.1 N NaOH+1% SDS). Samples were separated in 12% polyacrylamide gels, transferred onto Nitrocellulose membrane. For probing, mouse anti-GFP antibody (Roche) and the HRP conjugated goat anti-mouse secondary antibody (Bangalore Genei), were used at 1:3000 and 1:5000 dilution respectively in 2.5% skim milk powder in 1x-PBS. The blots were developed using Chemiluminescence Ultra substrate (BioRad) and imaged using VersaDoc imaging system.

PFGE analysis for *M. globosa* and *M. slooffiae*

For CHEF analysis of *M. globosa* (CBS7966) and *M. slooffiae* (CBS7956), the cells grown on solid mDixon medium were collected and resuspended in PBS. CHEF plugs were prepared as described in previous studies (7, 8). Chromosomes were separated in 1% Megabase certified agarose gel made with 0.5 × TBE, using a BioRad CHEF-DR II System running at 3.2 V/cm with linear ramping switching time from 90 to 360 seconds for 120 hours in 0.5 × TBE at 14 °C. The gel was stained with EtBr and visualized under UV.

For the chromoblot analyses of *M. globosa* (CBS7966), the gel from the CHEF analysis was first transferred to membrane, and the resulting chromoblots were then hybridized with four probes from the chromosomes 3, 4, 5, and 6 of the CBS7966 genome assembly, respectively (see the Table for the primer information), as described in previous studies (6, 9).

M. globosa genome assembly

Sequence reads were assembled using HGAP3 included in SMRTPortal v2.3 (PacBio, Menlo Park, CA, USA) with default parameters except for the genome size set to 9 Mb. Assembly completeness was evaluated by checking for telomeric repeats. Non-telomeric contig-ends were aligned to other contigs using BLAST and unique overlaps used to build complete chromosomes. Short telomere ends were extended using uniquely mapping reads longer than 10 kb and repolishing of the assembly using the resequencing pipeline in SMRTPortal v2.3. The assembly resulted in 19 contigs, with a total length of 9.2 Mb. 17 long and 1 short telomere could be identified. 6 contigs had telomeres on the 5'- and 3'-end thus representing full-length chromosomes (Chr1, 2, 3, 6, 7 and 8). 6 contigs had only one telomere and 7 contigs had no telomeric sequence. 2 contigs without telomeres were from the mitochondrion and 2 were from the ribosomal repeats. Chromosome 5 was constructed from 2 contigs ending in ribosomal repeats. The assembly contains 6 copies of the repeat, but read coverage suggested a length of 30-40 repeat units that cannot be resolved with the available read length. The remaining contigs were used to build chromosomes 4 and 9 that share highly similar 5'-ends. The two ends can be distinguished by two microsatellite expansions. Chromosome 4 had a very short 3'-telomere from the default assembly, but the raw data contained a uniquely mapping read that extended several repeat units past the assembly end. After polishing the reference, all 9 chromosomes had clear 5'- and 3'- telomeres.

M. slooffiae genome assembly

Sequence reads were assembled using HGAP3 included in SMRTPortal v2.3 (PacBio, Menlo Park, CA, USA) with default parameters. This resulted in an assembly with 14 scaffolds with telomeric repeats at both ends in 9 contigs. Of the remaining five contigs, three of them could be assigned to mitochondrial DNA based on BLAST analysis with *M. globosa.* The remaining two contigs of sized 5.8 kb and 2.3 kb respectively did not show BLAST hits against *M. globosa* or *M. sympodialis* genomes. Synteny analysis in this study was done using this assembly.

Synteny analysis

For synteny conservation across the centromeres as indicated in Figure S2B, 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 were done in context of ORFs flanking the centromeres of *M. sympodialis*. The protein sequences for each of these ORFs were used as query in BLAST analysis against the genome of other species. Local database for each genome was setup using Geneious software for this analysis. The percentage identity values for each ORFs are mentioned in the boxes in Figure S2B. Additionally, synteny analyses between *M. globosa* and *M. sympodialis* were conducted using megablast

(word size: 28) and plotted together with GC content (calculated as deviation from the genomic mean, in non-overlapping 1 kb windows), using Circos (v0.69-6) (10). Additional whole-genome alignments were conducted with Satsuma (11), with default parameters. The linear synteny comparisons shown in Figure 3C were generated with Python application EasyFig (12).

Supplementary figure legends

Figure S1. Properties of centromeres in *M. sympodialis* and *M. globosa*

(A) A 30 kb window of Mtw1 enrichment profile (*CEN*, red) plotted with the GC content (%GC, blue) and regions of transcription (RNA-seq, green). Ruler in each box indicates chromosomal coordinates in kb. The scales in *y*-axis are as follows: *CEN* (0-8000), %GC (0-80), RNA-seq reads (0-1500). Green arrows in each panel indicate predicted ORFs based on RNA-seq data. (B) Heat-map indicating the region having the maximum enrichment observed from the ChIP-seq data in *M. sympodialis*. *CEN*, Mtw1 enriched regions; %GC, GC content of the region; ORF, genes present in the Mtw1 enriched regions predicted from RNA-seq data of *M. sympodialis*. Pink colored box represent heat-maps showing maximum enrichment (dark red) at the intergenic regions. (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. Genome average considered for comparison is labeled as global. The *y*-axis represents Log_2RPKM values. (D) Fold change in Mtw1 and histone H3 enrichment at *CEN7* as compared to control region by qPCR analysis. Schematic representation of a 5 kb region of Chr7 with *CEN7* (Red box) is depicted below the graph. Black 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 Control- ORF region (190 kb away from *CEN1*). The *x*-axis indicates regions across the *CEN* probed by PCR and *y*-axis indicates fold reduction in histone H3 enrichment. Error bars indicate standard error of the mean (SEM). (E) Dot-plot generated by plotting sequences from the Mtw1 enriched regions in *M. sympodialis* against themselves is depicted. (F) Dot-plot generated by plotting sequences of the intergenic regions harboring the GC troughs (predicted centromeres) in *M. globosa* against themselves. Diagonal lines in C and D indicate self-identity.

Figure S2. Comparative analysis of *M. globosa* and *M. sympodialis* genomes.

