Inter-species conservation of organisation and function between non-homologous regional centromeres

Pin Tong[1]*, Alison L. Pidoux[1]*+, Nicholas R.T. Toda[1,3], Ryan Ard[1,4], Harald Berger[1,5], Manu Shukla[1], Jesus Torres-Garcia[1], Carolin A. Mueller[2], Conrad A. Nieduszynski[2], Robin C. Allshire[1]

1. Wellcome Centre *for* Cell Biology and Institute of Cell Biology,
School of Biological Sciences,
The University of Edinburgh,
Mayfield Road,
Edinburgh EH9 3BF, UK.

2. Sir William Dunn School of Pathology,
University of Oxford,
South Parks Road,
Oxford OX1 3RE, UK.

* These authors made an equal contribution

+ Co-corresponding authors:
alison.pidoux@ed.ac.uk
robin.allshire@ed.ac.uk

Present addresses:

3. UPMC CNRS, Roscoff Marine Station, Place Georges Teissier, 29680 Roscoff, France

4. Copenhagen Plant Science Centre, University of Copenhagen, Bülowsvej 34, 1870 Frederiksberg C, Denmark

5. Symbiocyte, Universität für Bodenkultur Wien, University of Natural Resources and Life Sciences, Vienna, Austria

Despite the conserved essential function of centromeres, centromeric DNA itself is not conserved[1-4]. The histone-H3 variant, CENP-A, is the epigenetic mark that specifies centromere identity[5-8]. Paradoxically, CENP-A normally assembles on particular sequences at specific genomic locations. To gain insight into the specification of complex centromeres we took an evolutionary approach, fully assembling genomes and centromeres of related fission yeasts. Centromere domain organization, but not sequence, is conserved between *Schizosaccharomyces pombe*, *S. octosporus* and *S. cryophilus* with a central CENP-A^{Cnp1} domain flanked by heterochromatic outer-repeat regions. Conserved syntenic clusters of tRNA genes and 5S rRNA genes occur across the centromeres of *S. octosporus* and *S. cryophilus,* suggesting conserved function. Remarkably, non-homologous centromere central-core sequences from *S. octosporus* are recognized in *S. pombe*, resulting in cross-species establishment of CENP-A^{Cnp1} chromatin and functional kinetochores. Therefore, despite the lack of sequence conservation, *Schizosaccharomyces* centromere DNA possesses intrinsic conserved properties that promote assembly of CENP-A chromatin. Thus, centromere DNA can be recognized and function over unprecedented evolutionary timescales.

Centromeres are the chromosomal regions upon which kinetochores assemble to mediate accurate chromosome segregation. Evidence suggests that both genetic and epigenetic influences define centromere identity[1,2,4,7,9]. *S. pombe,* a paradigm for dissecting complex regional centromere function, has demarcated centromeres (35-110 kb) with a central domain assembled in CENP-A^{Cnp1} chromatin, flanked by outer-repeat elements assembled in RNAi-dependent heterochromatin, in which histone-H3 is methylated on lysine-9 (H3K9)[10-13]. Heterochromatin is required for establishment but not maintenance of CENP-A^{Cnp1} chromatin[6,14]. We have proposed that it is not the sequence *per se* of *S. pombe* central-core that is key in its ability to establish CENP-A chromatin, but the properties programmed by it[15]. To investigate whether these properties are conserved we have determined the centromere sequences of other *Schizosaccharomyces* species and tested their cross-species functionality.

Long-read (PacBio) sequencing permitted complete assembly of the genomes across centromeres of *S. octosporus* (11.9 Mb) and *S. cryophilus* (12.0 Mb), extending genome sequences[16] to telomeric or subtelomeric repeats or rDNA arrays (**Supplementary Figs. 1-3, Supplementary Tables 1,2**). Consistent with their closer evolutionary relationship[16,17], *S. octosporus* and *S. cryophilus* (32 My separation, compared to 119 My separation from *S. pombe*) exhibit greatest synteny (**Fig. 1a**). Synteny is preserved adjacent to centromeres (**Fig. 1b**). Circos plots indicate a chromosome arm translocation occurred within two ancestral centromeres to generate *S. cryophilus cen2* (*S.cry-cen2*) and *S.cry-cen3* relative to *S. octosporus* and *S. pombe* (**Fig. 1b**). Despite centromere-adjacent synteny, *Schizosaccharomyces* centromeres lack detectable sequence homology (see below). All centromeres contain a central domain: central-core (*cnt*) surrounded by inverted repeat (*imr*) elements unique to each centromere (**Fig. 2, Supplementary Fig. 4, Supplementary Tables 3-6**). CENP-A^{Cnp1} localises to fission yeast centromeres (**Fig. 2a**) and ChIP-Seq indicates that central domains are assembled in

CENP-A^{Cnp1} chromatin, flanked by various outer-repeat elements assembled in H3K9me2-heterochromatin (**Fig. 2b,c**). Despite the lack of sequence conservation, *S. octosporus* and *S. cryophilus* centromere organisation is strongly conserved with that of *S. pombe*, having CENP-A^{Cnp1}-assembled central domains separated by clusters of tRNA genes from outer-repeats assembled in heterochromatin[10,11] (**Supplementary Fig. 5, Supplementary Tables 7,8**). In contrast, our analyses of partially-assembled, transposon-rich centromeres of *S. japonicus* reveals the presence of heterochromatin on all classes of transposons and CENP-A on only two classes (**Supplementary Fig. 5, Supplementary Table 9**)[16].

Numerous 5S rRNA genes are located in the heterochromatic outer-repeats of *S. octosporus* and *S. cryophilus* centromeres (but not *S. pombe*) (**Fig. 1a, Supplementary Tables 10,11**). Almost all (25/26; 20/20) are within Five-S-Associated Repeats (FSARs; 0.6-4.2 kb) (**Fig. 3a**), encompassing ~35% of outer-repeat regions. FSARs exhibit 90% intra-class homology (**Supplementary Table 12**), but no interspecies homology. The three types of FSAR repeats almost always occur together, in the same order and orientation, but vary in copy number: *S. octosporus*: (oFSAR-1)$_1$(oFSAR-2)$_{1-9}$(oFSAR-3)$_1$; *S. cryophilus*: (cFSAR-1)$_{1-3}$(cFSAR-2)$_{1-2}$(cFSAR-3)$_1$. Both sides of *S. octosporus* and *S. cryophilus* centromeres contain at least one FSAR-1-2-3 array, except the right side of *S.cry-cen2* with two lone cFSAR-3 elements (**Fig. 3a, Supplementary Fig. 4**). *S. cryophilus* cFSAR-2 and cFSAR-3 repeats share ~400 bp homology (88% identity), constituting *hsp16* heat-shock protein ORFs (**Fig. 3a,b, Supplementary Table 13**) that are intact, implying functionality, selection and expression in some situations. Phylogenetic gene trees indicate that cFSAR-3-*hsp16* genes are more closely related with each other than with those in subtelomeric regions or cFSAR-2s (**Fig. 3b**), consistent with repeat homogenisation[18-20]. cFSAR-1s contain an eroded ORF with homology to a small hypothetical protein and *S. octosporus* oFSAR-2s contain a region of homology with a family of membrane proteins (**Fig. 3a**). The functions of centromere-associated *hsp16* genes and other ORF-homologous regions remain to be explored.

S. cryophilus heterochromatic outer-repeats contain additional repetitive elements, including a 6.2 kb element (cTAR-14) with homology to the retrotransposon *Tcry1* and transposon remnants at the mating-type locus[16] (**Figs. 1a,2b**, **Supplementary Fig. 4** and **Supplementary Tables 3,4,14**). *Tcry1* is located in the chrIII-R subtelomeric region (**Supplementary Figs. 3,4, Supplementary Table 1**). Although no retrotransposons have been identified in *S. octosporus*, remnants are present in the mating-type locus and oTAR-14ex in *S.oct-cen3* outer-repeats (**Fig. 2c**, **Supplementary Figs. 1,4** and **Supplementary Tables 5,6,15**). Hence, transposon remnants, FSARs and other repeats are assembled in heterochromatin at *S. octosporus* and *S. cryophilus* centromeres and potentially mediate heterochromatin nucleation.

tDNA clusters occur at transitions between CENP-A and heterochromatin domains in two of three centromeres in *S. octosporus* (*S.oct-cen2, S.oct-cen3*) and *S. cryophilus* (*S.cry-cen1, S.cry-cen2*), and are associated with low levels of both H3K9me2 and CENP-A^{Cnp1} (**Fig. 2b,c**), suggesting that they may

act as boundaries, as in *S. pombe*[21-23]. No tDNAs demarcate the CENP-A/heterochromatin transition at *S.cry-cen3.* Instead, this transition coincides precisely with 270-bp LTRs (**Fig. 2b, Supplementary Tables 3,4,14**), which may also act as boundaries[24-26]. Like tDNAs, LTRs are regions of low nucleosome occupancy, which may counter spreading of heterochromatin[26,27]. In addition, tDNA clusters occur near the extremities of all centromeres in both species, separating heterochromatin from adjacent euchromatin. tDNAs and LTRs are thus likely to act as chromatin boundaries at fission yeast centromeres.

A high proportion (~32%) of tRNA genes in *S. pombe*, *S. octosporus,* and *S. cryophilus* genomes are located within centromere regions[28] (**Figs. 1a,3c; Supplementary Tables 16-18**). Centromeric tDNAs are intact and are conserved in sequence with their genome-wide counterparts, indicating that they are functional genes. Two major, conserved tDNA clusters reside exclusively within *S. octosporus* and *S. cryophilus* centromeres (p-value<0.00001; q-value<0.05) (**Fig. 3c,d**). Cluster1 comprises several sub-clusters of 2-3 tDNAs in various combinations of up to 8 tDNAs, whilst Cluster2 contains up to 5 tDNAs (**Fig. 3d**); 17 different tDNAs (14 amino-acids) are represented, none of which are unique to centromeres (**Fig. 3c**). Intriguingly, the order and orientation of tDNAs within clusters is conserved between species, but intervening sequence is not (**Fig. 3d,e**). Strikingly, as well as local tDNA cluster conservation, inspection of centromere maps reveals synteny of tDNAs and clusters across large portions of *S. octosporus* and *S. cryophilus* centromeres. For example, the tDNA order AIR-RKL-E-T-T-L-DVAIR-RKLEF-A-DV (single-letter code) is observed at *S.oct-cen1* and *S.cry-cen3* (**Supplementary Fig. 6**). This synteny, together with both possessing small central-cores and long *imrs* suggests that these two centromeres are ancestrally related (**Fig. 3f**). Similarly, at *S.oct-cen3* and *S.cry-cen2*, tDNAs occur in the order NME-DV-AIRKE-EKRIA-VD-EMN-RIAVD, and at *S.oct-cen2* and *S.cry-cen1* the same tDNAs are present in the *imr* repeats and beyond (FELK-KL-E-DV). Central-cores have similar sizes and structures in the two species, each containing long (oCNT-L(6.4 kb); cCNT-L(6.0 kb)) and short (oCNT-S(1.2 kb); cCNT-S(1.3 kb)) species-specific repeats (**Fig.3f**, **Supplementary Tables 3-6,19**). CNT-repeats are arranged head-to-tail at one centromere and head-to-head at the other centromere in each species. Together, these similarities suggest ancestral relationships between *S.oct-cen2* and *S.cry-cen1*, *So-cen3* and *Scry-cen2*. Further, in places where synteny appears to break down, patterns of tDNA clusters suggest specific centromeric rearrangements occurred between the species. For instance, tDNA clusters at the edges of *S.cry-cen2R* and *S.cry-cen3L* are consistent with an inter-centromere arm translocation relative to *S.oct-cen1R* and *S.oct-cen2R*, indicated by gene synteny maps (**Figs. 1b, 4a** and **Supplementary Fig. 6**).

No central-core sequence homology was revealed between species using BLASTN. To identify potential underlying centromere sequence features, k-mer frequencies (5-mers), normalized for centromeric AT-bias, were used in Principal Component Analysis. CENP-A-associated regions of all three genomes group together, distinct from the majority of non-centromere sequences (p-value, 9.3×10^{-7}) (**Fig. 4b,c**). Interestingly, *S. pombe* neocentromere-forming regions[29] also cluster separately from other genomic regions, sharing sequence features with centromeres.

K-mer analysis and conserved centromeric organisation prompted us to investigate cross-species functionality of protein and DNA components of *Schizosaccharomyces* centromeres. GFP-tagged CENP-A^{Cnp1} protein from each species localised to *S. pombe* centromeres and complemented the *cnp1-1* mutant[30] (**Fig. 5a-c**), indicating that heterologous CENP-A proteins assemble and function at *S. pombe* centromeres, despite normally assembling on non-homologous sequences in their respective organisms.

Introduction of *S. pombe* central-core (*S.pom-cnt*) DNA on minichromosomes into *S. pombe* results in the establishment and maintenance of CENP-A chromatin if *S.pom-cnt* is adjacent to heterochromatin, or if CENP-A is overexpressed[6,14,15,31]. *S.oct-cnt* regions (3.2-10 kb) or *S.pom-cnt2* (positive control) were placed adjacent to *S. pombe* outer-repeat DNA in mini-chromosome constructs (**Fig. 6a**) which were transformed into *S. pombe* cells expressing wild-type levels (wt-CENP-A) or overexpressing *S. pombe* GFP-CENP-A^{Cnp1} (hi-CENP-A^{Cnp1})[15]. Acquisition of centromere function is indicated by minichromosome retention on non-selective indicator plates (white/pale-pink colonies), and by the appearance of sectored colonies (**Fig. 6b,c**). The pHET-*S.pom-cnt2* minichromosome containing *S.pom-cnt2* established centromere function at high frequency immediately upon transformation in hi-CENP-A^{Cnp1} cells (90%) and at lower frequency in wt-CENP-A cells (15%; not shown). Centromere function was established on *S.oct-cnt*-containing minichromosomes in hi-CENP-A cells only (**Fig. 6d**). Centromere function was not due to minichromosomes gaining portions of *S. pombe* central-core DNA (data not shown). CENP-A^{Cnp1} ChIP-qPCR indicated that, for minichromosomes with established centromere function, CENP-A^{Cnp1} chromatin was assembled on non-homologous *S.oct-cnt* DNA, to levels similar to those at endogenous *S. pombe* centromeres and to *S.pom-cnt2* on a minichromosome (**Fig. 6e**). Minichromosomes containing *S.oct-cnt* provided efficient segregation function (**Fig. 6d**), no longer requiring CENP-A^{Cnp1} overexpression to maintain that function once established (**Fig. 6f**), consistent with the self-propagating ability of CENP-A chromatin[5,15]. These analyses indicate that *S.oct-cnt* is competent to establish CENP-A chromatin and centromere function in *S. pombe* when CENP-A^{Cnp1} is overexpressed, suggesting that *S. octosporus* central-core DNA has intrinsic properties that promote the establishment of CENP-A chromatin despite lacking sequence homology.

Based on conserved features, ancestral *Schizosaccharomyces* centromeres may have consisted of a CENP-A^{Cnp1}-assembled central-core surrounded by tDNA clusters and 5S rDNAs. We surmise that RNAPIII promoters perhaps provided targets for transposon integration[32], followed by heterochromatin formation to silence retrotransposons and preserve genome integrity[33,34]. The ability of heterochromatin to recruit cohesin[35,36], benefitting chromosome segregation selected for heterochromatin maintenance[37,38], rather than underlying sequence which evolved by repeat expansion and continuous homogenisation[18-20]. Because tDNAs performed important functions – as boundaries preventing heterochromatin spread into central-cores and perhaps in higher order centromere organisation and architecture – tDNA clusters were maintained[21]. In *S. pombe*, non-centromeric and centromeric tDNAs and 5S rDNAs cluster adjacent to centromeres in a TFIIIC-dependent manner[22,23]. The multiple tandem

centromeric 5S rDNAs and tDNAs could contribute to a robust, highly-folded heterochromatin structure promoting optimum kinetochore configuration for co-ordinated microtubule attachments and accurate chromosome segregation[38].

The lack of overt sequence conservation between centromeres of different species appears not to prevent functional conservation, which may be driven by underlying sequence features or properties such as the transcriptional landscape. Although maintenance of centromere function has been observed at a pre-established human centromere in chicken cells[39] (310 My divergence), CENP-A establishment on human alpha-satellite in mouse cells[40] (90 My divergence) is surpassed by the competence of *S. octosporus* central-core DNA to establish CENP-A chromatin in *S. pombe* from which it is separated by 119 My of evolution[16] (equivalent to 383 My using a chordate molecular clock). Thus, our analyses extend the evolutionary timescale over which cross-species establishment of CENP-A chromatin has been demonstrated.

Methods

Cell growth and manipulation

Standard genetic and molecular techniques were followed. Fission yeast methods were as described[41]. Strains used in this study are listed in **Supplementary Table 20**. All *Schizosaccharomyces* strains were grown at 32°C in YES, except *S. cryophilus* which was grown at 25°C unless otherwise stated. *S. pombe* cells carrying minichromosomes were grown in PMG-ade-ura. For low GFP-tagged CENP-A^{Cnp1} protein expression from episomal plasmids, cells were grown in PMG-leu with thiamine.

PacBio sequencing of genomic DNA

High molecular weight genomic DNA was prepared from *S. cryophilus*, *S. octosporus* and *S. japonicus* using a Qiagen Blood and Cell Culture DNA Kit (Qiagen), according to manufacturer's instructions. Pacific Biosciences (PacBio) sequencing was carried out at the CSHL Cancer Center Next Generation Genomics Shared Resource. Samples were prepared following the standard 20 kb PacBio protocol. Briefly: 10-20 μg of genomic material was sheared via g-tube (Covaris) to 20 kb. Samples were damage repaired via ExoVII (PacBio), damage repair mix and end repair mix using standard PacBio 20 kb protocol. Repaired DNA underwent blunt-end ligation to add SMRTbell adapters. For some libraries: 10-50 kb molecules from 1-2 μg SMRTbell libraries were size selected using BluePippin (Sage Science) after which samples were annealed to Pacbio SMRTbell primers per the standard PacBio 20 kb protocol. Annealed samples were sequenced on the PacBio RSII instrument with P4/C3 chemistry. Magbead loading was used to load each sample at a concentration between 50 to 200 pM. Additional PacBio sequencing (without BluePippin) was performed by Biomedical Research Core Facilities, University of Michigan. There, the following kits were used: DNA Sequencing Kit XL 1.0, DNA Template Prep Kit 2.0 (3Kb - 10Kb)" and DNA/Polymerase Binding Kit P4. MagBead Standard Seq v2 sequencing was performed using 10,000 bp size bin with no Stage Start with a 2 hour observation time on a PacBio RSII sequencer. A summary of PacBio sequencing performed is listed in **Supplementary Table 21**.

De novo whole genome assembly of PacBio sequence reads

PacBio reads were assembled using HGAP3 (The Hierarchical Genome Assembly Process version 3)[42]. Reads were first sorted by length, and the top 30% used as seed reads by HGAP3. All remaining reads of at least 1 kb in length were used to polish the seed reads. These polished reads were used to *de novo* assemble the genomes and Quiver software used to generate consensus genome contigs. Comparisons to the ChIP-seq input data and Broad Institute *Schizosaccharomyces* reference genomes[16] showed very high agreement with these datasets.

The *S. octosporus* and *S. cryophilus* chromosomes were named according to their sequence lengths, the longest chromosome being labelled as chromosome I in each case.

De novo assembly of the _S. pombe_ genome using nanopore technology

Genomic DNA was extracted as described previously[43]. DNA purity and concentration were assessed using a Nanodrop 2000 and the double-stranded high sensitivity assay on a Qubit fluorometer, respectively. Genomic DNA was sequenced using the MinION nanopore sequencer (Oxford Nanopore Technologies). Three sequencing libraries were generated using the 1D ligation kit SQK-LSK108, the 2D ligation kit SQK-NSK007 and the 1D Rapid sequencing kit SQK-RAD002, following manufacturers guidelines. Each library was sequenced on one MinION flow cell. Sequencing reads were base-called using Metrichor (1D and 2D ligation libraries) or Albacore (rapid sequencing library). The combined dataset incorporating reads from three flow cells was assembled using Canu v1.5. The assembly was computed using default Canu parameters and a genome size of 13.8 Mbp. QUAST v3.2 was used to evaluate the genome assembly.