(A) Electrophoretic karyotype of *M. globosa* and *S. cerevisiae*. Chromosome sizes marked for *M. globosa* are estimated from the genome assembly. The Asterisks indicate doublet bands. Chr5 that contains rDNA locus (902 kb) is a part of the doublet marked with an asterisk. (B) Electrophoretic karyotype of *M. slooffiae* and *S. cerevisiae*. Chromosome sizes marked for *M. globosa* are estimated from the genome

assembly. Asterisks indicate doublet bands. (C) Zoomed-in image of the synteny breakpoint shown for MgChr2 representing the conservation of synteny at the ORF level. Pink arrows represent ORFs from MgChr2 and their homologs on MsChr2 and MsChr4, with orange ribbons representing their synteny. Purple arrows represent ORFs from MgChr4 and their homologs in MsChr4 with blue ribbon connectors depicting synteny. Labels in black circles mark the synteny breakpoints. Synteny breakpoint of MgChr2 is marked as *MgCEN2*(III). The regions on MsChr2 and MsChr4 where the homologs of ORFs flanking the breakpoint are located are marked II and IV. The synteny block start site between MgChr4 and MsChr4 of *M. globosa* is labeled V.

Figure S3. The 12 bp AT-rich motif is enriched across centromeres of *Malassezia* species used in this study. (A-F) The genomes of *M. sympodialis, M. globosa, M. restricta, M. slooffiae, M. dermatis* and *M. nana* were scanned for matches to the 12bp 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.

Figure S4. Eight of the nine predicted centromeres of *M. globosa* and *M. slooffiae* are flanked by ORFs with synteny to *M. sympodialis* centromeres. Gene order and synteny of ORFs flanking the centromeres of 6 *Malassezia* species analyzed in this study using the protein sequences (MSYGxxxx) from *M. sympodialis* as query. Species included are as follows: Mna- *M. nana*, Mde- *M. dermatis*, Msy- *M. sympodialis*, Mgl- *M. globosa*, Msl- *M. slooffiae*, and Mre- *M. restricta*. Chromosome/scaffold numbers are indicated at the start of every track. BCL2.1 and BCK2.1 are abbreviations for scaffolds BCLA0104.1 and BCKX0104.1 respectively. Boxes represent ORFs with the numbers inside them indicating percentage identity from BLAST analysis. Broken lines marked with filled circle towards the ends indicate synteny break. Blue arrows indicate inverted orientation of the scaffold/chromosome.

Supplementary tables

Supplementary table S1. Identification of kinetochore proteins in *M. sympodialis* by BLAST

Sequences of *C. neoformans* homologs were used as query. Asterisk (*) indicates cases where sequences from *Ustilago maydis* homologs were used as query. ND- not detected. To detect proteins of Cbf3 complex, sequences of *S. cerevisiae* homologs were used as query.

	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	ND			
		Spc34	ND			
		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	ND			
		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	ND			
		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	ND			
		Ctf3	ND			

Supplementary table S2. Coordinates of centromeres and their GC content in *M. sympodialis*

Coordinates, length and GC content of Mtw1 enriched regions in comparison with that of the predicted centromeres in *M. sympodialis.*

Chr.	*CEN* start (bp)	*CEN* end (bp)	*CEN* length (bp)	GC content		Predicted *CEN* (AT-rich core)		
				CEN	Chr.	Coordinates	Start (bp)	End (bp)
Chr 1	784833	788599	3767	41	58	Chr 1	786541	787061
Chr 2	354218	357486	3269	41	58.6	Chr 2	355760	355841
Chr 3	235615	239940	4326	44	59.5	Chr 3	237534	238686
Chr 4	415985	420656	4672	58	60.7	Chr 4	418202	418728
Chr 5	100342	105251	4910	48	60	Chr 5	101950	102502
Chr 6	430028	433194	3167	48	60.3	Chr 6	431542	431987
Chr 7	22334	27476	5143	51	60.9	Chr 7	24694	25564
Chr R	123219	127284	4066	47	57.6	Chr R	125056	125220

Table S3. Coordinates of centromeres in closely related *Malassezia* species

Coordinates and length of centromeres predicted based on GC troughs and conservation of gene synteny with centromeres of *M. sympodialis*. '*' indicates scaffolds containing rDNA locus, '#' indicates partial synteny conservation which includes *CEN* and ORFs on one side of the centromeres.

	Chr./Scaffold	*CEN*	Start	End	Length (bp)	Syntenous *MsCEN*
M. globosa	Chr1	*CEN1*	981894	982242	349	*CEN3*
	Chr2	*CEN2*	362480	362807	327	-
	Chr3	*CEN3*	219647	220121	474	*CEN1*
	Chr4	*CEN4*	152635	152994	359	*CEN4*
	Chr5	*CEN5*	464007	464114	107	*CEN2*
	Chr6	*CEN6*	736701	737015	314	*CEN5*
	Chr7	*CEN7*	59472	59817	345	*CEN7*
	Chr8	*CEN8*	110988	114481	3493	*CEN6*
	ChrR*	*CENR*	215437	215595	158	*CENR*
M. slooffiae	unitig_1|quiver	*CEN2*	132,717	133,193	477	*CEN1*
	unitig_2|quiver	*CEN3*	367,665	368,177	513	*CEN3*
	unitig_3|quiver	*CEN4*	130,942	131,501	560	*CEN4*
	unitig_4|quiver	*CEN5*	183,442	183,981	540	-
	unitig_5|quiver	*CEN6*	411,984	412,552	569	*CEN2*
	unitig_6|quiver	*CEN7*	54,307	54,889	583	*CEN5*
	unitig_7|quiver	*CEN8*	497,637	498,149	513	*CEN6*
	unitig_8|quiver	*CEN9*	55,948	56,479	532	*CEN7*
	unitig_0|quiver*	*CENR*	138,919	139,465	547	*CENR*

M. restricta	Scaffold 1	*CEN1*	347,813	348,406	594	*CEN4*#
	Scaffold 2	*CEN2*	87,190	87,806	617	*CENR*
	Scaffold 3	*CEN3*	1,101,494	1,102,083	590	*CEN3*
	Scaffold 4	*CEN4*	754,356	754,989	634	*CEN1*
	Scaffold 6	*CEN6*	621,177	621,863	687	*CEN5*
	Scaffold 7	*CEN7*	390,657	391,286	630	*CEN2*
	Scaffold 8	*CEN8*	362,842	363,381	540	-
	Scaffold 9	*CEN9*	117,021	117,603	583	*CEN7*
	Scaffold 5*	*CENR*	70,306	70,913	608	*CEN6*
M. nana	BCLA01000001.1	*CEN1*	715,036	715,592	557	*CEN1*
	BCLA01000002.1	*CEN2*	349,428	350,120	693	*CEN2*
	BCLA01000003.1	*CEN3*	220,773	221,345	573	*CEN3*
	BCLA01000004.1	*CEN4*	410,594	411,387	794	*CEN4*
	BCLA01000005.1	*CEN5*	524,594	525,105	512	*CEN5*
	BCLA01000007.1	*CEN7*	408,363	409,067	705	*CEN6*
	BCLA01000008.1	*CEN8*	398,756	399,423	668	*CEN7*
	BCLA01000006.1*	*CENR*	133,647	134,324	678	*CENR*
M. dermatis	BCKX01000001.1	*CEN1*	711,456	711,978	523	*CEN1*
	BCKX01000002.1	*CEN2*	1,014,281	1,014,977	697	*CEN2*
	BCKX01000003.1	*CEN3*	232,065	232,795	731	*CEN3*
	BCKX01000004.1	*CEN4*	409,839	410,631	793	*CEN4*
	BCKX01000005.1	*CEN5*	94,520	95,018	499	*CEN5*
	BCKX01000006.1	*CEN6*	473,487	474,334	848	*CENR*
	BCKX01000007.1	*CEN7*	76,361	76,975	615	*CEN6*
	BCKX01000008.1	*CEN8*	17,893	18,540	648	*CEN7*

Supplementary table S4. List of key reagents used in this study

Antibodies and reagents	**Source**	**Identifier**
mouse anti-GFP	Roche	11814460001
mouse anti-PSTAIRE	Abcam	Cat. no.10345
goat anti-mouse HRP	Bangalore Genei	Cat. no. HP06
rabbit anti-H3	Abcam	Cat. no. ab1791
Lysing enzymes from *Trichoderma harzianum*	Sigma	Cat. no. L1412
Zymolyase 20T	MP biomedicals	Cat. no. 320921
GFP trap beads	ChromoTek	Cat. no. gta-20
Blocked agarose beads	ChromoTek	Cat. no. bab-20
Protein-A sepharose beads	Sigma	Cat. no. P3391
2-meraptoethanol	HiMedia	Cat. no. MB041

Supplementary table S5. List of strains used in this study

Yeast Strains	
ATCC42132	*Wild-type Malassezia sympodialis*
MSY001	GFP-Mtw1-NAT in *M. sympodialis* ATCC42132
CBS7966	Wild-type *Malassezia globosa*
CBS7956	Wild-type *Malassezia slooffiae*
BY4741	Wild-type *S. cerevisiae strain* (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*)

Supplementary table S6. List of plasmids used in this study

Plasmids	
Plasmid: pMHR04 (pGI3-GFP-Mtw1)	This study
Plasmid: pGI3	(13)
Plasmid: pVY7	(3)
Plasmid: pAIM1	(2)

Supplementary table S7. List of oligonucleotides used in this study

Oligonucleotides		
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	
MS1 F1	AAGAATTGATAACATTGTTGCAC	*CEN1* primers

MS1 R1	TAGAATAAAATGTCGCGAAGG	
MS2 F1	CTGAAGAAAAGAAACAAATTCG	*CEN2* primers
MS2 R1	TCGGAAATCCCGCAAAAG	
MS3 F1	CATATTCAGCCTCCACTAAG	*CEN3* primers
MS3 R1	CCTCTATCGAGTGCTCTAC	
MS4 F1	CGATATGGATTGGACTTATAAGTC	*CEN4* primers
MS4 R1	AAAAGCAATACGTAGACGG	
MSChrR F1	AAATTACCGACCAGAATTG	*CENR* primers
MSChrR R1	ATCTGTGTCCGCTCTCATC	
MS5 F1	TTTGACGCTTTATTTGTGTTTC	*CEN5* primers
MS5 R1	CACATATGCACGAATAATAAAACG	
MS6 F1	GATACATATTCTTACACTAATACTATTCG	*CEN6* primers
MS6 R1	GCATAGAGCTAATATCTGATATTC	
MS7 F1	GGAAGCATGAGATATTGG	*CEN7* primers
MS7 R1	AAACAAAGTAAAATTCTAATCACG	
MS7 LF1	CTCCTCCGATACGATTCAC	*CEN7* L1 primers
MS7 LR1	CAGCCATTATCTCCGACAC	
MS7 LF2	CTGGGTAGATTGAGAATGAG	*CEN7* L2 primers
MS7 LR2	CATGTATGTTCAGTCCCATG	
MS7 RF1	ATGATCCAAAAGAAAGCATAC	*CEN7* R1 primers
MS7 RR1	GAAGTATGTCTGGGTGAAGC	
MS C3	GAAGACGACAACGATACC	Control primers away from *CEN1*
MS C4	TAGCGAGTGAATAGCGTC	
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 present ChrR
Maglo_CBS7966_v2_Chr5_213001_213600_Reverse:	AGGCACACGTTCATCTGGTT	
Maglo_CBS7966_v2_Chr6_461001_461600_Forward:	TGCTCACCCAAAAGACGACC	Probe for present Chr5
Maglo_CBS7966_v2_Chr6_461001_461600_Reverse:	CGCGGACCTGGAACTGTATT	

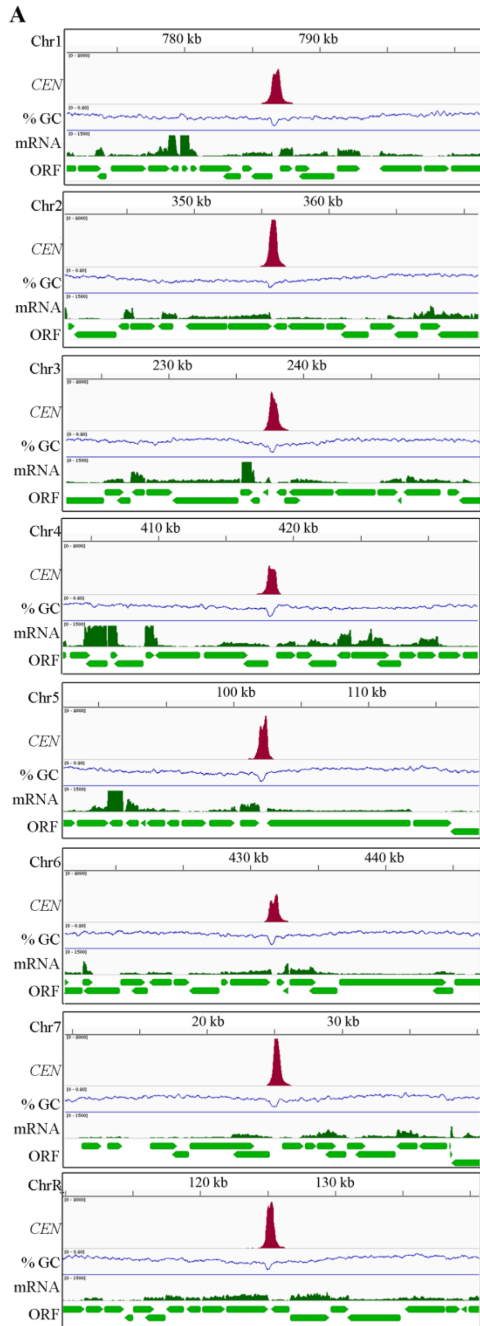
Supplementary table S8. List of software and algorithms used in this study