Genome annotation and chromosome structure

Genes were annotated onto the genome both *de novo,* using BLAST and the sequences of known genes, and by using liftover (https://genome-store.ucsc.edu) to carry over the previous gene annotation information from the Broad institute reference genomes (ref). CrossMap[44] was then used to lift the chain files over to the new, updated genome. The locations of tDNAs were predicted using tRNAscan[45,46]. Dfam 2.0[47] was used to annotate repetitive DNA elements. MUMmer3.23[48] was used to compare the genomes and annotate repeat elements and tandem repeat sequences, including those located in centromeric domain and telomere sequences. Centromeric repeat elements were manually identified using BLASTN and MEGABLAST (https://blast.ncbi.nlm.nih.gov). Each repeat element was named according to their sequence features (association with tDNA & rDNAs) and locations. The sequence of the wild-type (h^{90}) *S. pombe* mating-type locus was obtained by manually merging nanopore and PacBio contigs using available data[16], Supplementary Figure 10 and information at *www.pombase.org/status/mating-type-region*. Genome synteny alignment analysis was carried out using syMAP42[49,50], based on orthologous genes among the three genomes.

ChIP-qPCR

For analysis of CENP-A^{Cnp1} association with minichromosomes bearing *S. octosporus* central core DNA, three independent transformants with established centromere function (indicated by ability to form sectored colonies) for each minichromosome were grown in PMG-ade-ura cultures and fixed with 3.7% formaldehyde for 15 min at room temperature. Cells were lysed by bead-beating (Biospec) and ChIP was performed as previously described[51]. 10 μl anti-CENP-A^{Cnp1} sheep antiserum and 25 μl Protein-G-Agarose beads (Roche) were used per ChIP. qPCR was performed

using a LightCycler 480 and reagents (Roche) and analysed using Light Cycler 480 Software 1.5 (Roche). Primers used in qPCR are listed in **Supplementary Table 22**. Mean %IP ChIP values for *Sp-cnt* or *So-cnt* on minichromsomes were normalised to %IP for endogenous *S. pombe cnt1*. Error bars represent standard deviation.

ChIP-seq

A modified ChIP protocol was used. Briefly, pellets containing 7.5 x 10^8 cells were lysed by four 1 min pulses of bead beating in 500 µl of lysis buffer (50 mM HEPES-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate), with resting on ice in between. The insoluble chromatin fraction was pelleted by centrifugation at 6000 *g* and washed with 1 ml lysis buffer before resuspension in 300 µl lysis buffer containing 0.2% SDS. Chromatin was sheared by sonication using a Bioruptor (Diagenode) for 30 minutes (30 s on/off, high setting). 900 µl of lysis buffer (no SDS) was added and samples clarified by centrifugation at 17000 *g* for 20 minutes and the supernatant used for ChIP. 6 µl anti-H3K9me2 mouse monoclonal mAb5.1.1[52] (kind gift from Takeshi Urano) or 30 µl sheep anti-CENP-A^{Cnp1} antiserum[51] were used, along with protein G-dynabeads (ThermoFisher Scientific) or Protein-G agarose beads (Roche), respectively. (For neocentromere strains, cells were first treated with Zymolyase 100T, washed in sorbitol and permeablized. Chromatin was fragmented with incubation with micrococcal nuclease. Cell suspensions were adjusted to standard ChIP buffer conditions and extracted chromatin was processed as per standard ChIP.) Immunoprecipitated DNA was recovered using Qiagen PCR purification columns. ChIP-Seq libraries were prepared with 1-5 ng of ChIP or 10 ng of input DNA. DNA was end-repaired using NEB Quick blunting kit (E1201L). The blunt, phosphorylated ends were treated with Klenow-exo$^-$ (NEB, M0212S) and dATP. After ligation of NEXTflex adapters (Bioo Scientific) DNA was PCR amplified with Illumina primers for 12-15 cycles and library fragments of ~300 bp (insert plus adaptor sequences) were selected using Ampure XP beads (Beckman Coulter). The libraries were sequenced following Illumina HiSeq2000 work flow (or as indicated in **Supplementary Table 21**).

Defining fission yeast centromeres

CENP-A^{Cnp1} and H3K9me2 ChIP-seq data was generated to identify centromere regions. ChIP-Seq reads with mapping qualities lower than 30, or read pairs that were over 500-nt or less than 100-nt apart, were discarded. ChIP-seq data was normalized with respect to input data. Paired-end ChIP-seq data (single-end for *S. japonicus*) was aligned to the updated genome sequences using Bowtie2[53]. Samtools[54], Deeptools[55] and IGV [56] were subsequently used to generate sequence data coverage files and to visualize the data. MACS2[57] was used to detect CENP-A^{Cnp1} and heterochromatin-enriched regions of the genome.

Centromere tDNA cluster analysis

To test for the enrichment of tDNA clusters at centromere regions a greedy search approach was used to identify potential clusters. All tDNAs less than 1000 bp apart were grouped into clusters. To test for significant clustering of tDNAs at the centromere the locations of tDNAs across the genome were shuffled 1000 times. For each cluster observed in the real genome the proportion of permutations where the same cluster was observed at least as many times was calculated to provide estimates of significance. Following conversion of these p-values to q values to account for multiple testing, the centromere tDNA clusters each exhibited a q-value less than 0.005.

Hsp16 gene tree analysis

hsp16 paralogs from *S. octosporus* and *S. cryophilus* genomes were predicted using BLASTP. The predicted protein sequences from *hsp16* genes across all four fission yeasts were aligned together with those from *S. cerevisiae* using Clustal Omega. BEAST (Bayesian Evolutionary Analysis Sampling Trees)[58] and FigTree (http://tree.bio.ed.ac.uk/software/figtree/) was used to generate and view the *hsp16* gene phylogenetic tree.

5-mer frequency PCA analysis

The CENP-A^{Cnp1}-associated sequences in the *S. pombe*, *S. cryophilus* and *S. octosporus* genomes are all approximately 12 kb in length. Each genome was therefore split into 12 kb sliding windows with a 4.5 kb overlap. The frequencies of each 5-mer was calculated in each window using Jellyfish[59]. CENP-A^{Cnp1}-associated regions showed a general enrichment of AT base pairs relative to the genome as a whole. To normalize for GC content amongst the windows, all base pairs were randomized in each sequence window to generate 1000 artificial sequences with the same GC content. 5-mer frequencies were then recalculated for each of these 1000 artificial sequences and the true original 5-mer frequencies compared to these background frequencies by calculating a z-score. Consequently, these enrichment scores represent the k-mer enrichments in a given sequence normalized for GC content. Genome windows were split into 6 groups: CENP-A^{Cnp1}-associated sequences (CENP-A^{Cnp1} peaks covering more than 6 kb of sequence); outer repeat heterochromatin regions (more than half the window covered by H3K9me2 peaks adjacent to CENP-A domains); sub-telomeric regions (more than half the window covered by H3K9me peaks and close to the end of a chromosome); Mating-type locus, neo-centromere regions (identified using CENP-A^{Cnp1} ChIP-seq data on *S. pombe* neocentromere-containing strains[29]) and remaining genome sequences. Logistic regression and mean comparison were used to determine whether principal components were linked to the probability of a sequence belonging to a particular sequence group[60]. Logistic regression and mean comparison were used to determine whether principal components (FactoMineR) were linked to the probability of a sequence belonging to a particular sequence group.

Construction of minichromosomes

Regions of *S. octosporus* central core regions were amplified with primers indicated in **Supplementary Table 22**. Fragments were digested with *Bgl*II, *Nco*I or *Bam*HI, *Nco*I and ligated into *Bgl*II-*Nco*I-digested plasmid pK(5.6kb)-MCS-ΔBam which contains a 5.6 kb fragment of the *S. pombe* K (*dg*) outer repeat. To create plasmid pK-So-cnt2-10kb, an additional 3.6 kb region from *S.oct-cnt2* was inserted as a *Bam*HI-*Sal*I fragment into *Bgl*II-*Sal*I-digested pK-So-cnt2-6.5kb to make a 10 kb region of *S. octosporus* central core. For pKp plasmids, *S. octosporus* central core regions were by inserted as *Bgl*II-*Nco*I or S*al*I-*Bam*HI fragments into *Sal*I-*Bam*HI or *Nco*I-*Bam*HI digested plasmid pKp (pMC91) which contains 2 kb region from *S. pombe* K(*dg*) outer repeat. Plasmids are listed in **Supplementary Table 23**.

Centromere establishment assay

Strains A7373 or A7408, which contains integrated nmt41-GFP-CENP-A^{Cnp1} to allow high level expression of CENP-A[15],were grown in PMG-complete medium and transformed using sorbitol-electroporation method[61]. Cells were plated on PMG-uracil-adenine plates and incubated at 32^{o}C for 5-10 days until medium-sized colonies had grown. Colonies were replica-plated to PMG low adenine (10 μg/ml) plates to determine the frequency of establishment of centromere function. These indicator plates allow minichromosome loss (red) or retention (white/pale pink) to be determined. Minichromosome retention indicates that centromere function has been established and that minichromosomes segregate efficiently in mitosis. In the absence of centromere establishment, minichromosomes behave as episomes that are rapidly lost. Minichromosomes occasionally integrate giving a false positive white phenotype. To assess the frequency of such integration events and to confirm establishment of centromere segregation function, a proportion of colonies giving the white/pale-pink phenotype upon replica plating were re-streaked to single colonies on low-adenine plates – sectored colonies are indicative of segregation function with low levels of minichromosome loss, whereas pure white colonies are indicative of integration into endogenous chromosomes – and the establishment frequency adjusted accordingly.

Minichromosome stability assay

Minichromosome loss frequency was determined by half-sector assay. Briefly, transformants containing minichromsomes with established centromere function were grown in PMG-ade-ura to select for cells containing the minichromosome. Two transformants were analysed per minichromosome (four for pK-So-cnt2-4.7kb). Cells were plated on low-adenine containing plates and allowed to grow non-selectively for 4-7 days. Minichromosome loss is indicated by red sectors and retention by white sectors. To determine loss rate per division, all colonies were examined with a dissecting microscope. All colonies – except pure reds – were counted to give total number of colonies. Pure reds were checked for the absence of white sectors and were excluded because

they had lost the minichromosome before plating. To determine colonies that lost the minichromosome in the first division after plating, 'half-sectored' colonies were counted. This included any colony that was 50% or greater red (including those with only a tiny white sector). Loss rate per division is calculated as the number of half-sectored colonies as a percentage of all (non-pure-red) colonies.

Immunolocalisation

For localisation of CENP-A^{Cnp1}, *Schizosaccharomyces* cultures were fixed with 3.7% formaldehyde for 7 min, before processing for immunofluorescence as described [51]. Anti-CENP-A^{Cnp1} sheep antiserum[51] (raised to the N-terminal 19 amino-acids of *S. pombe* CENP-A^{Cnp1}) was used at 1:1000 dilution, and Alexa-488-coupled donkey-anti-sheep secondary antibody (A11015; Invitrogen) at 1:1000 dilution. Cells were stained with DAPI and mounted in Vectashield. Microscopy was performed with a Zeiss Imaging 2 microscope (Zeiss) using a 100x 1.4NA Plan-Apochromat objective, Prior filter wheel, illumination by HBO100 mercury bulb. Image acquisition with a Photometrics Prime sCMOS camera (Photometrics, https://www.photometrics.com) was controlled using Metamorph software (Universal Imaging Corporation). Exposures were 1500 ms for FITC/Alexa-488 channel and 300-1000 ms for DAPI. Images shown in Figure 2a are autoscaled.

To express GFP-tagged versions of *Schizosaccharomyces* CENP-A^{Cnp1} proteins in *S. pombe*, ORFs were amplified from relevant genomic DNA using primers listed in **Supplementary Table 22**. Fragments were digested with *Nde*I-*Bam*HI or *Nde*I-*Bgl*II and ligated into *Nde*I-*Bam*HI digested pREP41X-GFP vector[62] (**Supplementary Table 23**). For detection of GFP-tagged versions of *Schizosaccharomyces* CENP-A^{Cnp1} proteins in *S. pombe*, cells containing pREP41X-GFP-CENP-A^{Cnp1} episomal plasmids (variable copy number) were grown in PMG-leu + thiamine to allow low GFP-CENP-A^{Cnp1} expression. Cells were fixed, processed for immunolocalisation and microscopy as above. Anti-GFP antibody (A11122; Invitrogen) was used at 1:300, anti-Cdc11[51] (a spindle-pole body marker; gift from Ken Sawin) was used at 1:600. Secondary antibodies were, respectively, Alexa-488 coupled chicken anti-rabbit (A21441; Invitrogen) and Alexa-594 coupled donkey anti-sheep (A11016; Invitrogen) both at 1:1000. Exposures were FITC/488 channel: 1500 ms, TRITC/594 1000 ms, DAPI 500-1000 ms. For display of images in Figure 5C, TRITC/594 and FITC/488 images are scaled relative to the maximum intensity in the set of images, whilst DAPI images are autoscaled.

Data Availability

All data generated in this study have been submitted to GEO under accession number: GSE112454. SRA submission number for *S. pombe* nanopore sequencing data: SUB3761672. The following figures have associated raw data: 1, 2, S1, S2, S3, S5.

Acknowledgments

We thank Alastair Kerr, Shaun Webb and Daniel Robertson for bioinformatics support, David Kelly for microscopy support, and Ken Sawin and Takeshi Urano for antibodies, and Kojiro Ishii, Ken Sawin and Nick Rhind for yeast strains. We thank Robert Lyons, Joe Washburn, Christina McHenry (University of Michigan) and Greg J. Hannon, Richard McCombie, Eric Antoniou and Sara Goodwin (CSHL) for PacBio sequencing. We are grateful to Chris Ponting for advice and comments on the manuscript and Sandra Catania and other members of the Allshire and Heun labs for helpful discussions. N.R.T.T., R.A. and J.T-G. were supported by the Darwin Trust of Edinburgh. The Darwin Trust and a Principal's Career Development scholarship supported N.R.T.T. P.T. was partly supported by funding from the European Commission Network of Excellence EpiGeneSys- (HEALTH-F4-2010-257082) and a Wellcome Enhancement Award (095021) to R.C.A. R.C.A. is a Wellcome Principal Research Fellow (095021, 200885); the Wellcome Centre for Cell Biology is supported by core funding from Wellcome (203149). C.A.M. and C.A.N. are supported by Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/N016858/1 and Wellcome Investigator Award 110064/Z/15/Z.Pacific Biosciences (PacBio) sequencing carried out at the CSHL Cancer Center Next Generation Genomics Shared Resource, which is supported by the Cancer Center Support Grant 5P30CA045508 was paid for by a kind gift from Kathryn W. Davis to GJH.

Author Contributions

R.C.A. and A.L.P. designed the study. P.T. performed the PacBio genome assemblies and bioinformatics, ChIP-seq analysis and PCA analysis. C.M. performed the nanopore sequencing of *S. pombe* supervised by C.N. H.B., N.R.T.T. J.T.-G. and R.A. generated ChIP-seq data with contribution from M.S. A.L.P. performed cytology, analysis of repetitive regions, and experiments on cross-species functionality. R.C.A. supervised the study. A.L.P. wrote the manuscript with contributions from P.T., R.C.A. and other authors. All authors read and approved the final version of the manuscript.

Competing Financial Interests

The authors declare no competing financial interests.

Figure Legends

Figure 1: Genome organisation and synteny in *Schizosaccharomyces*

a) Circos plots depicting pairwise *S. pombe, S. octosporus* and *S. cryophilus* genome synteny. Rings from outside to inside represent: chromosomes; GC content (high: red, low: yellow); 5S rDNAs (red); tDNAs (black); LTRs (green); CENP-A^{Cnp1} ChIP-seq (purple); H3K9me2 ChIP-seq (orange); innermost ring and coloured connectors indicate regions of synteny between species. *S. pombe* chromosomes are indicated by blue (*S.pom-chr1*), green (*S.pom-chr2*), red (*S.pom-chr3*) in the left and right panels and regions of synteny on *S. octosporus* and *S. cryophilus* chromosomes, respectively, are indicated in corresponding colours. A similar designation is used for *S. octosporus* chromosomes in the middle panel.

b) Circos plot isolating regions adjacent to centromeres highlighting preserved synteny and an intra-centromeric chromosome arm swap involving *S. cryophilus cen2* and *cen3* relative to *S. pombe* and *S. octosporus.*

Figure 2: Domain organisation of *Schizosaccharomyces* centromeres

a) Immunostaining of centromeres in indicated *Schizosaccharomyces* species with anti-CENP-A^{Cnp1} antibody (green) and DNA staining (DAPI; red). Scale bar, 5 μm.

(b) *S. cryophilus* centromere organisation indicating DNA repeat elements. ChIP-seq profiles for CENP-A^{Cnp1} (purple) and H3K9me2-heterochromatin (orange) are shown above each centromere. Positions of tDNAs (single-letter code of cognate amino acid; black), 5S-rDNAs (red), and solo LTRs are indicated (pink). Central cores (cnt - purples) inner-most repeats (imr – blue shades). 5S-associated repeats (cFSARs – orange shades); tDNA-associated repeats (TARs) containing clusters of tDNAs (green shades); heterochromatic repeats (cHR) and TARs associated with single tDNAs (various colours: brown/pink/red). cTAR-14s, containing retrotransposon remnants (deep pink). For details, including individual repeat annotation, see **Supplementary Fig. 4** and **Supplementary Tables 3,4**.

(c) *S. octosporus* centromere organisation indicating DNA repeat elements. Labelling and shading as in (b). Only oTAR-14ex (pale pink part) contain retrotransposon remnants. Colouring is indicative of homology within each species but only of possible repeat equivalence (not homology) between species; see **Supplementary Table 5,6,19**.

Figure 3: *S. cryophilus* and *S. octosporus* contain conserved clusters of tDNAs and similar non-homologous repeat elements

(a) Schematic of *S. cryophilus* and *S. octosporus* FSAR repeats, indicating positions of 5S-rDNAs, *hsp16* genes and other ORFs. Copy number of each FSAR within centromeric arrays is indicated.

(b) Phylogenetic relationship of *S. cryophilus* centromeric *hsp16* genes with genomic *hsp16* and *hsp20* genes of *S. cryophilus*, *S. octosporus*, *S. pombe* and *S. japonicus*

(c) Heat map of tDNA frequency at centromeric and non-centromeric sites (blue shades) for *S. pombe*, *S. cryophilus*, and *S. octosporus*. Anticodons and cognate amino acids indicated right (purple: present at centromeres). Clusters containing these tDNAs indicated. Histogram (*top*): total tDNA frequencies in centromeres and non-centromeric sites of indicated species. Histogram (*left*): tDNA frequencies in each species.

(d) Depiction of centromeric tDNA clusters and sub-clusters. Combinations of 2 or 3 tDNAs subclusters present in both species (purple) or specific to *S. octosporus* (red) *or S. cryophilus* (blue) of are indicated (single-letter code of cognate amino-acid; arrows indicate plus or minus strand).

(e) *Top:* Dot-plot alignment (MEGABLAST) showing synteny between oTAR-4/oTAR-5 (DVAIR-Cluster 1) from *S.oct-cen1R* (chr1:3355194-3357165) with oTAR-4/oTAR-5 (DVAIR-Cluster 1) from *S.cry-cen3R* (chr3:964707-966623). *Bottom:* Dot-plot of oTAR-4/oTAR-5 (DVAIR-Cluster 1) from *S.oct-cen1R* (chr1:3355194-3357165) and oTAR-4/oTAR-5 (DVAIR-Cluster 1) from *S.oct-cen3L* (chr3:1791072-1793051).