Software and Algorithms		
Fiji	National Institute of Health	https:/fiji.sc/
Photoshop CS6	Adobe Systems	https:/www.adobe.com
Excel	Microsoft	https:/products.office.com/en-in/excel
Word	Microsoft	https:/products.office.com/en-in/word
Geneious 9.0	Biomatters Ltd.	http://www.geneious.com/
SyMap	(14)	
PhylloGibbs-MP	(15)	
Fast Statistical Alignment (FSA)	(16)	
Circos	(10)	
Satsuma	(11)	
Easyfig	(12)	

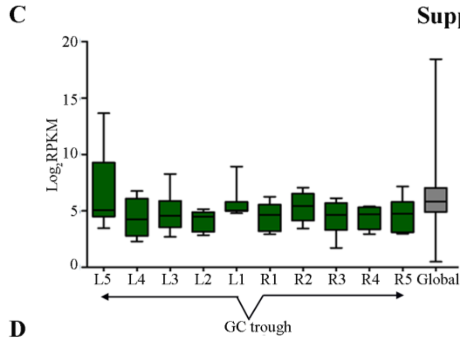
Supplementary figures

Supplementary figure S1

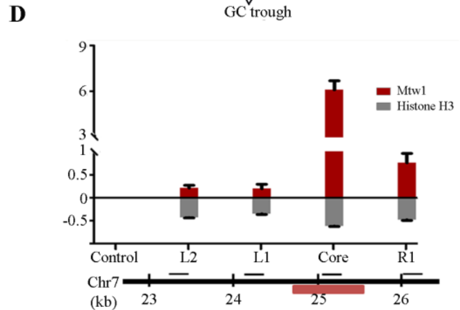
A



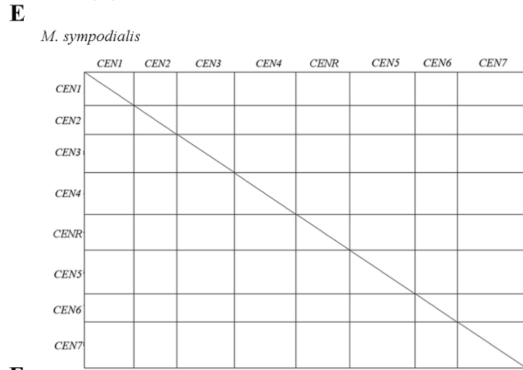
C



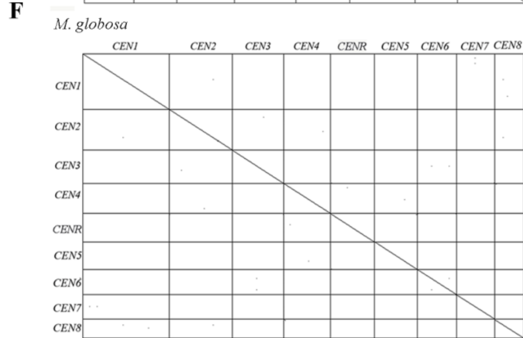
D



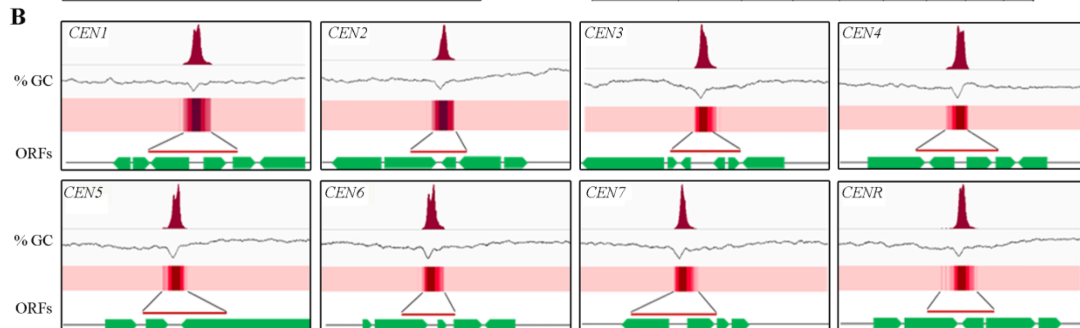
E



F

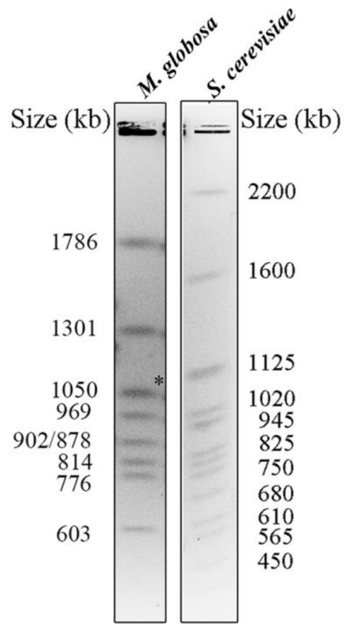


B

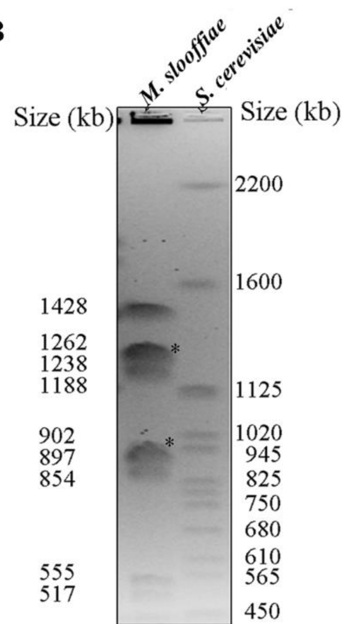


Supplementary figure S2

A



B



C

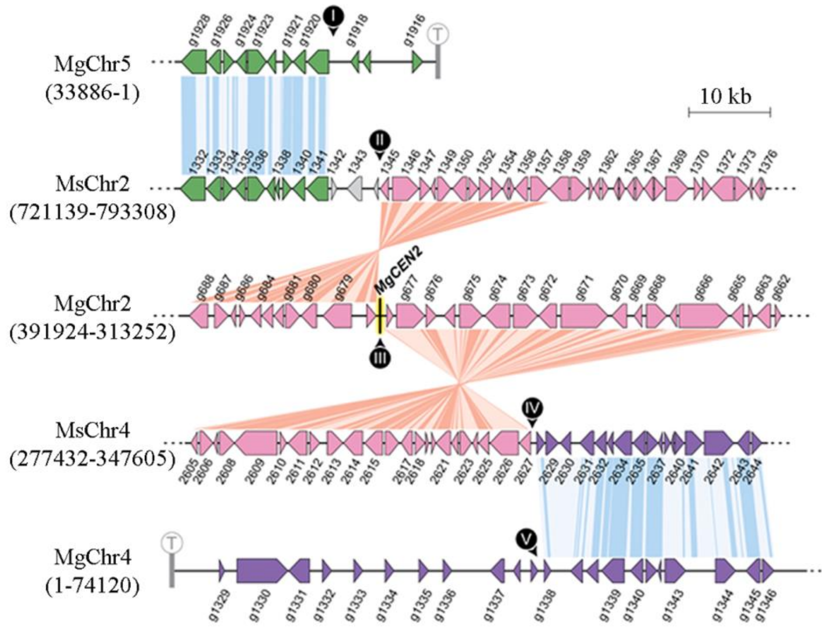
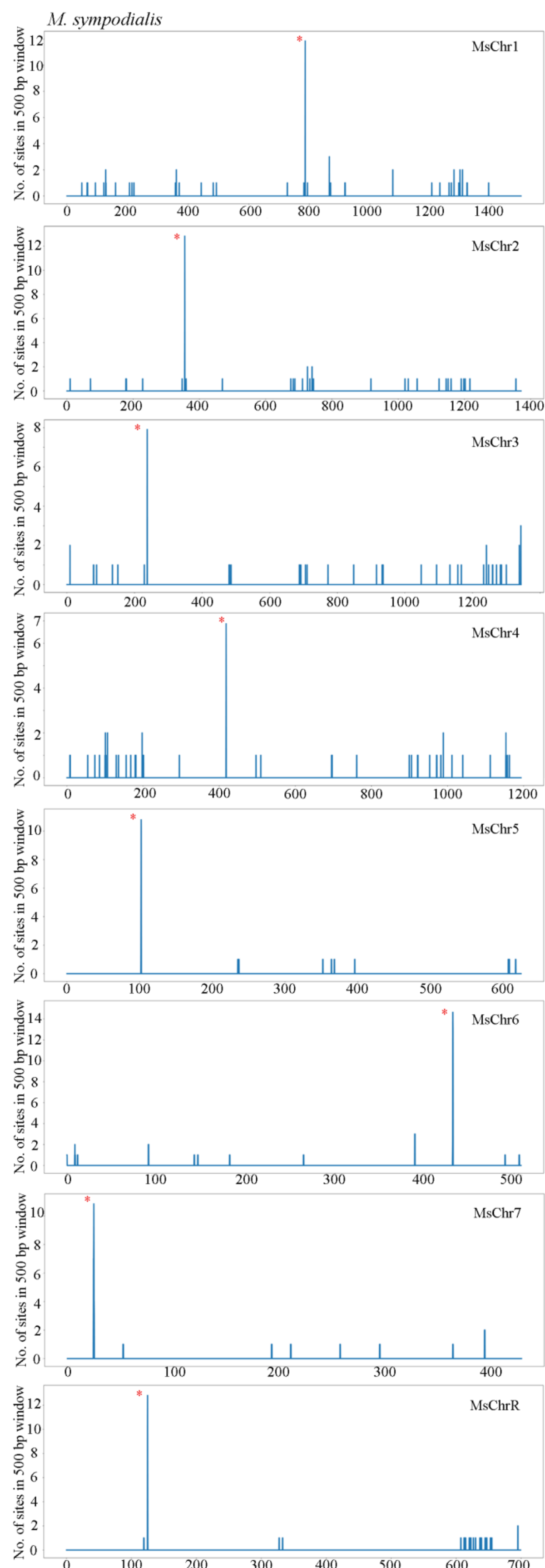


Figure S3

A



B

Figure S3

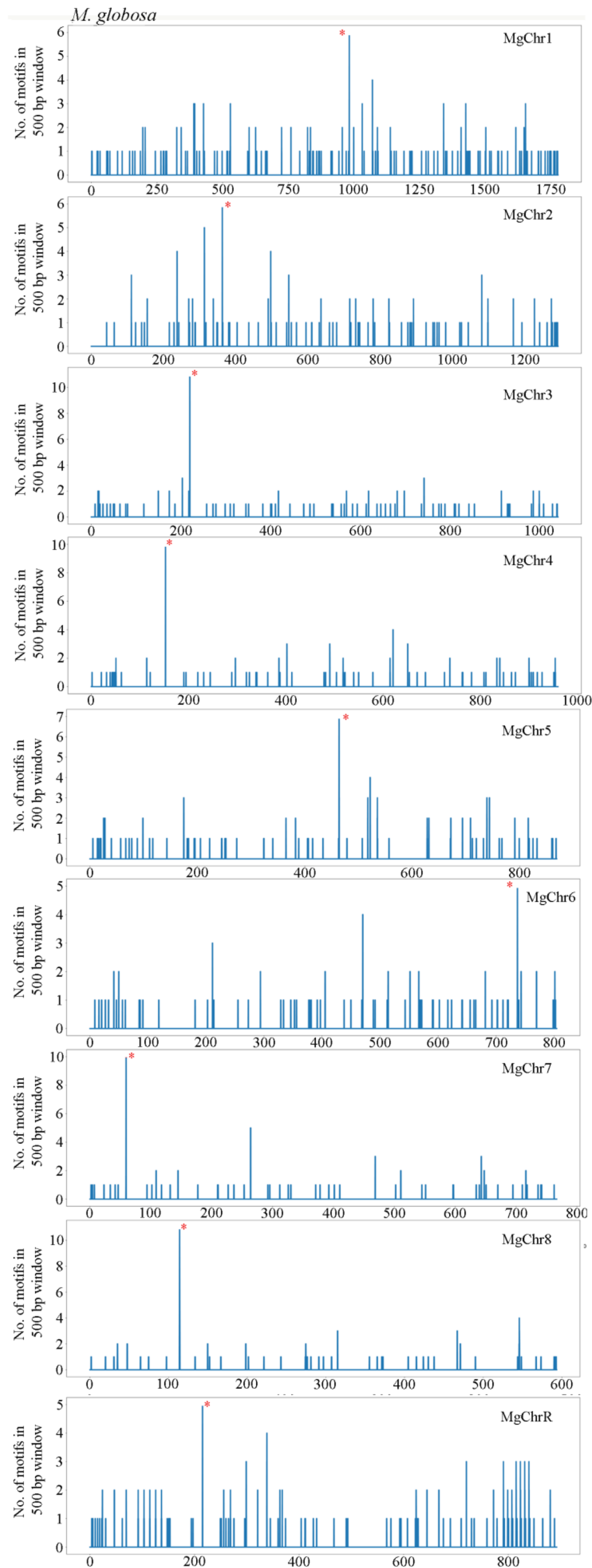
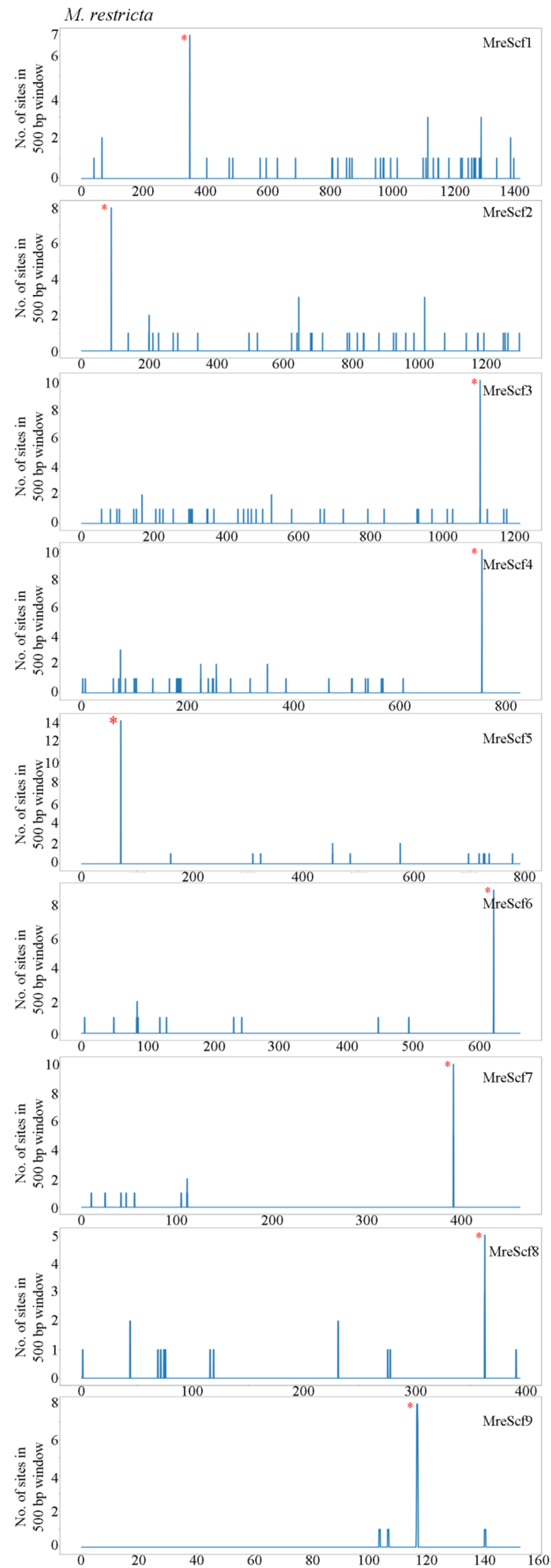