(f) Schematic of central domain similarity between species. Central cores (purple shades), *imr* (blues), TARs containing tDNA clusters (greens). Long (*CNT-L*) and short (*CNT-S*) central core repeats are indicated. tDNAs indicated in single-letter amino acid code. Colours highlight similarity of organisation between species and indicates homology within, not between, species.

Figure 4: *Schizosaccharomyces* centromeres share ancestry and sequence features

(a) Structural alignment of putatively equivalent centromere repeat elements of *S. cryophilus* and *S. octosporus* to highlight potential centromere rearrangements during evolution

(b) Principal Component Analysis PC1 and PC2 of 5-mer frequencies of three fission yeast genomes. Genome regions (12 kb window) were assigned to one of 5 specific annotated groups (CENP-A^{Cnp1}-associated (purple), centromeric heterochromatin (orange), mating-type locus (blue), subtelomeres (yellow), neocentromere-forming regions[29] (red), or other genome regions (grey). For each group the oval line encloses 95% of the data points.

(c) Boxplot Principal Component PC1 of each group. Colours as in b. Mean comparison between groups was used (p-value: >0.05, ns; >0.01, *; >0.001, **; >0.0001, ***; <0.0001, ****)[60]. Centre line, medium; box limits, upper and lower quartiles; whiskers, 1.5 x interquartile range; points, outliers.

Figure 5: Cross-species functionality of CENP-A^{Cnp1} proteins

(a) Alignment of *Schizosaccharomyces* CENP-A^{Cnp1} proteins. Positions of alpha helices (yellow), N-terminal tail (green) and CENP-A-targeting domain (CATD; red) are indicated.

(b) *S. pombe* temperature sensitive *cnp1-1* cells expressing plasmid-borne GFP-CENP-A^{Cnp1} from the indicated species (*Sp, S. pombe*; *So, S. octosporus*; *Sc, S. cryophilus*; *Sj, S. japonicus*), or GFP alone, were spotted on phloxine B-containing plates and incubated for 2-5 days at the indicated temperatures.

(c) Localisation of GFP-tagged CENP-A^{Cnp1} from indicated *Schizosaccharomyces* species in *S. pombe*. Wild-type *S. pombe* cells bearing plasmids described in (a) were grown at 32°C before fixation and staining with anti-GFP (green), anti-Cdc11 (red, spindle-pole body) and DAPI (blue, DNA). Centromeres cluster at the spindle-pole body in *S. pombe*. Scale bar, 5 μm.

Figure 6: *S. octosporus* central core DNA establishes CENP-A^{Cnp1} chromatin upon introduction into *S. pombe*

(a) Indicated regions of *S. octosporus* central core DNA placed adjacent to a portion of *S. pombe* heterochromatin-forming outer repeat sequence on a plasmid.

(b) *S. pombe* transformants containing minichromsome plasmids were replica-plated to low adenine non-selective plates: colonies retaining the chimeric minichromosome plasmid are white/pale-pink, those that lose it are red. Representative plate showing pKp-So-cnt3-6.5kb-containing colonies.

(c) *S. pombe* cells containing pKp-So-cnt3-6.5kb chimeric minichromosome were streaked to single colonies. Red colour indicates loss of minichromosome; small red sectors indicate low frequency minichromosome loss and mitotic segregation function.

(d) Establishment frequency of chimeric minichromosomes in *S. pombe* hi-CENP-A^{Cnp1} cells. Establishment frequency determined by replica plating of transformants (Methods) as shown in b (n=number of transformants analysed). Chromosome loss rate of established minichromosomes was determined by half-sector assay (Methods). Two transformants containing established centromeres were analysed for each minichromosome and the mean loss rate determined, n=number of colonies screened.

(e) ChIP-qPCR for CENP-A^{Cnp1} on *S. pombe* hi-CENP-A^{Cnp1} cells containing chimeric minichromosomes with established centromere function. Three independent transfomants were analysed for each minichromosome. ChIP enrichment on *S.pom-cnt2* and *S.oct-cnt*-bearing minichromosomes is normalised to the level at endogenous *S. pombe cnt1*. Error bars, standard deviation.

(f) Propagation of chimeric minichromosome stability. Cells containing pK(5.6kb)-So-cnt2-10kb were streaked on low adenine-containing plates with or without thiamine which results in repression or expression of high levels of *S. pombe* CENP-A^{Cnp1}.

Supplementary Figure Legends

Figure S1: *S. octosporus* and *S. cryophilus* genome assembly statistics

a) Histograms of SMRT cell subread lengths (green) and the sum of subread length (black) for the indicated genomes.

b) Summary of PacBio subreads and final assemblies.

c) Dot plot comparison of new assemblies with previously published assemblies for *S. octosporus*, *S. cryophilus and S. japonicus*[16].

d) Organisation of mating-type loci in *S. pombe*, *S. octosporus* and *S. cryophilus*[16,22]. ChIP-seq profiles for H3K9me2-heterochromatin (orange) are shown. Positions of mating-type loci (blue) and mating-type genes (white), mating-type associated repeat elements (H1: dark red; H2: red; H3: pink; abc: yellow); transposon remnants (pink), inverted IR repeats (grey) and other genes (black) are indicated. *cenH* region (orange) homologous to *S. pombe* centromeric *dg/dh* repeats is indicated. Blue shading indicates homologous genes between species.

Figure S2: Structure of *S. octosporus* chromosome ends

a) Overview of *S. octosporus* chromosomes, indicating organisation of subtelomeres. Multiple copies of terminal repeats (black arrows; GGGTTACTT) are detected at the end of chr1L (and internally). Combinations of subtelomeric repeats, including telomere-associated sequences (oTAS, dark red); RecQ type DNA helicase genes (*tlh*) and associated repeats (oTLH-R) and other subtelomeric repeats (oSTR-4 etc; turquoise); details in (b). A partial atypical rDNA repeat is detected at: chr1R, 2R and 3R (light green arrow). Due to repetitive nature of this region, assemblies are incomplete at all chromosome ends (denoted by grey star), except chr1L.

b) Details of terminal 100 kb of chr1L, chr2L and chr3L. Two copies of telomere-associated sequences (oTAS; dark red), and oTLH-R (containing RecQ type DNA helicase genes (*tlh*)) are present at chr1L, along with multiple copies of GGGTTACTT repeats. Numerous other subtelomeric repeats, designated oSTR-4 etc (blue/turquoise) are indicated, mostly by number designation only due to space constraints.

c) Top, structure of atypical rDNA repeat unit (lacking the full NTS seen in typical rDNA repeats, see (d). Atypical rDNA repeat units are present at centromere-proximal side of chromosome ends: chr1R, 2R, 3R. Due to repetitive nature of these regions, full assembly was not achieved and the number of rDNA repeat units present at each chromosome end is unknown. From ChIP input read counts the total number of rDNA repeat units is estimated to be approximately 150 copies.

d) Homology of rDNA repeat unit between *Schizosaccharomyces* species.

S. pombe, *S. cryophilus* and *S. octosporus* rDNA repeats are shown. *S. pombe* elements were previously defined[63]. Homology indicated by grey blocks: darker grey indicates higher homology (65%-92%).

Figure S3: Structure of *S. cryophilus* chromosome ends

a) *Left*, overview of *S. cryophilus* chromosomes, indicating organisation of subtelomeres. Multiple copies of terminal repeats (black arrows; GGGTTACTT) are present at the ends of 1R, 2L and 3R, along with combinations of subtelomeric repeats, including telomere-associated sequences (cTAS, red); RecQ type DNA helicase genes (*tlh*) and associated repeats (cTLH-R) and other subtelomeric repeats (cSTR-4 etc; shades of pink/brown); details in c. rDNA repeats are located at: 1L, 2R and 3L. Centromere-proximal rDNA repeat is atypical (light green arrow) and associated with a 5S rDNA (details in b). Distal to that, assemblies of chromosome 1 and 3 indicate a partial standard rDNA repeat (no associated 5S rDNA; dark green). Due to repetitive nature of this region, assemblies are incomplete at 1L, 2R and 3L (denoted by grey star). *Middle*, two classes of terminal rDNA-containing contigs were also identified. Both types contain multiple copies of the terminal GGGTTACTT repeat. In class A these directly abut rDNA repeat units. In class B, cTAS and cTLH-R elements are located between the terminal repeats and rDNA repeat unit. A similar arrangement has recently been described for the rDNA-containing ends of *S. pombe* chromosome 3[64]. *Right*, key to repeat elements.

b) Two types of rDNA repeat are present in *S. cryophilus*. *Top*, standard repeat of 11.1 kb, present in tandem arrays. 28S, 18S and 5.8S genes are indicated (green), along with putative associated elements external transcribed spacers (5'ETS, turquoise; 3'ETS, yellow) and non-transcribed spacer (NTS, teal). *Bottom*, atypical rDNA repeat located at centromere-proximal location of all three rDNA-containing chromosome ends. In place of standard NTS element it is associated with a 7.5 kb repeat (pale blue) containing a 5S rDNA gene (red arrow).

c) *Upper 3 panels:* Subtelomere-repeat-containing chromosome ends are assembled in heterochromatin. Terminal 100 kb of arms 1R, 2L and 3R are shown. Location of repeat elements are indicated, along with positions of *hsp16* genes. Smaller repeat elements are indicated by vertical bars (see key, bottom right). H3K9me2 ChIP-seq profile is shown (orange). cT-180 repeats (green bars) coincide with deep troughs in H3K9me2 reads, suggesting that these elements could perform a boundary function. cTLH-R contains intact *tlh* genes, whereas cSTR-4 elements contain degraded copies of *tlh*. The partial cSTR-4 element has weak homology to cSTR-4 at 2L and 3R and a highly degraded copy of *tlh*. Location of genes indicated in black. The non-heterochromatic portion of the subtelomeres contain several paralogous genes, homologues of which are also found in the subtelomeric regions of *S. octosporus*. The intact retrotransposon *Tcry1* (magenta; LTRs, pink arrows) is located in the subtelomeric region of 3R.

Lower 3 panels: Chromosome ends with rDNA repeat units: terminal 100 kb of 1L, 2R, 3L shown. Feature colours as in a, b. H3K9me2 ChIP-seq profile (orange) indicates enrichment over the non-transcribed spacer. Note that, as with all repetitive regions, numbers of ChIP-seq reads mapped represent an average over all repeats. Locations of 5S rDNAs, LTRs (pink arrow) and a partial *Tcry1* retrotransposon (magenta) are indicated. Assembly of full rDNA-containing chromosome ends was not possible due to its repetitive nature, consequently the number of rDNA repeat units

present at each chromosome end remains unknown. However, the total number of rDNA repeat units is estimated to be 150 from ChIP input read counts.

Figure S4: Centromere repeat organisation in *S. cryophilus* and *S. octosporus*

Structure and organisation of *S. cryophilus* (a) and *S. octosporus* (b) centromeres are shown. Location and names of centromeric repeats indicated. Repeat colour indicates identity/high degree of homology within species. Repeats with the same colour between species do not show sequence homology, but are present in similar contexts with respect to chromatin status, association with particular tDNAs or occurrence and location within centromeres. The only detectable homology is between tDNAs themselves, and between cTAR-14 (pink; all *S.cry-cens*) and extended oTAR-14-ex (*S.oct-chr3*) elements which have weak homology to the retrotransposon *Tcry1* and retransposon remnants at the mating-type loci of both species and at other locations in the genomes (see **Supplementary Tables 14,15,19**).

Central core regions are indicated in shades of purple, and positions of long (*cCNT-L* and *oCNT-L*) and short (*cCNT-S* and *oCNT-S*) repeats are indicated. Innermost (*imr*) inverted repeats are shown in shades of blue and in some cases contain smaller 'boundary type' repeat elements (greens). tDNAs are shown as vertical black bars and the cognate amino acid shown in single letter code below. Small repeat elements associated with clustered tDNAs which may have boundary function are shown in shades of green and turquoise (tDNA-associated repeats; TARs). 5S rDNAs are shown as vertical red bars. Heterochromatic Five-S-associated repeats (cFSARs and oFSARs) are shown in shades of orange (see **Figure 3a**). Longer repeats associated with single tDNAs (unlikely to have boundary function) are indicated in shades of red/brown/plum (TARs 11-14). Other heterochromatic repeats (HR) are indicated in shades of yellow/brown/pink. Genes flanking the centromeres are indicated in black.

(c) Retrotransposon remnants are present within *S. cryophilus* and *S. octosporus* centromeres and other genomic locations**.** *Left panel:* Dot-plot alignment (BLASTN) of *Tcry1* retrotransposon[16] with *S. cryophilus* chromosome 3. Pink triangle indicates position of *Tcry1* itself at subtelomeric region of chr3R. Purple triangle indicates position of *S.oct-cen3*. Homology lies within cTAR-14 elements (also present at *S. cryophilus cen1* and *cen3*). Other regions of homology on chromosome arms are due to partial copies of *Tcry1* retrotransposon. *Right panel:* Dot-plot alignment (BLASTN) of *Tcry1* retrotransposon with *S. octosporus* chromosome 3. Purple triangle indicates position of *S.oct-cen3*. Homology within centromere is with oTAR-14ex elements, present only at *S.oct-cen3*. Region of homology in subtelomere is ~135 kb from chromosome end.

Figure S5: Domain organisation of *S. pombe* centromeres and transposon-rich centromeres of *S. japonicus*

(a) *S. pombe* genome was assembled from Oxford nanopore sequencing. Updated *S. pombe* centromere organisation indicating repeat elements. ChIP-seq profiles for CENP-A^{Cnp1} (purple) and H3K9me2-heterochromatin (orange) are shown above each centromere. Positions of tDNAs (single-letter code of cognate amino acid; black) are indicated. Central cores are shown in shades of purple (TM element common to *cnt1* and *cnt3* shown in mauve); innermost repeats (*imr*) in blues, heterochromatic outer repeat elements are shown in grey (*dg*) and orange (*dh*).

(b) Representative H3K9me2 and CENP-A-associated *S. japonicus* contigs. ChIP-seq profiles for CENP-A^{Cnp1} (purple) and H3K9me2-heterochromatin (orange) are shown above each contig. Due to their repetitive nature, full assembly of centromere regions was not possible; association with CENP-A^{Cnp1} is strong indicator that a contig is centromere located. Retrotransposons mapping to contigs are indicated; colour-coding as previously published[16] (key, bottom), full-length or almost full-length retrotransposons are indicated by a black outline. An additional putative retrotransposon was identified which has been named Tj11 (see **Supplementary Table 9**). Positions of tDNA clusters (single-letter code of cognate amino acids; black) are indicated. Top/middle: Chromosome arm-sized contigs which terminate within retrotransposon arrays. Bottom: smaller contigs containing only retrotransposons and other repetitive elements could not be incorporated into genome assembly.

Figure S6: Synteny of tDNA clusters and organisation of repeat elements suggests common ancestry of *S. cryophilus* and *S. octosporus* centromeres, despite lack of sequence conservation

(a) Centromeres of *S. cryophilus* and *S. octosporus* paired with the most similar centromere from the opposite species. *In silico* rearrangements of *S. cryophilus* centromeres closely recapitulate the organisation of *S. octosporus* centromeres, providing support for their evolution from common ancestral centromeres. Dashed boxes and arrows indicate rearrangement that that would increase the structural similarity between *S.cry-cen1* and *S.oct-cen2*.

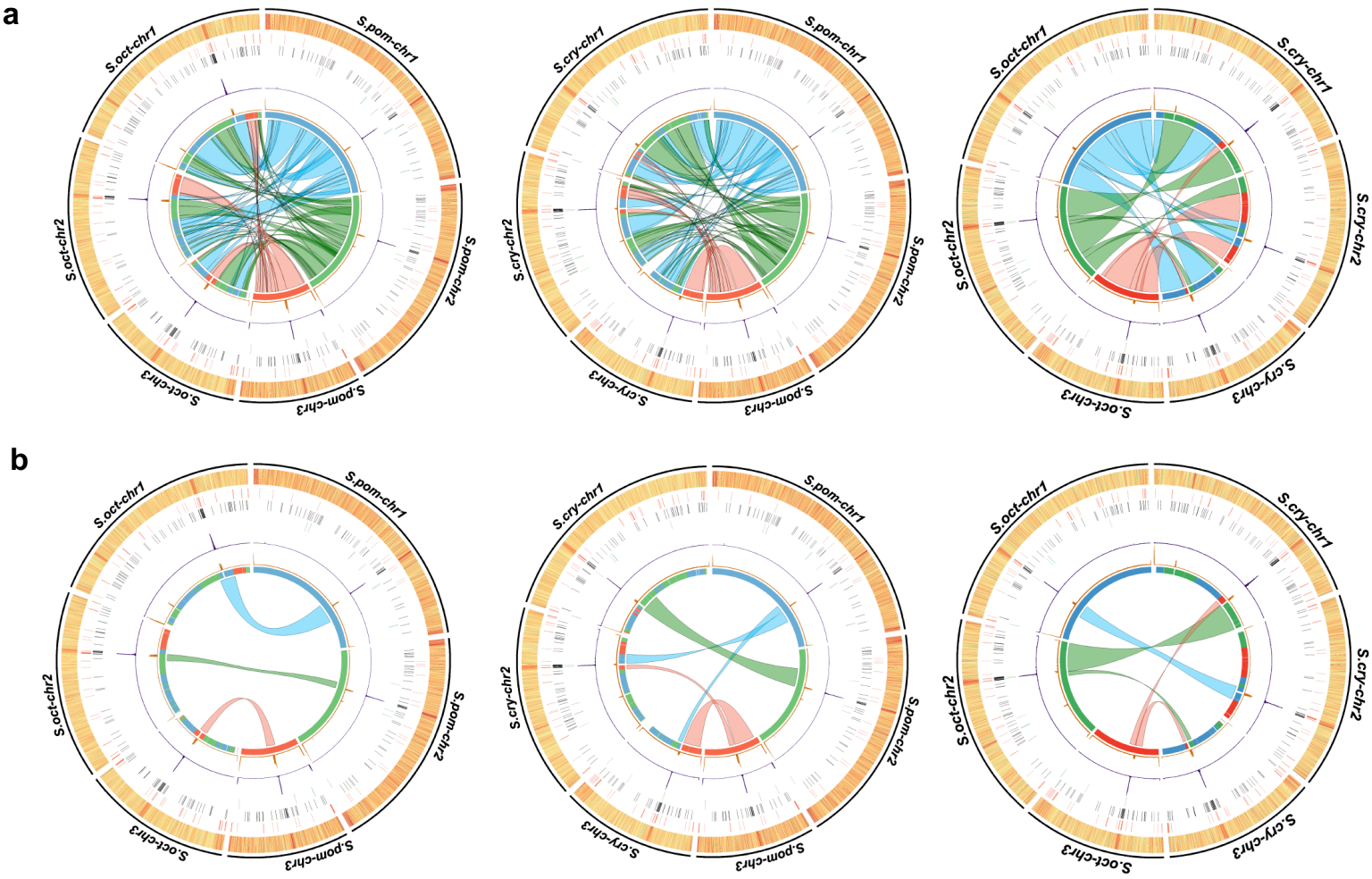
b and c) A rearrangement involving arm swap of *S.cry-cen2R* and *S.cry-cen3L* would produce synteny of genes on either side of *S.cry-cen2* and *S.cry-cen3*, and *Soct-cen3* and *Soct-cen1* respectively (centre, dashed double-headed arrow). Partial inversions (curved double-headed arrows) would increase similarity between *S. cryophilus* and *S. octosporus* centromeres (labelled compare). See **Figure 4a**.

References:

1. Buscaino, A., Allshire, R. & Pidoux, A. Building centromeres: home sweet home or a nomadic existence? *Curr Opin Genet Dev* **20,** 118–126 (2010).
2. Plohl, M., Mestrović, N. & Mravinac, B. Centromere identity from the DNA point of view. *Chromosoma* **123,** 313–325 (2014).
3. Malik, H. S. & Henikoff, S. Major evolutionary transitions in centromere complexity. *Cell* **138,** 1067–1082 (2009).
4. Dumont, M. & Fachinetti, D. DNA Sequences in Centromere Formation and Function. *Prog. Mol. Subcell. Biol.* **56,** 305–336 (2017).
5. Mendiburo, M. J., Padeken, J., Fülöp, S., Schepers, A. & Heun, P. Drosophila CENH3 is sufficient for centromere formation. *Science* **334,** 686–690 (2011).
6. Folco, H. D., Pidoux, A. L., Urano, T. & Allshire, R. C. Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. *Science* **319,** 94–97 (2008).
7. Gómez-Rodríguez, M. & Jansen, L. E. Basic properties of epigenetic systems: lessons from the centromere. *Curr Opin Genet Dev* **23,** 219–227 (2013).
8. McKinley, K. L. & Cheeseman, I. M. The molecular basis for centromere identity and function. *Nat Rev Mol Cell Biol* **17,** 16–29 (2015).
9. Karpen, G. H. & Allshire, R. C. The case for epigenetic effects on centromere identity and function. *Trends Genet* **13,** 489–496 (1997).
10. Partridge, J. F., Borgstrøm, B. & Allshire, R. C. Distinct protein interaction domains and protein spreading in a complex centromere. *Genes Dev* **14,** 783–791 (2000).
11. Allshire, R. C. & Ekwall, K. Epigenetic Regulation of Chromatin States in Schizosaccharomyces pombe. *Cold Spring Harb Perspect Biol* **7,** (2015).
12. Martienssen, R. & Moazed, D. RNAi and Heterochromatin Assembly.
13. Reyes-Turcu, F. E. & Grewal, S. I. Different means, same end-heterochromatin formation by RNAi and RNAi-independent RNA processing factors in fission yeast. - PubMed - NCBI. *Curr Opin Genet Dev* **22,** 156–163 (2012).
14. Kagansky, A. *et al.* Synthetic heterochromatin bypasses RNAi and centromeric repeats to establish functional centromeres. *Science* **324,** 1716–1719 (2009).
15. Catania, S., Pidoux, A. L. & Allshire, R. C. Sequence Features and Transcriptional Stalling within Centromere DNA Promote Establishment of CENP-A Chromatin. *PLoS Genet* **11,** e1004986 (2015).
16. Rhind, N. *et al.* Comparative functional genomics of the fission yeasts. *Science* **332,** 930–936 (2011).
17. Helston, R. M., Box, J. A., Tang, W. & Baumann, P. Schizosaccharomyces cryophilus sp. nov., a new species of fission yeast. *FEMS Yeast Res* **10,** 779–786 (2010).
18. Dawe, R. K. & Henikoff, S. Centromeres put epigenetics in the driver's seat. *Trends Biochem Sci* **31,** 662–669 (2006).
19. Roizès, G. Human centromeric alphoid domains are periodically homogenized so that they vary substantially between homologues. Mechanism and implications for centromere functioning. *Nucleic Acids Res* **34,** 1912–1924 (2006).
20. Sharma, A., Wolfgruber, T. K. & Presting, G. G. Tandem repeats derived from centromeric retrotransposons. *BMC Genomics* **14,** 142 (2013).
21. Scott, K. C., Merrett, S. L. & Willard, H. F. A Heterochromatin Barrier Partitions the Fission Yeast Centromere into Discrete Chromatin Domains. *Current Biology* **16,** 119–129 (2006).
22. Noma, K.-I., Cam, H. P., Maraia, R. J. & Grewal, S. I. S. A role for TFIIIC transcription factor complex in genome organization. *Cell* **125,** 859–872 (2006).
23. Mizuguchi, T., Barrowman, J. & Grewal, S. I. S. Chromosome domain architecture and dynamic organization of the fission yeast genome. *FEBS Lett* **589,** 2975–2986 (2015).
24. Donze, D., Adams, C. R., Rine, J. & Kamakaka, R. T. The boundaries of the silenced HMR domain in Saccharomyces cerevisiae. *Genes Dev* **13,** 698–708 (1999).
25. Carabana, J., Watanabe, A., Hao, B. & Krangel, M. S. A barrier-type insulator forms a boundary between active and inactive chromatin at the murine TCRβ locus. *J. Immunol.* **186,** 3556–3562 (2011).
26. Valenzuela, L. & Kamakaka, R. T. Chromatin insulators. *Annu Rev Genet* **40,** 107–138 (2006).
27. Steglich, B. *et al.* The Fun30 chromatin remodeler Fft3 controls nuclear organization and chromatin structure of insulators and subtelomeres in fission yeast. *PLoS Genet* **11,** e1005101 (2015).
28. Iben, J. R. & Maraia, R. J. tRNAomics: tRNA gene copy number variation and codon use provide

bioinformatic evidence of a new anticodon:codon wobble pair in a eukaryote. *RNA* **18,** 1358–1372 (2012).
29. Ishii, K. *et al.* Heterochromatin integrity affects chromosome reorganization after centromere dysfunction. *Science* **321,** 1088–1091 (2008).
30. Takahashi, K., Chen, E. S. & Yanagida, M. Requirement of Mis6 centromere connector for localizing a CENP-A-like protein in fission yeast. *Science* **288,** 2215–2219 (2000).
31. Baum, M., Ngan, V. K. & Clarke, L. 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 (1994).
32. Guo, Y., Singh, P. K. & Levin, H. L. A long terminal repeat retrotransposon of Schizosaccharomyces japonicus integrates upstream of RNA pol III transcribed genes. *Mob DNA* **6,** 19 (2015).
33. Plohl, M., Luchetti, A., Mestrović, N. & Mantovani, B. Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* **409,** 72–82 (2008).
34. Nishibuchi, G. & Déjardin, J. The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. *Chromosome Res* **25,** 77–87 (2017).
35. Nonaka, N. *et al.* Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nat Cell Biol* **4,** 89–93 (2002).
36. Bernard, P. *et al.* Requirement of heterochromatin for cohesion at centromeres. *Science* **294,** 2539–2542 (2001).
37. Peters, J.-M. & Nishiyama, T. Sister chromatid cohesion. *Cold Spring Harb Perspect Biol* **4,** a011130–a011130 (2012).
38. Gregan, J. *et al.* The kinetochore proteins Pcs1 and Mde4 and heterochromatin are required to prevent merotelic orientation. *Curr Biol* **17,** 1190–1200 (2007).
39. Shen, M. H., Yang, J. W., Yang, J., Pendon, C. & Brown, W. R. The accuracy of segregation of human mini-chromosomes varies in different vertebrate cell lines, correlates with the extent of centromere formation and provides evidence for a trans-acting centromere maintenance activity. *Chromosoma* **109,** 524–535 (2001).
40. Okada, T. *et al.* CENP-B controls centromere formation depending on the chromatin context. *Cell* **131,** 1287–1300 (2007).
41. Moreno, S., Klar, A. & Nurse, P. Molecular genetic analysis of fission yeast Schizosaccharomyces pombe. *Meth Enzymol* **194,** 795–823 (1991).
42. Chin, C.-S. *et al.* Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* **10,** 563–569 (2013).
43. Murray, J. M., Watson, A. T. & Carr, A. M. Molecular Genetic Tools and Techniques in Fission Yeast. *Cold Spring Harb Protoc* **2016,** pdb.top087601 (2016).
44. Zhao, H. *et al.* CrossMap: a versatile tool for coordinate conversion between genome assemblies. *Bioinformatics* **30,** 1006–1007 (2014).
45. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25,** 955–964 (1997).
46. Schattner, P., Brooks, A. N. & Lowe, T. M. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* **33,** W686–9 (2005).
47. Hubley, R. *et al.* The Dfam database of repetitive DNA families. *Nucleic Acids Res* **44,** D81–9 (2016).
48. Kurtz, S. *et al.* Versatile and open software for comparing large genomes. *Genome Biol* **5,** R12 (2004).
49. Soderlund, C., Nelson, W., Shoemaker, A. & Paterson, A. SyMAP: A system for discovering and viewing syntenic regions of FPC maps. *Genome Res* **16,** 1159–1168 (2006).
50. Soderlund, C., Bomhoff, M. & Nelson, W. M. SyMAP v3.4: a turnkey synteny system with application to plant genomes. *Nucleic Acids Res* **39,** e68–e68 (2011).
51. Castillo, A. G. *et al.* Plasticity of fission yeast CENP-A chromatin driven by relative levels of histone H3 and H4. *PLoS Genet* **3,** e121 (2007).
52. Nakagawachi, T. *et al.* Silencing effect of CpG island hypermethylation and histone modifications on O6-methylguanine-DNA methyltransferase (MGMT) gene expression in human cancer. *Oncogene* **22,** 8835–8844 (2003).
53. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9,** 357–359 (2012).
54. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).

55. Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* **42,** W187–91 (2014).
56. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* **29,** 24–26 (2011).
57. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9,** R137 (2008).
58. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7,** 214 (2007).
59. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27,** 764–770 (2011).
60. Kruskal, W. H. & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Amer. Stat. Soc.* **47,** 583–621 (1952).
61. Suga, M. & Hatakeyama, T. High efficiency transformation of Schizosaccharomyces pombe pretreated with thiol compounds by electroporation. *Yeast* **18,** 1015–1021 (2001).
62. Craven, R. A. *et al.* Vectors for the expression of tagged proteins in Schizosaccharomyces pombe. *Gene* **221,** 59–68 (1998).
63. Sanchez, J. A., Kim, S. M. & Huberman, J. A. Ribosomal DNA replication in the fission yeast, Schizosaccharomyces pombe. *Experimental Cell Research* **238,** 220–230 (1998).
64. Tashiro, S., Nishihara, Y., Kugou, K., Ohta, K. & Kanoh, J. Subtelomeres constitute a safeguard for gene expression and chromosome homeostasis. *Nucleic Acids Res* **45,** 10333–10349 (2017).



S. octosporus vs S. pombe | *S. cryophilus vs S. pombe* | *S. octosporus vs S. cryophilus*

Figure 1, Tong et al.

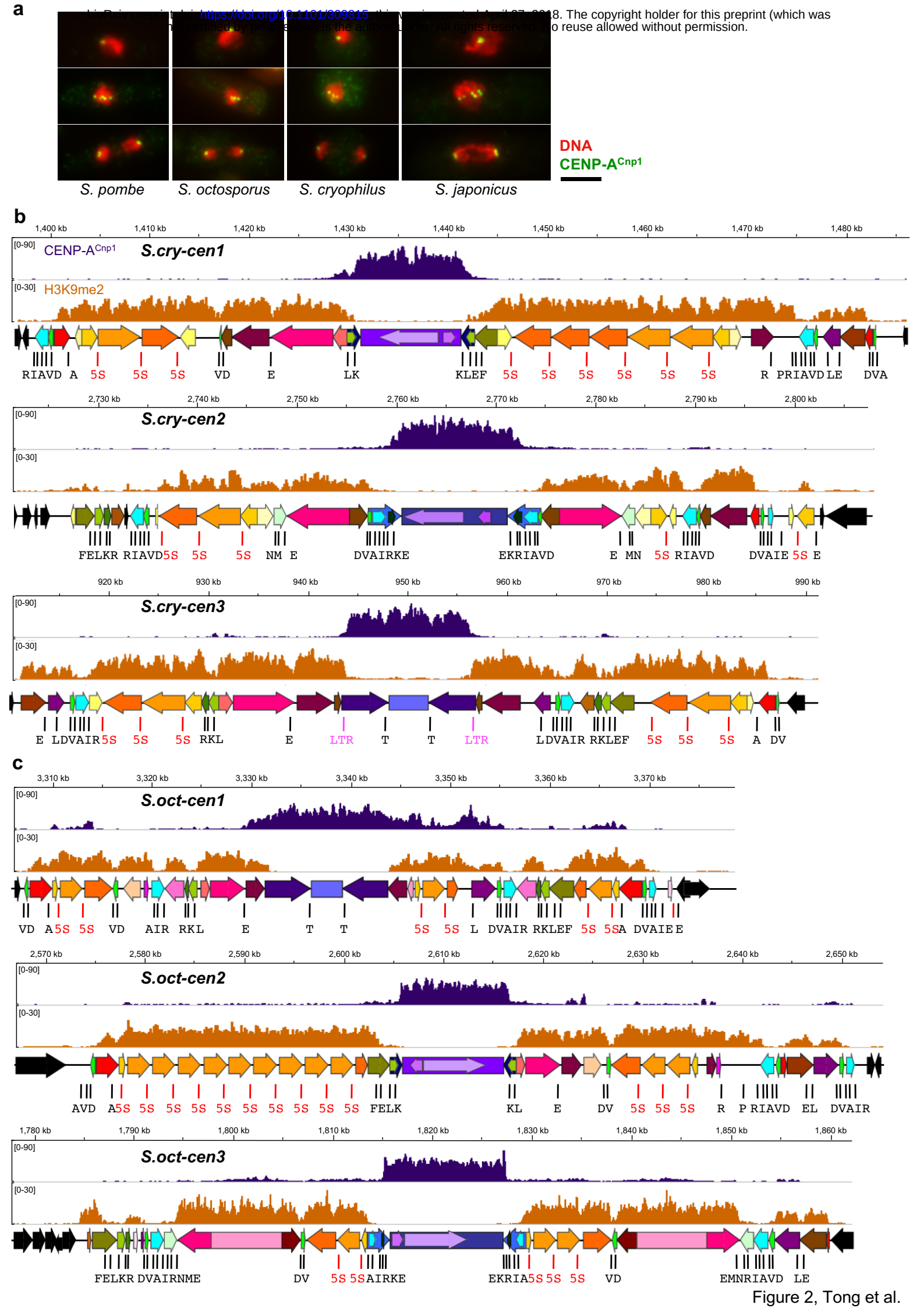


Figure 2, Tong et al.

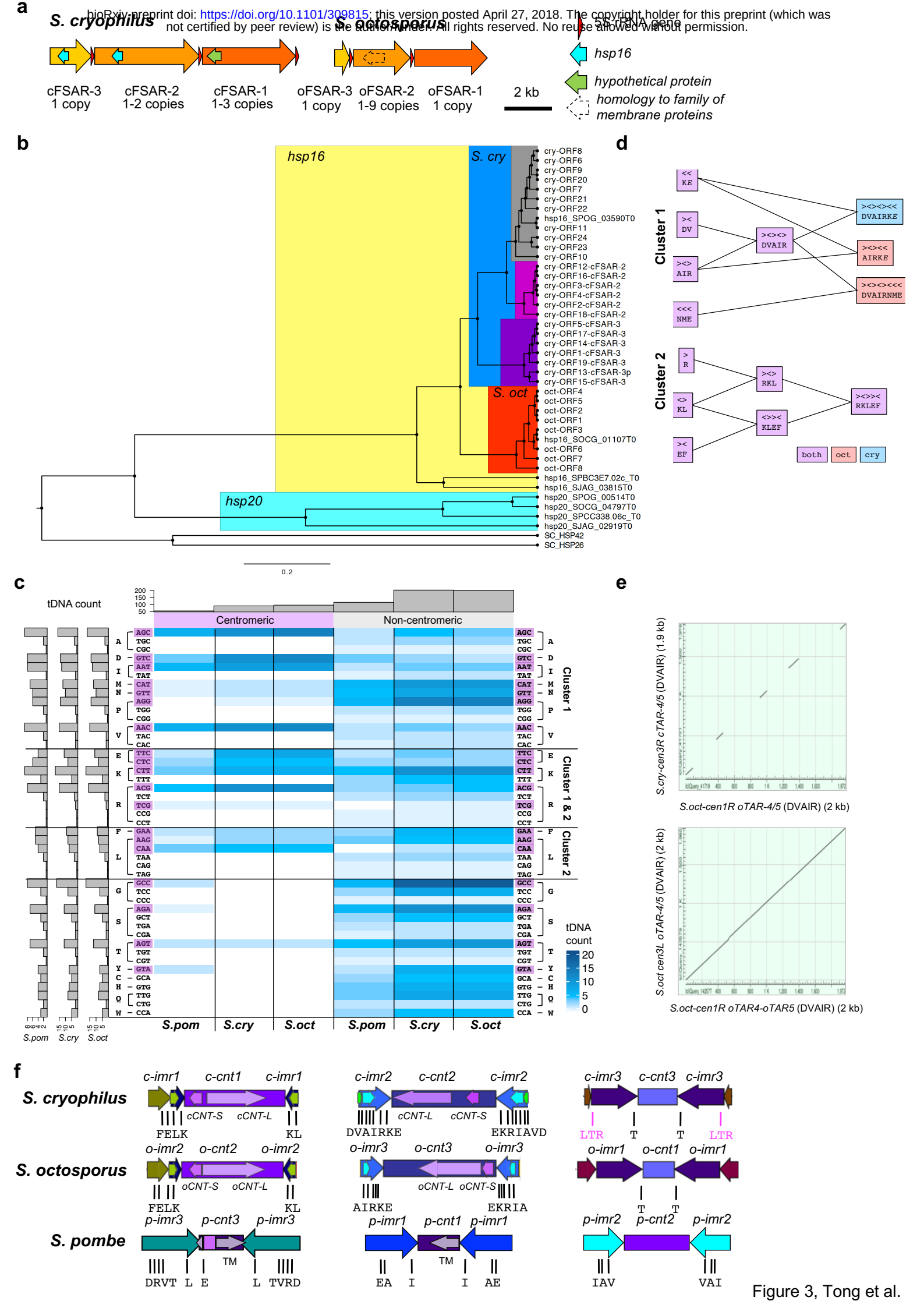
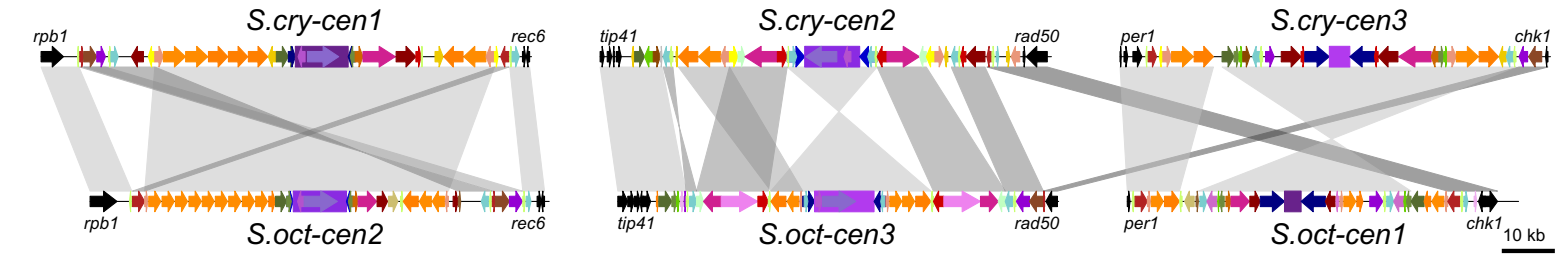
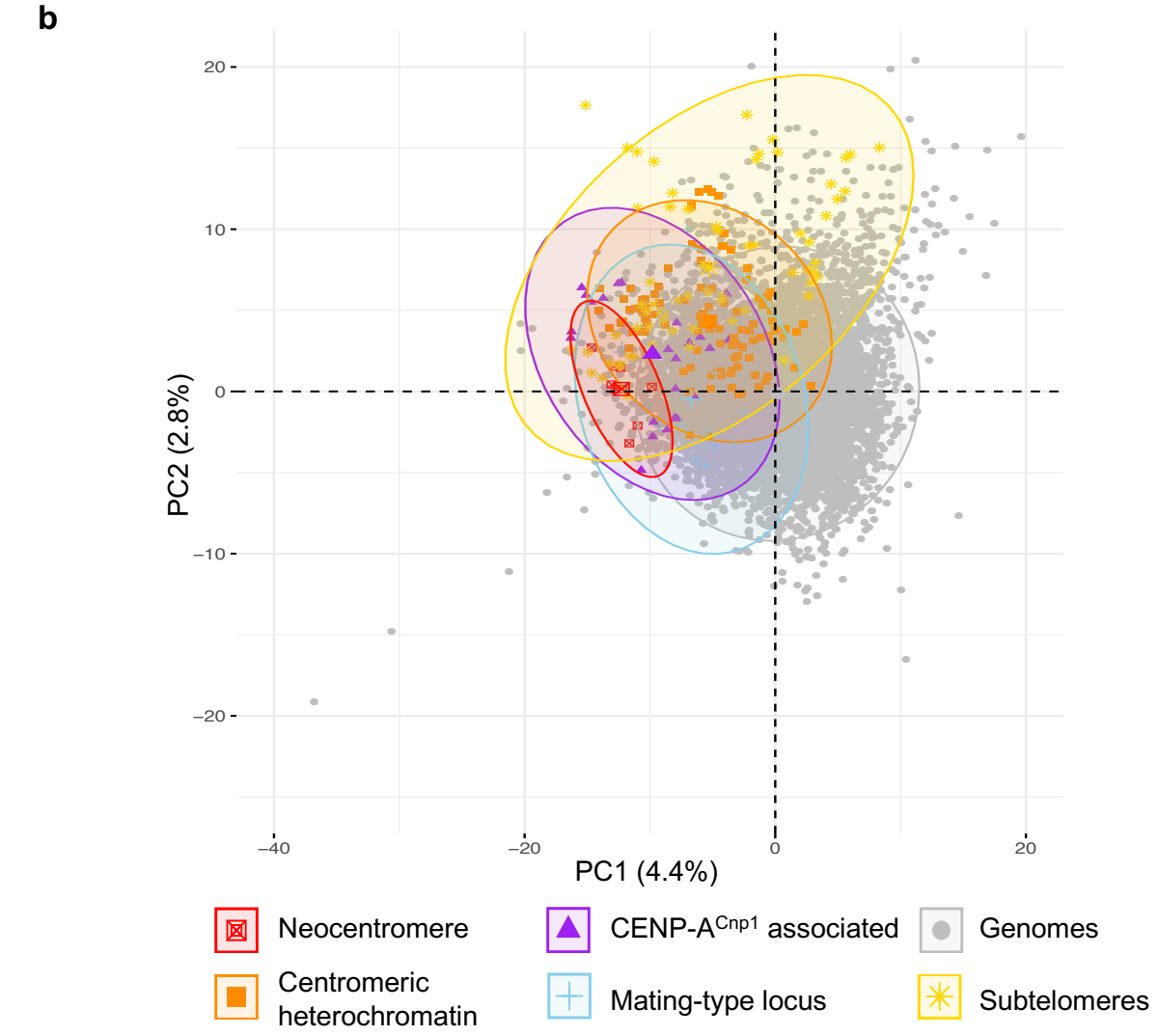


Figure 3, Tong et al.