Figure S3



D

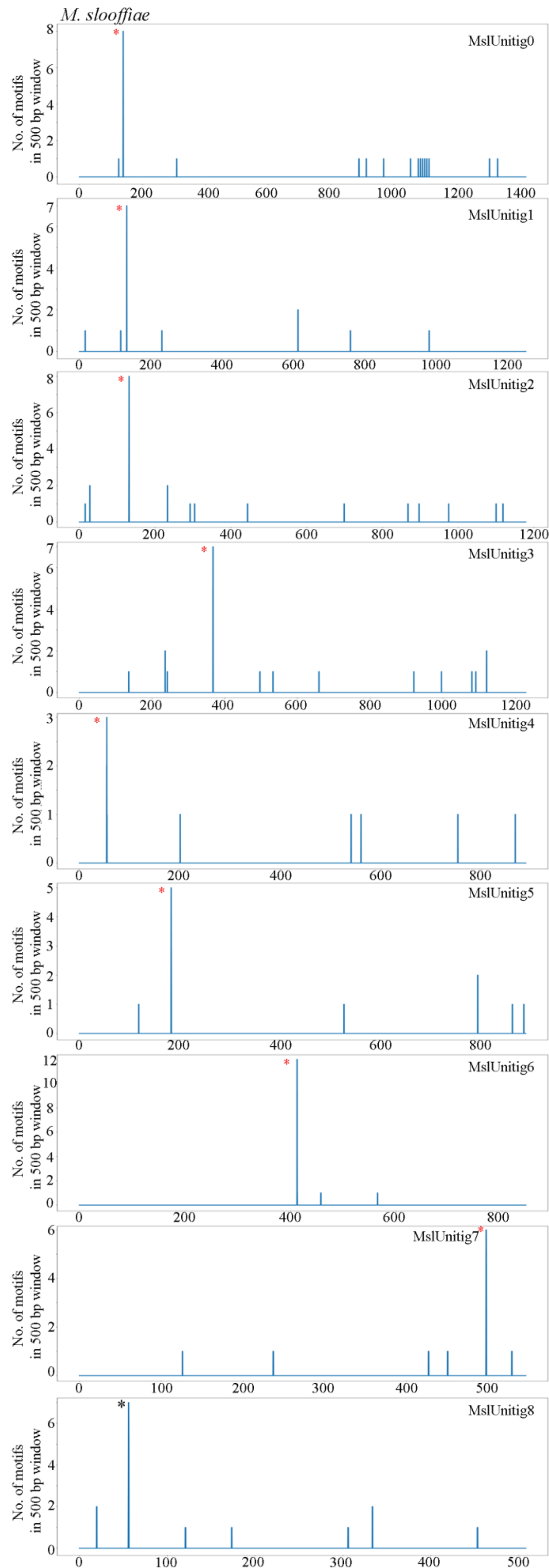
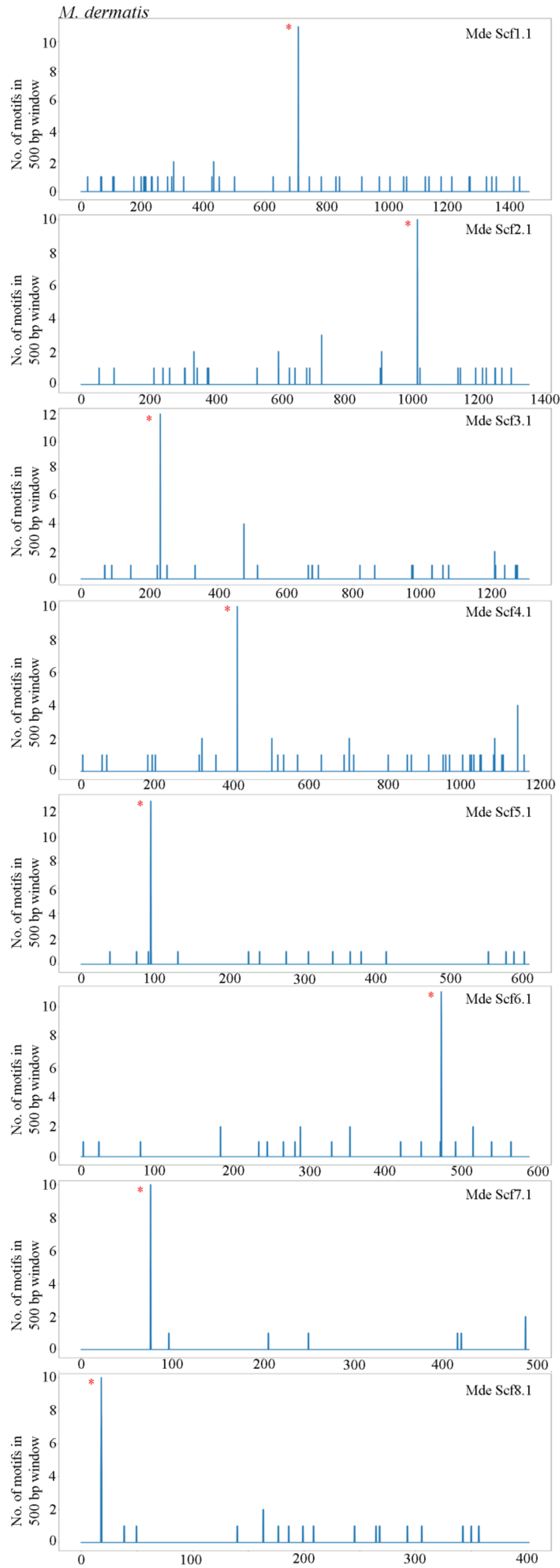