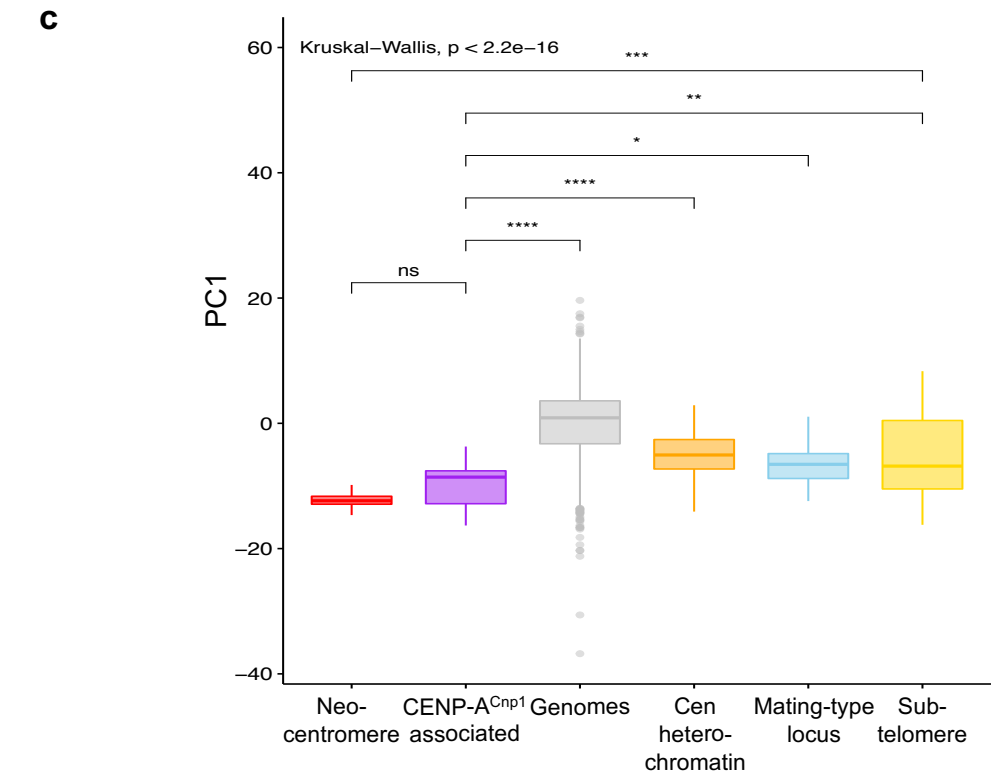
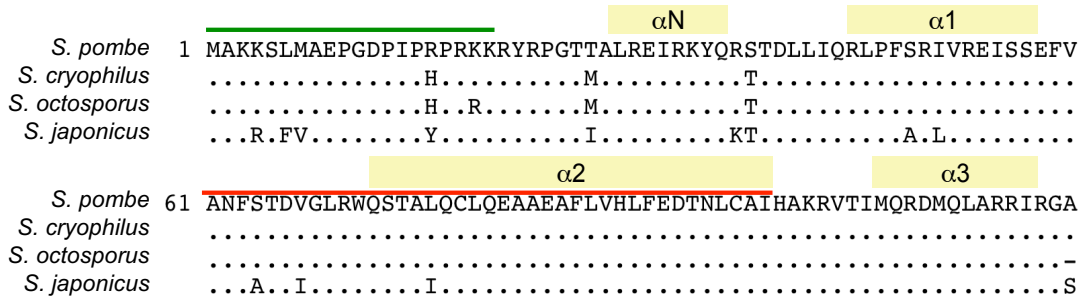
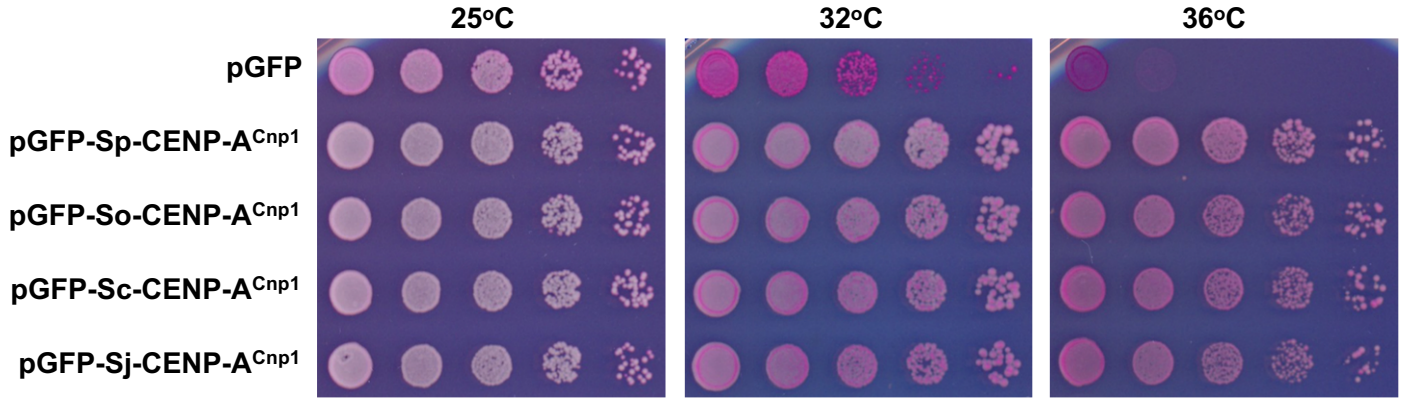


Figure 4, Tong et al.

a



b



c

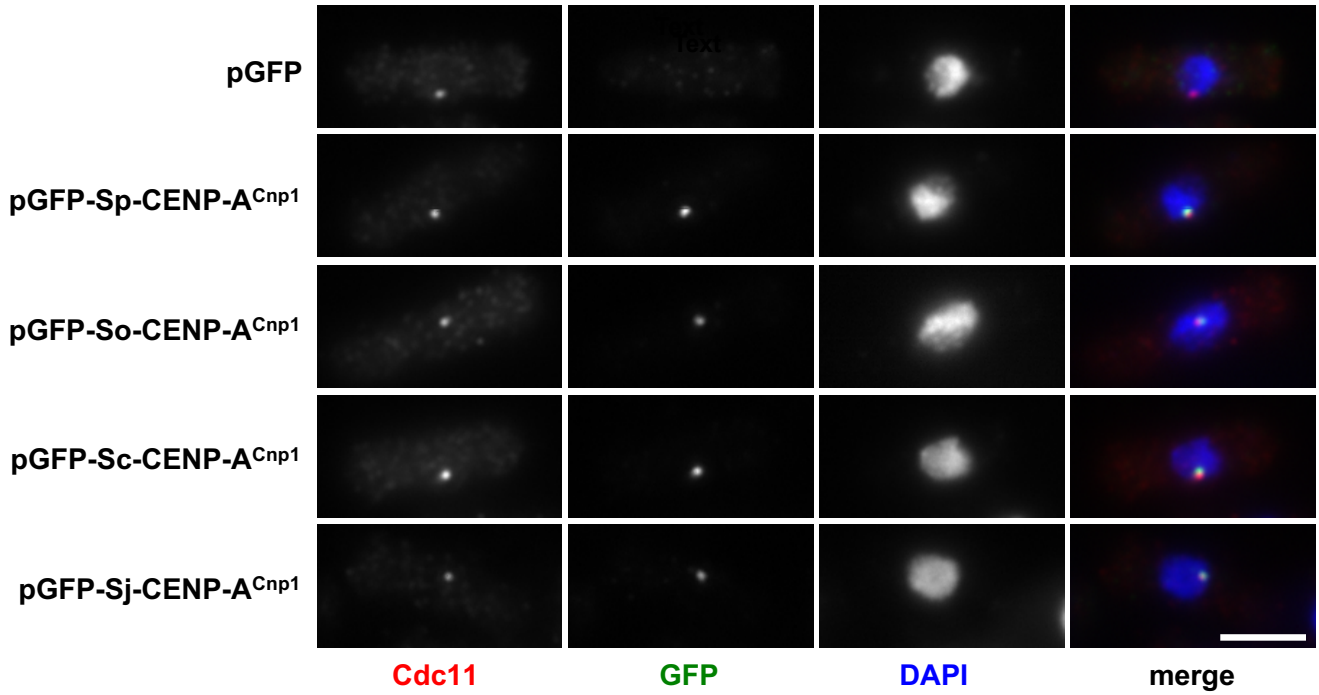
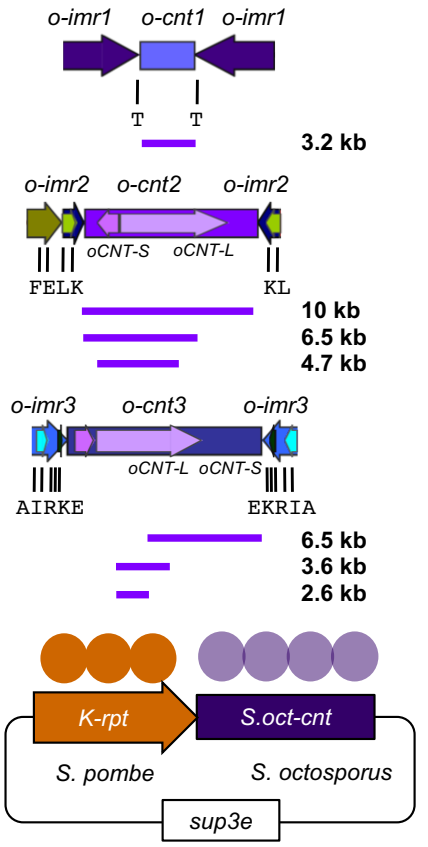


Figure 5, Tong et al.

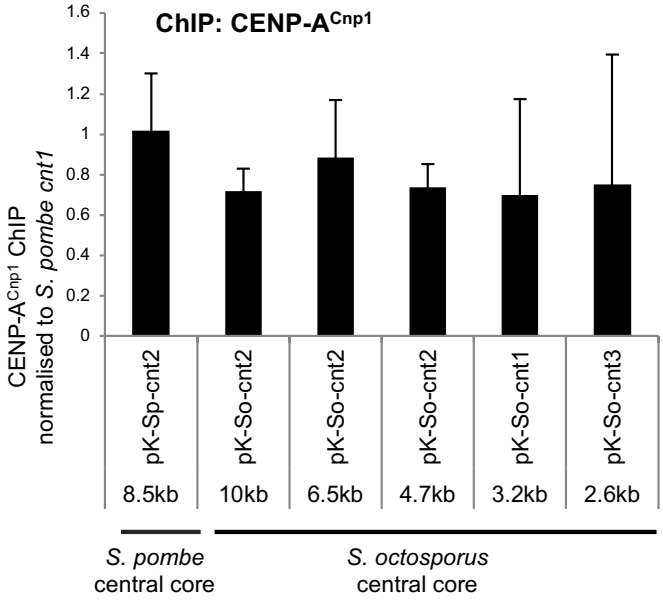
a



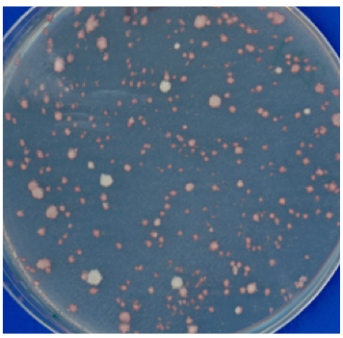
d

Plasmid	Establishment frequency % (n)	Loss rate per division % (n)
pK-Sp-cnt2-8.5kb	94 (217)	5.8 (3284)
pK-So-cnt2-10kb	65.6 (88)	11.2 (1636)
pK-So-cnt3-6.5kb	40 (262)	5.8 (3705)
pK-So-cnt2-6.5kb	32.6 (208)	6.6 (3621)
pK-So-cnt2-4.7kb	7.5 (973)	6.8 (7176)
pK-So-cnt1-3.2kb	2.3 (1529)	11.5 (2017)
pK-So-cnt3-2.6kb	2.1 (1443)	21.1 (1237)
pK-So-cnt3-3.6kb	1.1 (1939)	ND
pKp-So-cnt3-6.5kb	12.5 (916)	15.6 (2099)
pKp-So-cnt3-3.6kb	0 (1538)	NA
pKp	0 (295)	NA

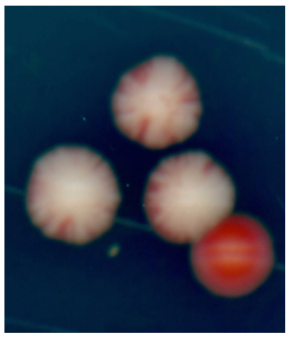
e



b



c



f

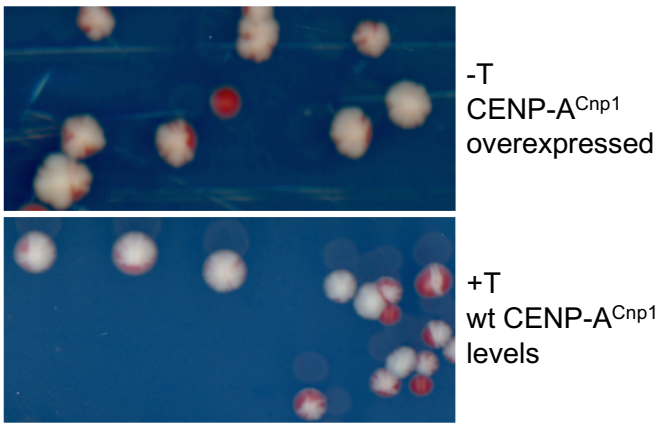
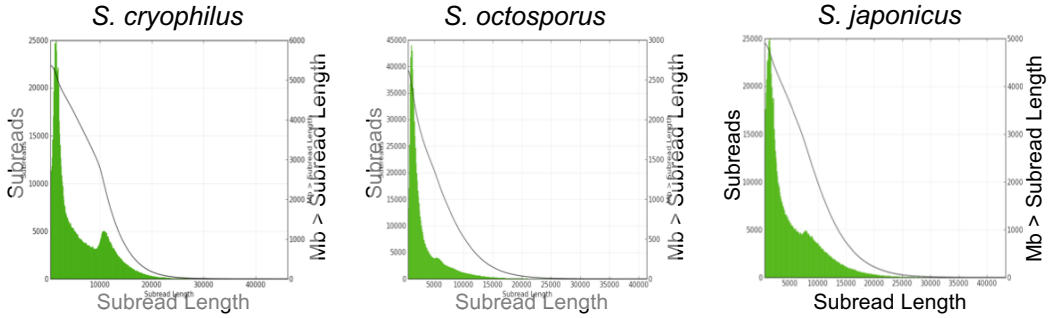