Figure S3

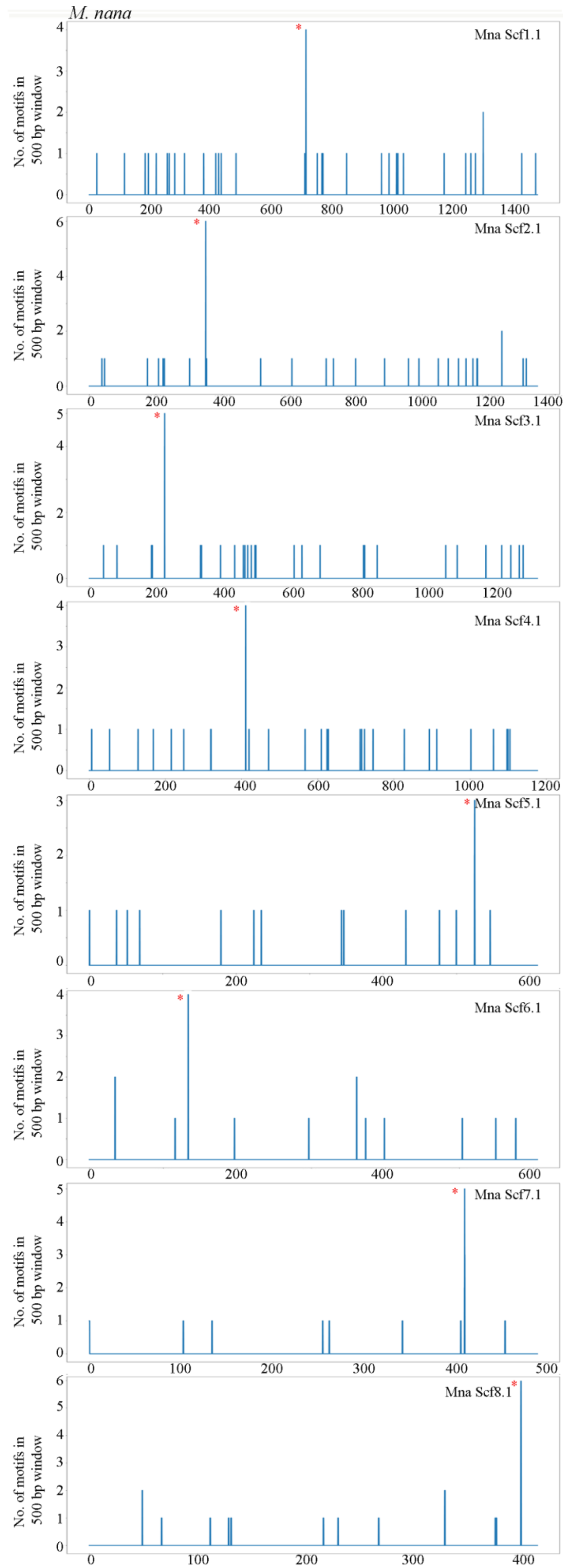
E

Figure S3

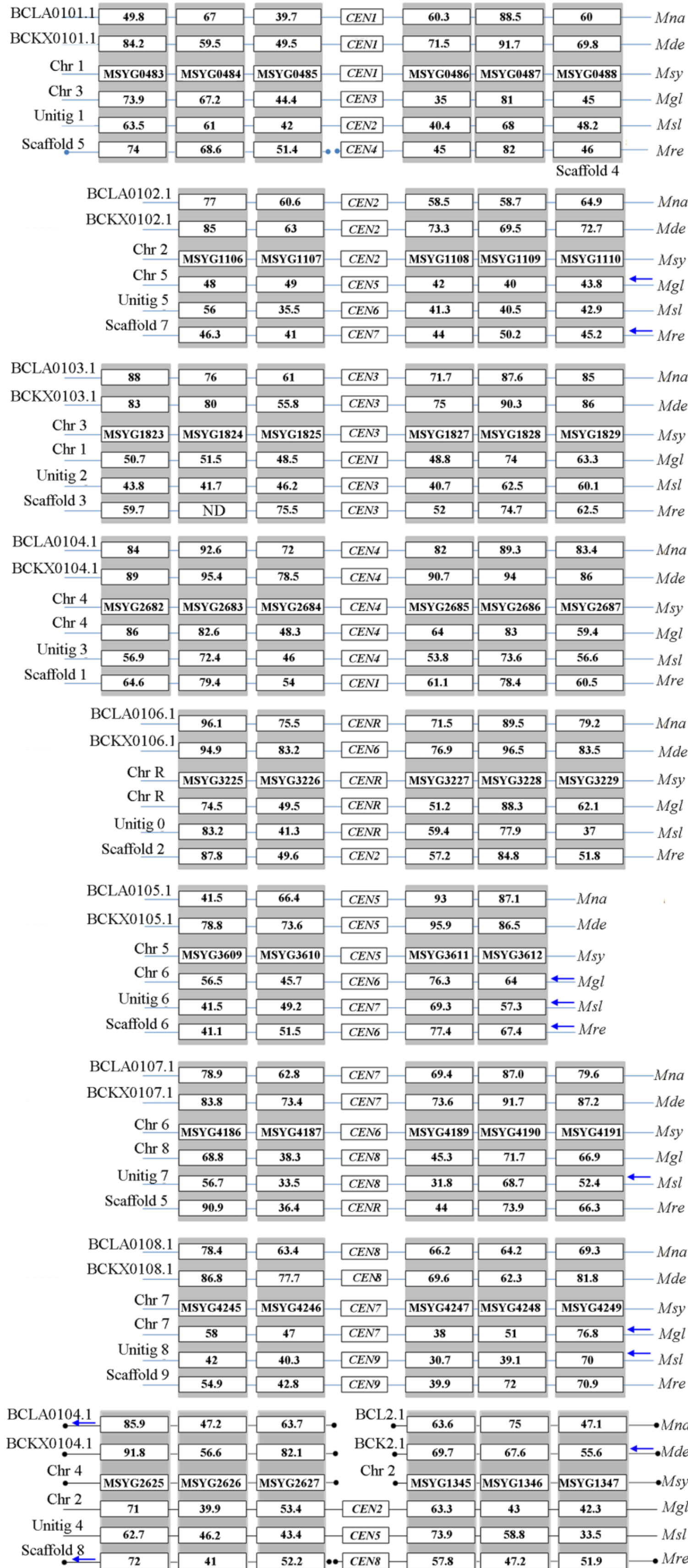


F

Figure S3



Supplementary figure S4



Supplementary references

1. Eckert-Boulet N, Pedersen ML, Krogh BO, & Lisby M (2012) Optimization of ordered plasmid assembly by gap repair in *Saccharomyces cerevisiae*. *Yeast* 29(8):323-334.
2. Ianiri G, Averette AF, Kingsbury JM, Heitman J, & Idnurm A (2016) Gene function analysis in the ubiquitous human commensal and pathogen *Malassezia* Genus. *mBio* 7(6).
3. Kozubowski L*, et al.* (2013) Ordered kinetochore assembly in the human-pathogenic basidiomycetous yeast *Cryptococcus neoformans*. *mBio* 4(5):e00614-00613.
4. Ianiri G, Applen Clancey S, Lee SC, & Heitman J (2017) FKBP12-dependent inhibition of calcineurin mediates immunosuppressive antifungal drug action in *Malassezia*. *mBio* 8(5).
5. Boekhout TaN, E.S., (2003) *Pulsed Field Gel Electrophoresis (PFGE) of Yeasts* pp pp. 93–103.
6. Yadav V*, et al.* (2018) RNAi is a critical determinant of centromere evolution in closely related fungi. *Proc Natl Acad Sci U S A* 115(12):3108-3113.
7. Sun S, Billmyre RB, Mieczkowski PA, & Heitman J (2014) Unisexual reproduction drives meiotic recombination and phenotypic and karyotypic plasticity in *Cryptococcus neoformans*. *PLoS Genet* 10(12):e1004849.
8. Sun S*, et al.* (2017) Fungal genome and mating system transitions facilitated by chromosomal translocations involving intercentromeric recombination. *PLoS Biol* 15(8):e2002527.
9. Findley K*, et al.* (2012) Discovery of a modified tetrapolar sexual cycle in *Cryptococcus amylolentus* and the evolution of MAT in the *Cryptococcus* species complex. *PLoS Genet* 8(2):e1002528.
10. Krzywinski M*, et al.* (2009) Circos: an information aesthetic for comparative genomics. *Genome Res* 19(9):1639-1645.
11. Grabherr MG*, et al.* (2010) Genome-wide synteny through highly sensitive sequence alignment: Satsuma. *Bioinformatics* 26(9):1145-1151.
12. Sullivan MJ, Petty NK, & Beatson SA (2011) Easyfig: a genome comparison visualizer. *Bioinformatics* 27(7):1009-1010.
13. Ianiri G, Boyce KJ, & Idnurm A (2017) Isolation of conditional mutations in genes essential for viability of *Cryptococcus neoformans*. *Curr Genet* 63(3):519-530.
14. Soderlund C, Bomhoff M, & Nelson WM (2011) SyMAP v3.4: a turnkey synteny system with application to plant genomes. *Nucleic Acids Res* 39(10):e68.
15. Siddharthan R (2008) PhyloGibbs-MP: module prediction and discriminative motif-finding by Gibbs sampling. *PLoS Comput Biol* 4(8):e1000156.
16. Bradley RK*, et al.* (2009) Fast statistical alignment. *PLoS Comput Biol* 5(5):e1000392.