Figure 6, Tong et al

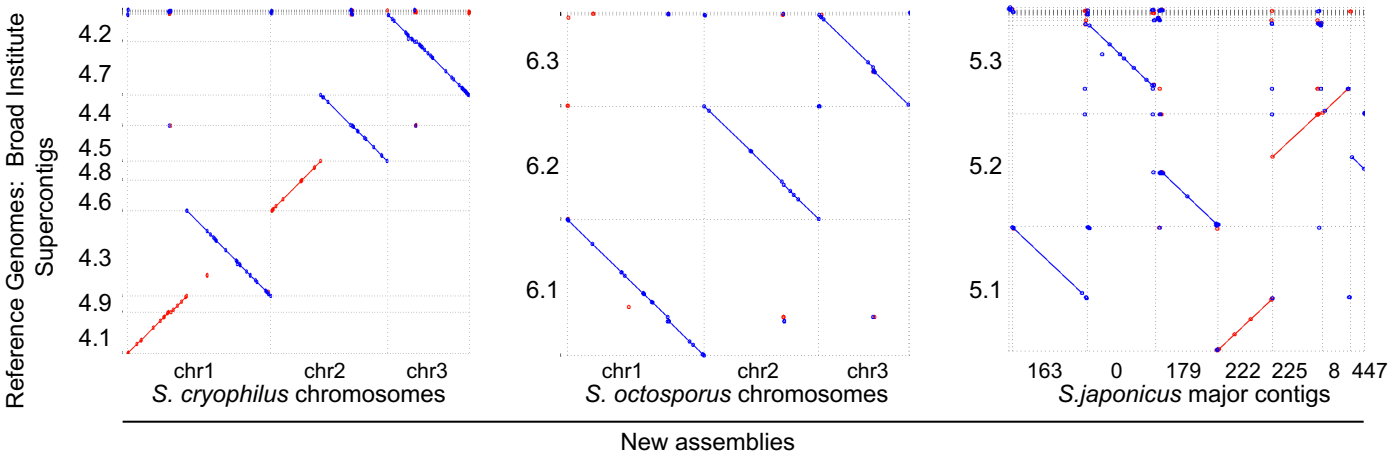
a



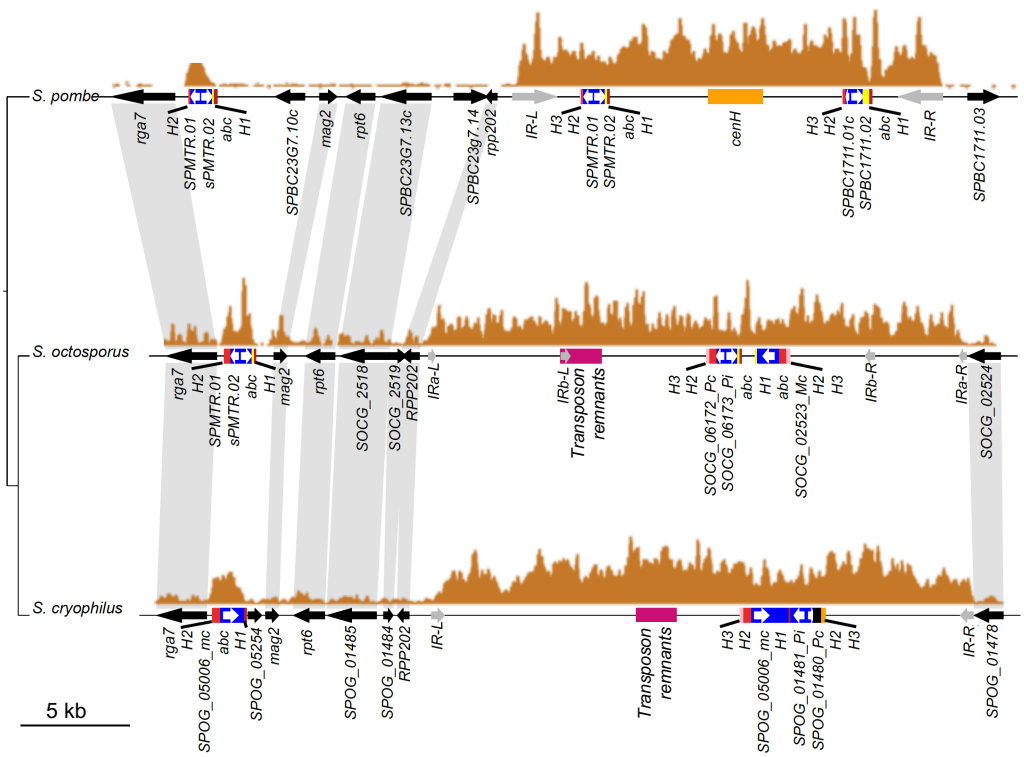
b

	S. cryophilus	*S. octosporus*	*S. japonicus*
Contig N50	2,797,701	3,794,048	1,051,694
Max contig length	4,890,599	4,519,984	2,096,611
Number of contigs	12	39	280
Sum of contig lengths	11,965,400	11,871,057	16,757,317
Published genome length (Broad)	11,589,478	11,678,700	11,813,213
Number of SMRT cells	11	11	15
PacBio reads N50	10,335	5,323	8,715
PacBio reads N90	16,558	12,893	16,805
PacBio Longest read	46,162	41,378	43,216

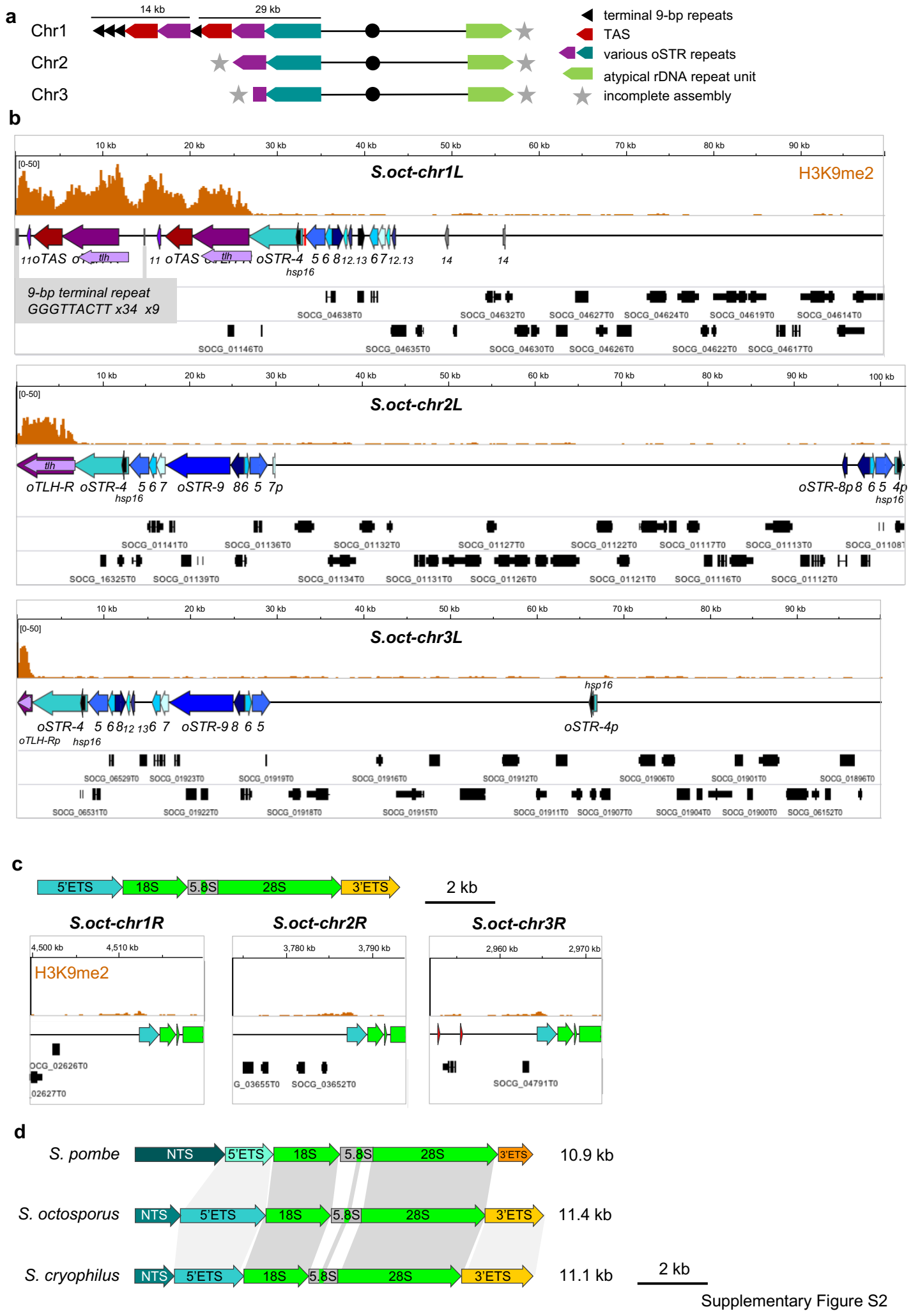
c



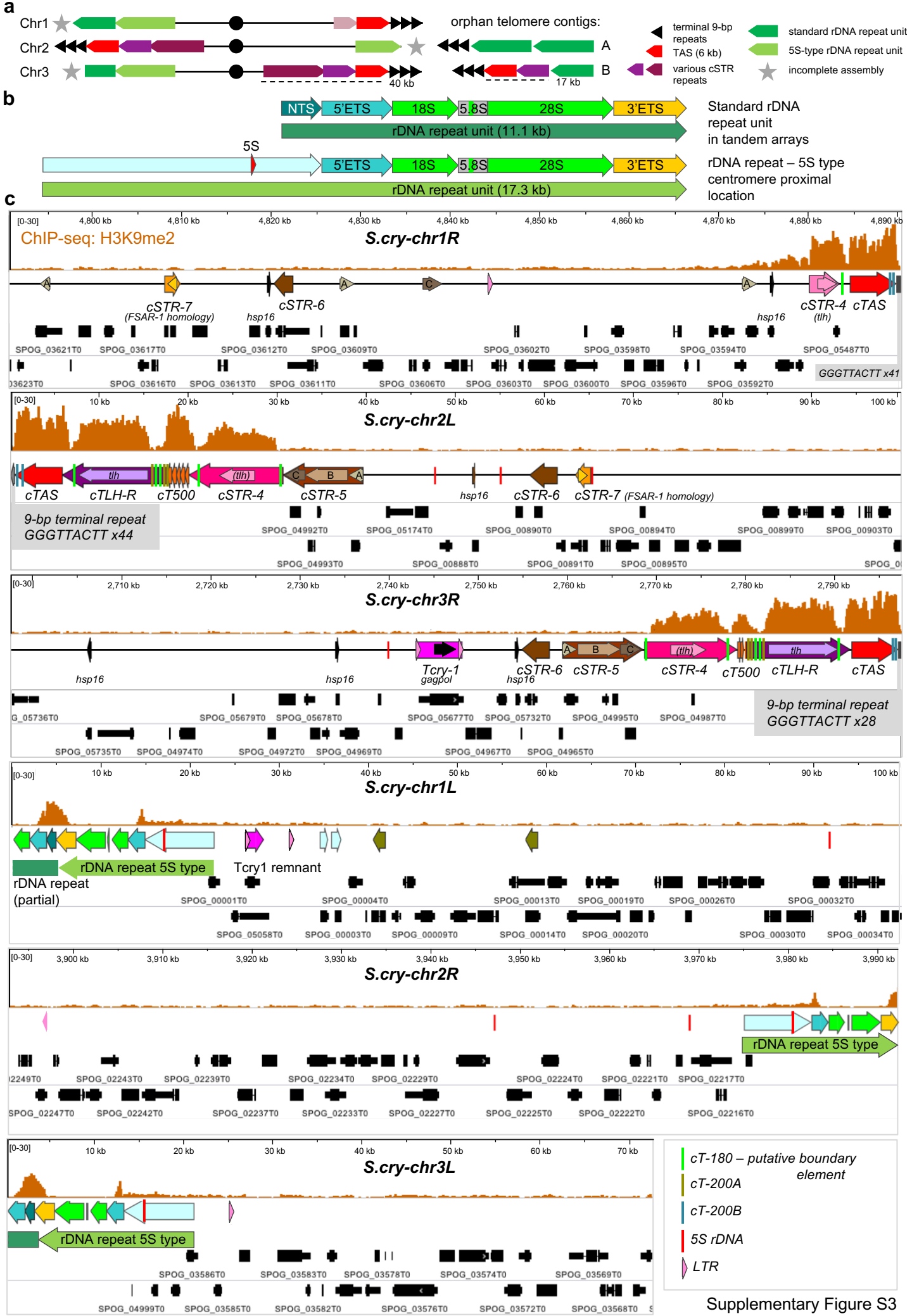
d



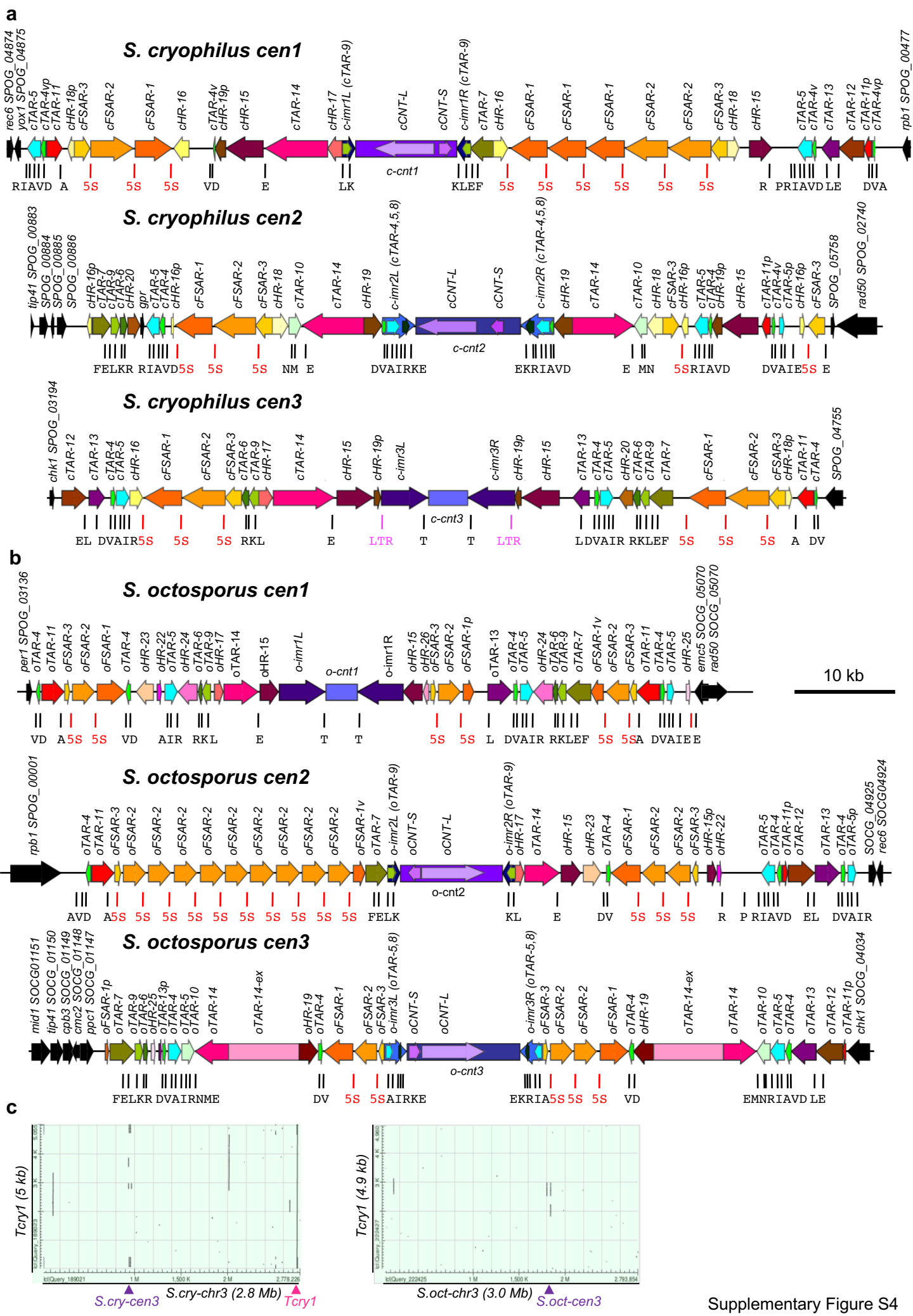
Supplementary Figure S1



Supplementary Figure S2

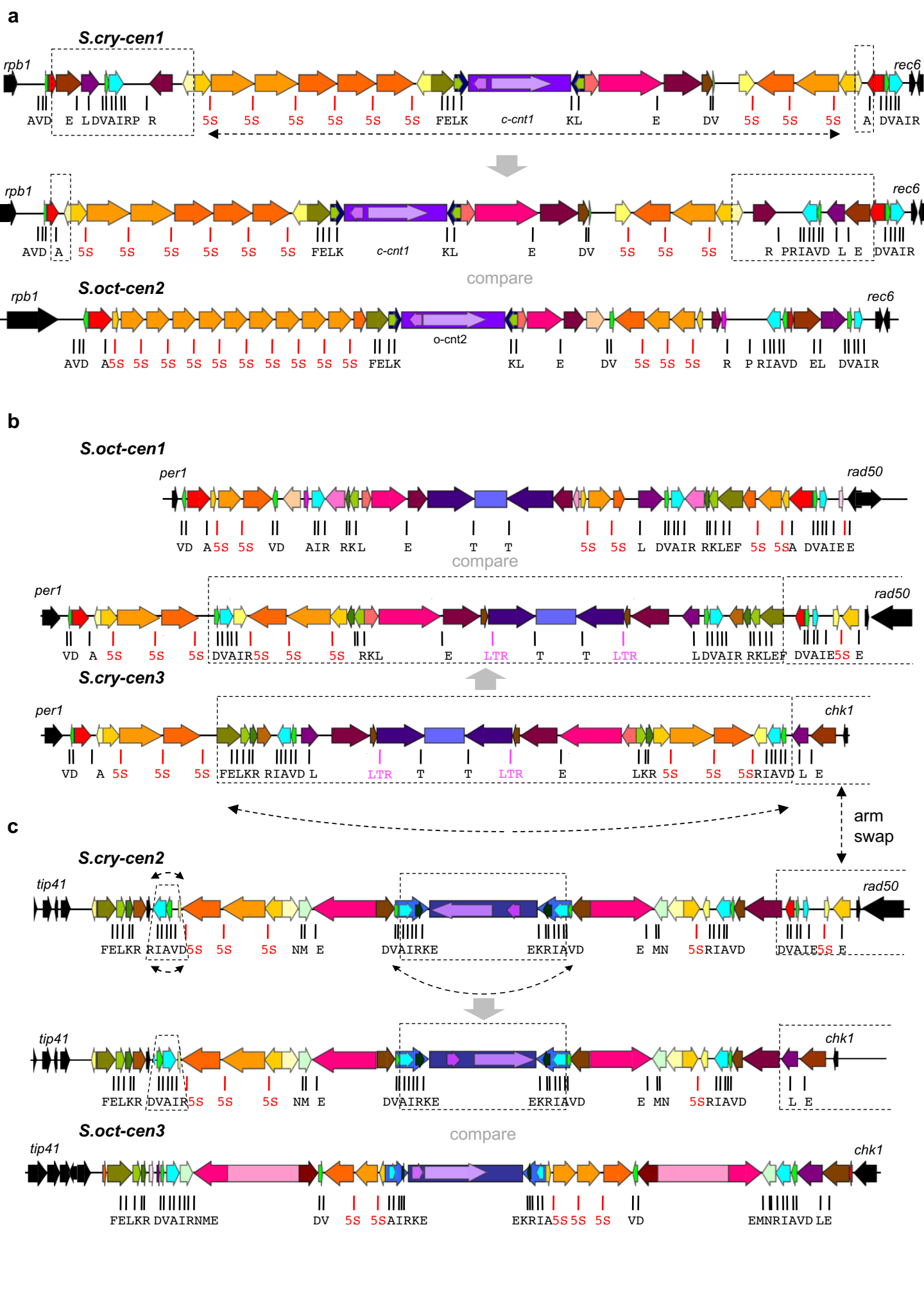


Supplementary Figure S3



Supplementary Figure S4

Supplementary Figure S5



Supplementary Figure S6

Supplementary Table 1: S. cryophilus telomere repeat annotation

chromosom	feature	start	end	size	strand	type
cry-chr1	cSTR-5C par	4790599	4790599	1216	-	repeat
cry-chr1	cSTR-7	4790599	4790599	1778	+	repeat
cry-chr1	cFSAR-1 hor	4790599	4790599	1166	-	repeat
cry-chr1	hsp16	4790599	4790599	429	+	gene
cry-chr1	cSTR-6	4790599	4790599	2128	-	repeat
cry-chr1	cSTR-5a	4790599	4790599	1811	+	repeat
cry-chr1	cSTR-5C	4790599	4790599	2135	+	repeat
cry-chr1	LTR	4790599	4790599	295	+	LTR
cry-chr1	9bp-rpt*	4790599	4790599	9	+	repeat
cry-chr1	cSTR-5a	4790599	4790599	1395	+	repeat
cry-chr1	hsp16	4790599	4790599	429	+	gene
cry-chr1	cSTR-4p	4790599	4790599	3389	+	repeat
cry-chr1	degraded tlh	4790599	4790599	2282	+	repeat
cry-chr1	weak 180-rpt	4790599	4790599	180	+	repeat
cry-chr1	cTAS	4790599	4790599	5237	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	-	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	-	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	-	repeat
cry-chr1	rpt-200B	4790599	4790599	198	+	repeat
cry-chr1	rpt-200B	4790599	4790599	205	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat

cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr1	9bp-rpt	4790599	4790599	9	+	repeat
cry-chr2	9bp-rpt	5	13	9	-	repeat
cry-chr2	9bp-rpt	14	22	9	-	repeat
cry-chr2	9bp-rpt	23	31	9	-	repeat
cry-chr2	9bp-rpt	32	40	9	-	repeat
cry-chr2	9bp-rpt	41	49	9	-	repeat
cry-chr2	9bp-rpt	58	66	9	-	repeat
cry-chr2	9bp-rpt	67	75	9	-	repeat
cry-chr2	9bp-rpt	76	84	9	-	repeat
cry-chr2	9bp-rpt	85	93	9	-	repeat
cry-chr2	9bp-rpt	94	102	9	-	repeat
cry-chr2	9bp-rpt	103	111	9	-	repeat
cry-chr2	9bp-rpt	112	120	9	-	repeat
cry-chr2	9bp-rpt	121	129	9	-	repeat
cry-chr2	9bp-rpt	130	138	9	-	repeat
cry-chr2	9bp-rpt	139	147	9	-	repeat
cry-chr2	9bp-rpt	148	156	9	-	repeat
cry-chr2	9bp-rpt	157	165	9	-	repeat
cry-chr2	9bp-rpt	166	174	9	-	repeat
cry-chr2	9bp-rpt	175	183	9	-	repeat
cry-chr2	9bp-rpt	184	192	9	-	repeat
cry-chr2	9bp-rpt	193	201	9	-	repeat
cry-chr2	9bp-rpt	202	210	9	-	repeat
cry-chr2	9bp-rpt	211	219	9	-	repeat
cry-chr2	9bp-rpt	220	228	9	-	repeat
cry-chr2	9bp-rpt	229	237	9	-	repeat
cry-chr2	9bp-rpt	238	246	9	-	repeat
cry-chr2	9bp-rpt	247	255	9	-	repeat
cry-chr2	9bp-rpt	256	264	9	-	repeat
cry-chr2	9bp-rpt	265	273	9	-	repeat
cry-chr2	9bp-rpt	282	290	9	-	repeat
cry-chr2	9bp-rpt	291	299	9	-	repeat
cry-chr2	9bp-rpt	300	308	9	-	repeat
cry-chr2	9bp-rpt	309	317	9	-	repeat
cry-chr2	9bp-rpt	318	326	9	-	repeat
cry-chr2	9bp-rpt	327	335	9	-	repeat
cry-chr2	9bp-rpt	336	344	9	-	repeat
cry-chr2	9bp-rpt	345	353	9	-	repeat
cry-chr2	9bp-rpt	354	362	9	-	repeat
cry-chr2	9bp-rpt	363	371	9	-	repeat
cry-chr2	9bp-rpt	372	380	9	-	repeat
cry-chr2	9bp-rpt	381	389	9	-	repeat
cry-chr2	9bp-rpt	390	398	9	-	repeat
cry-chr2	9bp-rpt	399	407	9	-	repeat
cry-chr2	9bp-rpt	408	416	9	-	repeat
cry-chr2	cTAS	423	5712	5290	-	repeat
cry-chr2	cT-200B	539	743	205	-	repeat
cry-chr2	cT-200B	862	1059	198	-	repeat
cry-chr2	9bp-rpt	4016	4024	9	+	repeat
cry-chr2	9bp-rpt	4902	4910	9	+	repeat

cry-chr2	9bp-rpt	5076	5084	9	+	repeat
cry-chr2	cTLH-R	5713	17,539	11,827	-	repeat
cry-chr2	cT-180	6968	7147	180	-	repeat
cry-chr2	CTTAGA-rep	7494	7568	75	+	repeat
cry-chr2	tlh gene	7738	15,485	7748	-	gene
cry-chr2	cT-300	15,815	16,109	295	-	repeat
cry-chr2	cT-200	15,854	16,051	198	-	repeat
cry-chr2	cT-180p	16,134	16,278	145	-	repeat
cry-chr2	cT-180	16,284	16,464	181	-	repeat
cry-chr2	cT-300	16,491	16,788	298	-	repeat
cry-chr2	cT-200	16,529	16,728	200	-	repeat
cry-chr2	cT-200	17,046	17,245	200	-	repeat
cry-chr2	cT-500	17,540	18,047	508	-	repeat
cry-chr2	cT-500	18,048	18,552	505	-	repeat
cry-chr2	cT-500	18,553	19,045	493	-	repeat
cry-chr2	cT-500	19,046	19,559	514	-	repeat
cry-chr2	cT-500p	19,560	19,923	364	-	repeat
cry-chr2	cSTR-4	19,924	30,491	10,568	-	repeat
cry-chr2	cT-180	20,991	21,170	180	-	repeat
cry-chr2	CTTAGA-rep	21,492	21,545	54	+	repeat
cry-chr2	degraded tlh	23,412	27,370	3959	-	misc_feature
cry-chr2	9bp-rpt	24,613	24,621	9	+	repeat
cry-chr2	cT-180p	30,153	30,324	172	-	repeat
cry-chr2	cSTR-5	30,492	39,517	9026	-	repeat
cry-chr2	cSTR-5C	30,725	32,898	2174	-	repeat
cry-chr2	cSTR-5B	32,899	37,672	4774	-	repeat
cry-chr2	cSTR-5A	37,673	39,517	1845	-	repeat
cry-chr2	5S rDNA	47,503	47,623	121	+	rRNA
cry-chr2	cSTR-5B par	51,718	52,104	387	+	repeat
cry-chr2	5S rDNA	54,859	54,979	121	+	rRNA
cry-chr2	cSTR-6	58,176	61,238	3063	-	repeat
cry-chr2	cSTR-7	63,271	64,979	1709	-	repeat
cry-chr2	cFSAR-1 hor	63,640	64,760	1121	+	repeat
cry-chr2	5S rDNA	65,079	65,199	121	+	rRNA
cry-chr3	hsp16	2697701	2697701	429	-	gene
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	-	repeat
cry-chr3	hsp16	2697701	2697701	429	+	gene
cry-chr3	5S rRNA	2697701	2697701	115	-	rRNA
cry-chr3	Tcry1-1 retro	2697701	2697701	5055	+	retrotransposon
cry-chr3	Tcry1-LTR	2697701	2697701	374	+	LTR
cry-chr3	Tcry1-LTR	2697701	2697701	374	+	LTR
cry-chr3	hsp16	2697701	2697701	429	+	gene
cry-chr3	cSTR-6	2697701	2697701	3018	-	repeat
cry-chr3	cSTR-5	2697701	2697701	9033	+	repeat
cry-chr3	cSTR-5A	2697701	2697701	1828	+	repeat
cry-chr3	cSTR-5B	2697701	2697701	4806	+	repeat
cry-chr3	cSTR-5C	2697701	2697701	2166	+	repeat
cry-chr3	cSTR-4	2697701	2697701	10,548	+	repeat
cry-chr3	cT-180p	2697701	2697701	172	+	repeat
cry-chr3	tlh gene	2697701	2697701	337	+	gene
cry-chr3	9bp-rpt	2697701	2697701	9	-	repeat
cry-chr3	cT-180	2697701	2697701	180	+	repeat
cry-chr3	cT-500p	2697701	2697701	271	+	repeat
cry-chr3	cT-500p	2697701	2697701	419	+	repeat

cry-chr3	cT-500p	2697701	2697701	155	+	repeat
cry-chr3	cTLH-R	2697701	2697701	11,730	+	repeat
cry-chr3	cT-180	2697701	2697701	181	+	repeat
cry-chr3	cT-180p	2697701	2697701	145	+	repeat
cry-chr3	tlh gene	2697701	2697701	7748	+	gene
cry-chr3	cT-180	2697701	2697701	180	+	repeat
cry-chr3	cTAS	2697701	2697701	5289	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	-	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	-	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	-	repeat
cry-chr3	cT-200B	2697701	2697701	198	+	repeat
cry-chr3	cT-200B	2697701	2697701	205	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat
cry-chr3	9bp-rpt	2697701	2697701	9	+	repeat

* terminal telomere repeat, GGGTTACTT, ref 16

Supplementary Table 2: S. octosporus telomere repeat annotation

chromosome	feature	start	end	size	strand	type	
oct-chr1	9bp-rpt*	6	14	9	-	repeat	
oct-chr1	9bp-rpt	15	23	9	-	repeat	
oct-chr1	9bp-rpt	24	32	9	-	repeat	
oct-chr1	9bp-rpt	33	41	9	-	repeat	
oct-chr1	9bp-rpt	42	50	9	-	repeat	
oct-chr1	9bp-rpt	51	59	9	-	repeat	
oct-chr1	9bp-rpt	60	68	9	-	repeat	
oct-chr1	9bp-rpt	77	85	9	-	repeat	
oct-chr1	9bp-rpt	86	94	9	-	repeat	
oct-chr1	9bp-rpt	95	103	9	-	repeat	
oct-chr1	9bp-rpt	104	112	9	-	repeat	
oct-chr1	9bp-rpt	113	121	9	-	repeat	
oct-chr1	9bp-rpt	122	130	9	-	repeat	
oct-chr1	9bp-rpt	131	139	9	-	repeat	
oct-chr1	9bp-rpt	140	148	9	-	repeat	
oct-chr1	9bp-rpt	149	157	9	-	repeat	
oct-chr1	9bp-rpt	158	166	9	-	repeat	
oct-chr1	9bp-rpt	167	175	9	-	repeat	
oct-chr1	9bp-rpt	176	184	9	-	repeat	
oct-chr1	9bp-rpt	185	193	9	-	repeat	
oct-chr1	9bp-rpt	202	210	9	-	repeat	
oct-chr1	9bp-rpt	211	219	9	-	repeat	
oct-chr1	9bp-rpt	220	228	9	-	repeat	
oct-chr1	9bp-rpt	229	237	9	-	repeat	
oct-chr1	9bp-rpt	238	246	9	-	repeat	
oct-chr1	9bp-rpt	247	255	9	-	repeat	
oct-chr1	9bp-rpt	256	264	9	-	repeat	
oct-chr1	9bp-rpt	265	273	9	-	repeat	
oct-chr1	9bp-rpt	274	282	9	-	repeat	
oct-chr1	9bp-rpt	283	291	9	-	repeat	
oct-chr1	9bp-rpt	292	300	9	-	repeat	
oct-chr1	9bp-rpt	301	309	9	-	repeat	
oct-chr1	9bp-rpt	310	318	9	-	repeat	
oct-chr1	9bp-rpt	324	332	9	-	repeat	
oct-chr1	oSTR-11	1237	1746	510	-	repeat	
oct-chr1	oTAS	2000	5292	3293	-	repeat	
oct-chr1	oTLH-R	5293	11,855	6563	-	repeat	
oct-chr1	tlh gene	6024	12,350	6327	-	gene	
oct-chr1	9bp-rpt	14,717	14,725	9	-	repeat	
oct-chr1	9bp-rpt	14,726	14,734	9	-	repeat	
oct-chr1	9bp-rpt	14,735	14,743	9	-	repeat	
oct-chr1	9bp-rpt	14,744	14,752	9	-	repeat	
oct-chr1	9bp-rpt	14,761	14,769	9	-	repeat	
oct-chr1	9bp-rpt	14,770	14,778	9	-	repeat	
oct-chr1	9bp-rpt	14,779	14,787	9	-	repeat	
oct-chr1	9bp-rpt	14,788	14,796	9	-	repeat	
oct-chr1	9bp-rpt	14,802	14,810	9	-	repeat	
oct-chr1	oSTR-11	16,192	16,697	506	-	repeat	
oct-chr1	oTAS	17,177	20,263	3087	-	repeat	
oct-chr1	oTLH-R	20,264	26,833	6570	-	repeat	
oct-chr1	tlh gene	20,996	26,995	6000	-	gene	
oct-chr1	oSTR-4	26,834	33,000	6167	-	repeat	
oct-chr1	hsp16	32,337	32,764	428	-	gene	
oct-chr1	5S rRNA	33,223	33,349	127	+	rRNA	
oct-chr1	oSTR-5	33,350	35,528	2179	-	repeat	

oct-chr1	oSTR-6	35,529	36,427	899	-	repeat	
oct-chr1	oSTR-8	36,428	37,643	1216	+	repeat	
oct-chr1	oSTR-12	37,650	38,062	413	-	repeat	
oct-chr1	oSTR-13	38,198	38,657	460	-	repeat	
oct-chr1	oSTR-6	40,665	41,599	935	-	repeat	
oct-chr1	oSTR-7	41,600	42,479	880	-	repeat	
oct-chr1	oSTR-12	42,635	43,040	406	-	repeat	
oct-chr1	oSTR-13	43,224	43,683	460	-	repeat	
oct-chr1	oSTR-14	49,350	49,752	403	-	repeat	
oct-chr1	oSTR-14	55,981	56,383	403	+	repeat	
oct-chr2	oTLH-R	1	6560	6560	-	repeat	
oct-chr2	tlh gene	725	6722	5998	-	gene	
oct-chr2	oSTR-4	6561	12,977	6417	-	repeat	
oct-chr2	hsp16	12,079	12,506	428	-	gene	
oct-chr2	oSTR-5	13,092	15,263	2172	-	repeat	
oct-chr2	oSTR-6	15,264	16,197	934	-	repeat	
oct-chr2	oSTR-7	16,198	17,122	925	-	repeat	
oct-chr2	oSTR-9	17,246	24,578	7333	-	repeat	
oct-chr2	oSTR-8	24,708	26,271	1564	-	repeat	
oct-chr2	oSTR-6	26,272	26,906	635	+	repeat	
oct-chr2	oSTR-5	26,907	29,054	2148	+	repeat	
oct-chr2	oSTR-7p	29,560	29,877	318	+	repeat	
oct-chr2	oSTR-8p	95,123	95,594	472	-	repeat	
oct-chr2	oSTR-8	96,825	98,318	1494	-	repeat	
oct-chr2	oSTR-6	98,319	98,920	602	+	repeat	
oct-chr2	oSTR-5	98,921	101,069	2149	+	repeat	
oct-chr2	oSTR-4p	101,184	102,090	907	+	repeat	
oct-chr2	hsp16	101,655	101,979	325	+	gene	
oct-chr3	tlh gene	1	1755	1755	-	gene	
oct-chr3	oTLH-R	1	1755	1755	-	repeat	
oct-chr3	oSTR-4	1756	8156	6401	-	repeat	
oct-chr3	hsp16	7260	7687	428	-	gene	
oct-chr3	oSTR-5	8271	10,447	2177	-	repeat	
oct-chr3	oSTR-6v	10,448	11,348	901	-	repeat	
oct-chr3	oSTR-8	11,349	12,544	1196	+	repeat	
oct-chr3	oSTR-12	12,571	12,983	413	-	repeat	
oct-chr3	oSTR-13	13,114	13,573	460	-	repeat	
oct-chr3	oSTR-6	15,601	16,535	935	-	repeat	
oct-chr3	oSTR-7	16,536	17,460	925	-	repeat	
oct-chr3	oSTR-9	17,584	24,923	7340	-	repeat	
oct-chr3	oSTR-8	25,053	26,302	1250	-	repeat	
oct-chr3	oSTR-6v	26,332	27,232	901	+	repeat	
oct-chr3	oSTR-5	27,233	29,383	2151	+	repeat	
oct-chr3	oSTR-4p	66,187	67,142	956	-	repeat	
oct-chr3	hsp16	66,344	66,771	428	-	gene	
* terminal telomere repeat, GGGTTACTT, ref 16							

Supplementary Table 3: S. cryophilus centromere repeat annotation

Repeats	Features	tRNA-anticodon	size (kb)	% GC	comments
cFSAR-1	5S rDNA; ORF		3.7	35	ORF - hypothetical protein
cFSAR-2	5S rDNA; hsp16		4.1	35	
cFSAR-3	5S rDNA; hsp16		1.7	34	
cTAR-4	tDNA: DV	AspGTC, ValAAC	0.5	38	several variants
cTAR-5	tDNA: AIR	AlaAGC, IleAAT, ArgACG	1.4	33	
cTAR-6	tDNA: R	ArgACG	0.7	34	as RKL usually
cTAR-7	tDNA: EF	GluCTC, PheGAA	2.3	34	
cTAR-8	tDNA: KE	LysCTT, GluCTC	0.7	34	Part of imr3
cTAR-9	tDNA: LK	LeuCAA, LysCTT	0.9	31	
cTAR-10	tDNA: NM	AsnGTT, MetCAT,	1.3	33	always as NME (but E sometimes alone in cTAR-14)
cTAR-11	tDNA: A	AlaAGC	1.6	31	
cTAR-12	tDNA: E	GluTTC	2.5	31	
cTAR-13	tDNA: L	LeuAAG	1.8	31	
cTAR-14	tDNA: E	GluTTC	6.2	34	retrotransposon remnant, sometimes alone, sometimes as NME
cHR-15			3.8	34	
cHR-16			1.4	35	
cHR-17			1.5	34	
cHR-18			0.7	31	
cHR-19			1.9	33	
cHR-20			1.4	33	
cHR-21			0.7	33	
cCNT-L			6	33	
cCNT-S			1.3	31	
c-cnt1			10.1	33	Contains cCNT-L and cCNT-S
c-cnt2			10.5	32	Contains cCNT-L and cCNT-S
c-cnt3			3.9	34	
c-imr1	tDNA: LK	LeuCAA, LysCTT	1.3	31	Contains cTAR-9
c-imr2	tDNA: DVAIRKE	AspGTC, ValAAC, AlaAGC, IleAAT, ArgACG, LysCTT, GluCTC	1.5	32	Contains cTAR-4, cTAR-5, cTAR-8
c-imr3	tDNA: T; LTR	ThrAGT	4.7	31	

Supplementary Table 5: S. cryophilus centromere repeat cordinates

chromosome	feature	start	end	size	strand	type	
cry-chr1	rec6 SPOG_04874	1,396,362	1,396,991	630	+	gene	
cry-chr1	yox1 SPOG_0475	1,397,122	1,397,733	612	-	gene	
cry-chr1	cTAR-5	1,398,354	1,399,731	1378	-	repeat	
cry-chr1	cTAR-4vp	1,399,732	1,400,178	447	-	repeat	
cry-chr1	cTAR-11	1,400,337	1,401,910	1574	+	repeat	
cry-chr1	cHR-18p	1,402,402	1,403,095	694	-	repeat	
cry-chr1	cFSAR-3-1	1,403,096	1,404,738	1643	+	repeat	
cry-chr1	cFSAR-2-1	1,404,747	1,408,996	4250	+	repeat	
cry-chr1	cFSAR-1-1	1,409,127	1,412,833	3707	+	repeat	
cry-chr1	cHR-16	1,412,949	1,414,488	1540	-	repeat	
cry-chr1	cTAR-4v	1,416,885	1,417,121	237	-	repeat	
cry-chr1	cHR-19p	1,417,136	1,418,158	1023	-	repeat	
cry-chr1	cHR-21	1,417,438	1,418,158	721	-	repeat	
cry-chr1	cHR-15	1,418,160	1,421,951	3792	-	repeat	
cry-chr1	cTAR-14	1,422,098	1,428,300	6203	-	repeat	
cry-chr1	cHR-17	1,428,302	1,429,800	1499	-	repeat	
cry-chr1	c-imr1L	1,429,801	1,431,143	1343	+	repeat	
cry-chr1	cTAR-9	1,429,801	1,430,748	948	+	repeat	
cry-chr1	c-cnt1	1,431,144	1,441,262	10,119	+	repeat	
cry-chr1	cCNT-L	1,432,959	1,438,927	5969	-	repeat	
cry-chr1	cCNT-S	1,439,401	1,440,740	1340	+	repeat	
cry-chr1	c-imr1R	1,441,263	1,442,605	1343	-	repeat	
cry-chr1	cTAR-9	1,441,658	1,442,605	948	-	repeat	
cry-chr1	cTAR-7	1,442,606	1,444,909	2304	-	repeat	
cry-chr1	cHR-16	1,444,910	1,446,448	1539	+	repeat	
cry-chr1	cFSAR-1-2	1,446,564	1,450,272	3709	-	repeat	
cry-chr1	cFSAR-1-3	1,450,388	1,454,096	3709	-	repeat	
cry-chr1	cFSAR-1-4	1,454,212	1,457,920	3709	-	repeat	
cry-chr1	cFSAR-2-2	1,458,051	1,462,294	4244	-	repeat	
cry-chr1	cFSAR-2-3	1,462,318	1,466,560	4243	-	repeat	
cry-chr1	cFSAR-3-2	1,466,597	1,468,206	1610	-	repeat	
cry-chr1	cHR-18	1,468,207	1,469,451	1245	+	repeat	
cry-chr1	cHR-15	1,470,383	1,472,704	2322	+	repeat	
cry-chr1	cTAR-5	1,475,228	1,476,619	1392	-	repeat	
cry-chr1	cTAR-4v	1,476,620	1,477,077	458	-	repeat	
cry-chr1	cTAR-13	1,477,668	1,479,324	1657	-	repeat	
cry-chr1	cTAR-12	1,479,325	1,481,832	2508	-	repeat	
cry-chr1	cTAR-11p	1,481,833	1,482,586	754	-	repeat	
cry-chr1	cTAR-4vp	1,482,741	1,483,022	282	+	repeat	
cry-chr1	rpb1 SPOG_00477	1,485,710	1,490,000	4291	-	gene	
cry-chr2	tip41 SPOG_00883	2,721,410	2,722,177	768	+	gene	
cry-chr2	SPOG_00884	2,722,775	2,723,620	846	+	gene	
cry-chr2	SPOG_00885	2,723,885	2,724,201	317	-	gene	
cry-chr2	SPOG_00886	2,724,541	2,725,464	924	+	gene	
cry-chr2	cHR-16	2,727,478	2,728,027	550	-	repeat	
cry-chr2	cTAR-7	2,728,033	2,729,914	1882	+	repeat	
cry-chr2	cTAR-9	2,729,916	2,730,864	948	+	repeat	
cry-chr2	cTAR-6	2,730,874	2,731,593	719	+	repeat	
cry-chr2	cHR-20	2,731,594	2,732,966	1373	+	repeat	
cry-chr2	gagpol-p	2,732,967	2,733,295	329	+	repeat	
cry-chr2	cTAR-5	2,733,495	2,734,756	1262	-	repeat	
cry-chr2	cTAR-4	2,734,757	2,735,361	605	-	repeat	
cry-chr2	cHR-16p	2,735,966	2,736,292	327	+	repeat	
cry-chr2	cFSAR-1-5	2,736,409	2,740,137	3729	-	repeat	
cry-chr2	cFSAR-2-4	2,740,268	2,744,527	4260	-	repeat	
cry-chr2	cFSAR-3p-3	2,744,536	2,746,194	1659	-	repeat	
cry-chr2	cHR-18	2,746,195	2,747,844	1650	+	repeat	
cry-chr2	cTAR-10	2,747,845	2,749,157	1313	+	repeat	
cry-chr2	cTAR-14	2,749,158	2,755,381	6224	-	repeat	
cry-chr2	cHR-19	2,755,382	2,757,269	1888	+	repeat	
cry-chr2	c-imr2L	2,757,270	2,760,621	3352	+	repeat	
cry-chr2	cTAR-4	2,757,271	2,757,699	429	+	repeat	
cry-chr2	cTAR-5	2,757,700	2,759,088	1389	+	repeat	
cry-chr2	cTAR-8	2,759,302	2,760,076	775	+	repeat	
cry-chr2	c-cnt2	2,760,622	2,771,156	10,535	+	repeat	
cry-chr2	cCNT-L	2,760,812	2,766,778	5967	-	repeat	
cry-chr2	cCNT-S	2,768,136	2,769,479	1344	-	repeat	
cry-chr2	c-imr2R	2,771,157	2,774,508	3352	-	repeat	
cry-chr2	cTAR-8	2,771,702	2,772,476	775	-	repeat	
cry-chr2	cTAR-5	2,772,690	2,774,078	1389	-	repeat	
cry-chr2	cTAR-4	2,774,079	2,774,507	429	-	repeat	
cry-chr2	cHR-19	2,774,509	2,776,396	1888	-	repeat	
cry-chr2	cTAR-14	2,776,397	2,782,614	6218	+	repeat	
cry-chr2	cTAR-10	2,782,615	2,783,927	1313	-	repeat	
cry-chr2	cHR-18	2,783,928	2,785,577	1650	-	repeat	
cry-chr2	cFSAR-3-4	2,785,578	2,787,220	1643	+	repeat	
cry-chr2	cHR-16p	2,787,430	2,788,062	633	-	repeat	
cry-chr2	cTAR-5	2,788,702	2,790,072	1371	-	repeat	
cry-chr2	cTAR-4	2,790,073	2,790,501	429	-	repeat	
cry-chr2	cHR-19p	2,790,503	2,791,525	1023	-	repeat	
cry-chr2	cHR-21	2,790,805	2,791,524	720	-	repeat	
cry-chr2	cHR-15	2,791,526	2,795,172	3647	-	repeat	
cry-chr2	cTAR-11p	2,795,543	2,796,381	839	-	repeat	
cry-chr2	cTAR-4v	2,796,515	2,796,951	437	+	repeat	
cry-chr2	cTAR-5p	2,797,322	2,797,821	500	+	repeat	
cry-chr2	cHR-16p	2,799,301	2,799,933	633	+	repeat	
cry-chr2	cFSAR-3-5	2,800,197	2,801,835	1639	-	repeat	
cry-chr2	SPOG_05758	2,802,504	2,802,876	373	+	gene	
cry-chr2	rad50 SPOG_2740	2,803,090	2,807,117	4028	-	gene	
cry-chr3	chk1 SPOG_03194	910,001	910,466	466	+	gene	
cry-chr3	cTAR-12	911,204	913,718	2515	+	repeat	
cry-chr3	cTAR-13	913,902	915,551	1650	+	repeat	
cry-chr3	cTAR-4	916,128	916,632	505	+	repeat	
cry-chr3	cTAR-5	916,633	918,014	1382	+	repeat	
cry-chr3	cHR-16	918,016	919,371	1356	+	repeat	
cry-chr3	cFSAR-1-6	919,488	923,194	3707	-	repeat	
cry-chr3	cFSAR-2-5	923,325	927,568	4244	-	repeat	
cry-chr3	cFSAR-3-6	927,577	929,236	1660	-	repeat	
cry-chr3	cTAR-6	929,243	929,969	727	-	repeat	
cry-chr3	cTAR-9	929,970	930,904	935	-	repeat	
cry-chr3	cHR-17	930,905	932,404	1500	+	repeat	
cry-chr3	cTAR-14	932,405	938,616	6212	+	repeat	
cry-chr3	cHR-15	938,763	942,551	3789	+	repeat	
cry-chr3	cHR-19p	942,553	943,273	721	+	repeat	
cry-chr3	c-imr3L	943,274	947,980	4707	+	repeat	
cry-chr3	LTRp	943,312	943,583	272	+	partial LTR	
cry-chr3	c-cnt3	947,981	951,902	3922	+	repeat	
cry-chr3	c-imr3R	951,903	956,609	4707	-	repeat	
cry-chr3	LTRp	956,300	956,571	272	-	partial LTR	
cry-chr3	cHR-19p	956,610	957,330	721	-	repeat	
cry-chr3	cHR-15	957,332	961,120	3789	-	repeat	
cry-chr3	cTAR-13	962,481	964,130	1650	-	repeat	
cry-chr3	cTAR-4	964,707	965,241	535	+	repeat	
cry-chr3	cTAR-5	965,242	966,625	1384	+	repeat	
cry-chr3	cHR-20	967,145	968,500	1356	-	repeat	
cry-chr3	cTAR-6	968,501	969,222	722	-	repeat	
cry-chr3	cTAR-9	969,232	970,161	930	-	repeat	
cry-chr3	cTAR-7	970,163	972,467	2305	-	repeat	
cry-chr3	cFSAR-1-7	974,118	977,823	3706	-	repeat	
cry-chr3	cFSAR-2-6	977,954	982,138	4185	-	repeat	
cry-chr3	cFSAR-3-7	982,172	983,811	1640	-	repeat	
cry-chr3	cHR-18p	983,812	984,507	696	+	repeat	
cry-chr3	cTAR-11	985,020	986,609	1590	-	repeat	
cry-chr3	cTAR-4	986,743	987,093	351	+	repeat	
cry-chr3	per1 SPOG_04755	987,829	989,556	1728	-	gene	
cry-chr3	SPOG_04754	990,540	991,264	725	-	gene	
cry-chr3	SPOG_4753	991,787	992,027	241	-	gene	
p denotes partial repeat element							
v denotes variant repeat element							

Supplementary Table 6: S. octosporus centromere repeat coordinates

chromosome	feature	start	end	size	strand	type
oct-chr1	per1 SOCG_03136	3,306,097	3,306,629	533	+	gene
oct-chr1	oTAR-4	3,307,008	3,307,469	462	-	repeat
oct-chr1	oTAR-11	3,307,595	3,309,952	2358	+	repeat
oct-chr1	oFSAR-3-1	3,309,953	3,310,543	591	+	repeat
oct-chr1	oFSAR-2-1	3,310,713	3,312,996	2284	+	repeat
oct-chr1	oFSAR-1-1	3,313,225	3,316,090	2866	+	repeat
oct-chr1	oTAR-4	3,316,123	3,316,600	478	-	repeat
oct-chr1	oHR-23	3,317,091	3,318,909	1819	-	repeat
oct-chr1	oHR-22	3,319,280	3,319,726	447	+	repeat
oct-chr1	oTAR-5	3,320,010	3,321,349	1340	+	repeat
oct-chr1	oHR-24	3,321,350	3,323,280	1931	-	repeat
oct-chr1	oTAR-6	3,323,421	3,323,773	353	-	repeat
oct-chr1	oTAR-9	3,323,774	3,324,667	894	-	repeat
oct-chr1	oHR-17	3,325,002	3,325,981	980	+	repeat
oct-chr1	oTAR-14	3,325,982	3,329,505	3524	+	repeat
oct-chr1	oHR-15	3,329,586	3,331,584	1999	+	repeat
oct-chr1	o-imr1L	3,331,585	3,336,231	4647	+	repeat
oct-chr1	o-cnt1	3,336,232	3,339,409	3178	+	repeat
oct-chr1	o-imr1R	3,339,410	3,344,057	4648	-	repeat
oct-chr1	oHR-15	3,344,058	3,345,925	1868	-	repeat
oct-chr1	oHR-26	3,346,079	3,346,632	554	-	repeat
oct-chr1	oFSAR-3-2	3,346,842	3,347,430	589	+	repeat
oct-chr1	oFSAR-2-2	3,347,606	3,349,886	2281	+	repeat
oct-chr1	oFSAR-1p-2	3,350,095	3,351,258	1164	+	repeat
oct-chr1	oTAR-13	3,352,574	3,355,073	2500	+	repeat
oct-chr1	oTAR-4	3,355,194	3,355,655	462	+	repeat
oct-chr1	oTAR-5	3,355,817	3,357,165	1349	+	repeat
oct-chr1	oHR-24	3,357,166	3,359,090	1925	-	repeat
oct-chr1	oTAR-6	3,359,091	3,359,583	493	-	repeat
oct-chr1	oTAR-9	3,359,584	3,360,481	898	-	repeat
oct-chr1	oTAR-7	3,360,483	3,362,978	2496	-	repeat
oct-chr1	oFSAR-1v-3	3,362,979	3,364,249	1271	-	repeat
oct-chr1	oFSAR-2-3	3,364,458	3,366,754	2297	-	repeat
oct-chr1	oFSAR-3-3	3,366,926	3,367,516	591	-	repeat
oct-chr1	oTAR-11	3,367,517	3,369,862	2346	-	repeat
oct-chr1	oTAR-4	3,369,988	3,370,450	463	+	repeat
oct-chr1	oTAR-5p	3,370,612	3,371,381	770	+	repeat
oct-chr1	oHR-25	3,372,576	3,373,051	476	+	repeat
oct-chr1	emc5 SOCG_05070	3,373,418	3,374,803	1386	-	gene
oct-chr1	rad50 SOCG_03135	3,374,191	3,377,063	2873	+	gene
oct-chr2	rpb1 SOCG_00001	2,566,950	2,572,118	5169	+	gene
oct-chr2	oTAR-4	2,574,485	2,574,947	463	-	repeat
oct-chr2	oTAR-11	2,575,073	2,577,411	2339	+	repeat
oct-chr2	oFSAR-3-4	2,577,412	2,578,018	607	+	repeat
oct-chr2	oFSAR-2-4	2,578,190	2,580,518	2329	+	repeat
oct-chr2	o-FSAR-2	2,580,748	2,583,077	2330	+	repeat
oct-chr2	oFSAR-2-5	2,583,313	2,585,626	2314	+	repeat
oct-chr2	oFSAR-2-6	2,585,896	2,588,283	2388	+	repeat
oct-chr2	oFSAR-2-7	2,588,457	2,590,737	2281	+	repeat
oct-chr2	oFSAR-2-8	2,590,955	2,593,285	2331	+	repeat

oct-chr2	oFSAR-2-9	2,593,521	2,595,870	2350	+	repeat
oct-chr2	oFSAR-2-10	2,596,140	2,598,456	2317	+	repeat
oct-chr2	oFSAR-2-11	2,598,692	2,600,960	2269	+	repeat
oct-chr2	oFSAR-1v-4	2,601,200	2,602,485	1286	+	repeat
oct-chr2	oTAR-7	2,602,486	2,604,619	2134	+	repeat
oct-chr2	o-imr2L	2,604,620	2,605,893	1274	+	repeat
oct-chr2	oTAR-9	2,604,620	2,605,525	906	+	repeat
oct-chr2	o-cnt2	2,605,894	2,616,081	10,188	+	repeat
oct-chr2	oCNT-S	2,606,711	2,607,948	1238	-	repeat
oct-chr2	oCNT-L	2,607,974	2,614,343	6370	+	repeat
oct-chr2	o-imr2R	2,616,082	2,617,355	1274	-	repeat
oct-chr2	oTAR-9	2,616,450	2,617,355	906	-	repeat
oct-chr2	oHR-17	2,617,356	2,618,329	974	+	repeat
oct-chr2	oTAR-14	2,618,330	2,621,873	3544	+	repeat
oct-chr2	oHR-15	2,621,955	2,623,970	2016	+	repeat
oct-chr2	oHR-23	2,624,138	2,625,950	1813	+	repeat
oct-chr2	oTAR-4	2,626,442	2,626,934	493	+	repeat
oct-chr2	oFSAR-1-5	2,626,963	2,629,836	2874	-	repeat
oct-chr2	oFSAR-2-12	2,630,047	2,632,350	2304	-	repeat
oct-chr2	oFSAR-2-13	2,632,524	2,634,851	2328	-	repeat
oct-chr2	oFSAR-3-5	2,635,023	2,635,620	598	-	repeat
oct-chr2	oHR-15p	2,636,534	2,637,512	979	+	repeat
oct-chr2	oHR-22	2,637,534	2,637,991	458	+	repeat
oct-chr2	oTAR-5	2,641,921	2,643,287	1367	-	repeat
oct-chr2	oTAR-4	2,643,478	2,643,931	454	-	repeat
oct-chr2	oTAR-11p	2,644,056	2,644,641	586	+	repeat
oct-chr2	oTAR-12	2,644,642	2,647,319	2678	+	repeat
oct-chr2	oTAR-13	2,647,320	2,649,817	2498	+	repeat
oct-chr2	oTAR-4	2,649,888	2,650,341	454	+	repeat
oct-chr2	oTAR-5p	2,650,532	2,651,461	930	+	repeat
oct-chr2	SO_04925	2,652,629	2,653,395	767	+	gene
oct-chr2	rec6 SPOG_04878	2,653,417	2,654,052	636	-	gene
oct-chr3	mid1 SOCG_01151	1,777,969	1,779,872	1904	+	gene
oct-chr3	tip41 SOCG_01150	1,779,875	1,781,112	1238	+	gene
oct-chr3	cbp3 SOCG_01149	1,781,151	1,782,275	1125	+	gene
oct-chr3	cmc2 SOCG_01148	1,782,185	1,782,942	758	-	gene
oct-chr3	ppc1 SOCG_01147	1,782,884	1,784,251	1368	+	gene
oct-chr3	oFSAR-1p	1,785,370	1,785,700	331	+	repeat
oct-chr3	oTAR-7	1,785,822	1,788,321	2500	+	repeat
oct-chr3	oTAR-9	1,788,326	1,789,230	905	+	repeat
oct-chr3	oTAR-6	1,789,231	1,789,724	494	+	repeat
oct-chr3	oHR-25	1,790,007	1,790,467	461	+	repeat
oct-chr3	oTAR-13p	1,790,697	1,791,000	304	+	repeat
oct-chr3	oTAR-4	1,791,072	1,791,534	463	+	repeat
oct-chr3	oTAR-5	1,791,714	1,793,051	1338	+	repeat
oct-chr3	oTAR-10	1,793,052	1,794,395	1344	+	repeat
oct-chr3	oTAR-14	1,794,397	1,797,791	3395	-	repeat
oct-chr3	oTAR-14-extended	1,797,792	1,804,813	7022	+	repeat
oct-chr3	cHR-19	1,804,815	1,806,814	2000	+	repeat
oct-chr3	oTAR-4	1,806,815	1,807,283	469	+	repeat
oct-chr3	oFSAR-1-6	1,807,313	1,810,187	2875	-	repeat

oct-chr3	oFSAR-2-14	1,810,416	1,812,645	2230	-	repeat
oct-chr3	oFSAR-3-6	1,812,821	1,813,411	591	-	repeat
oct-chr3	cTAR-11p	1,813,411	1,813,763	353	-	repeat
oct-chr3	o-imr3L	1,813,412	1,815,675	2264	+	repeat
oct-chr3	oTAR-5	1,813,684	1,814,395	712	+	repeat
oct-chr3	oTAR-8	1,814,945	1,815,260	316	+	repeat
oct-chr3	o-cnt3	1,815,676	1,827,078	11,403	+	repeat
oct-chr3	oCNT-S	1,815,912	1,817,149	1238	+	repeat
oct-chr3	oCNT-L	1,817,194	1,823,569	6376	+	repeat
oct-chr3	c-imr3R	1,827,079	1,829,348	2270	-	repeat
oct-chr3	oTAR-8	1,827,491	1,827,806	316	-	repeat
oct-chr3	oTAR-5p	1,828,356	1,829,067	712	-	repeat
oct-chr3	cTAR-11p	1,828,988	1,829,349	362	+	repeat
oct-chr3	oFSAR-3-7	1,829,349	1,829,939	591	+	repeat
oct-chr3	oFSAR-2-15	1,830,100	1,832,401	2302	+	repeat
oct-chr3	oFSAR-2-16	1,832,564	1,834,843	2280	+	repeat
oct-chr3	oFSAR-1-7	1,835,072	1,837,960	2889	+	repeat
oct-chr3	oTAR-4	1,837,990	1,838,458	469	-	repeat
oct-chr3	oHR-19	1,838,459	1,840,458	2000	-	repeat
oct-chr3	oTAR-14-extended	1,840,459	1,847,481	7023	+	repeat
oct-chr3	cTAR-14	1,847,482	1,850,874	3393	+	repeat
oct-chr3	oTAR-10	1,850,876	1,852,219	1344	-	repeat
oct-chr3	oTAR-5	1,852,220	1,853,606	1387	-	repeat
oct-chr3	oTAR-4	1,853,787	1,854,285	499	-	repeat
oct-chr3	oTAR-13	1,854,357	1,856,862	2506	-	repeat
oct-chr3	oTAR-12	1,856,863	1,859,534	2672	-	repeat
oct-chr3	oTAR-11p	1,859,535	1,859,781	247	-	repeat
oct-chr3	chk1 SOCG_04034	1,860,011	1,862,233	2223	-	gene
p denotes partial repeat element						
v denotes variant repeat element						

Supplementary Table 7: S. pombe centromere repeat annotation

Feature	size (kb)	% GC
cc1	4.0	29.1
cc2	6.6	28.9
cc3	4.7	29.0
cc1 & cc3 homology region (TM element)	3.2	28.0
dg	4.1-4.5	33-34
dh	4.1-6.7	32-34
dh variant	2.1	36.0
dh partial	0.3-0.4	29-37
imr1	5.1	28.9
imr2	4.1	29.1
imr3	5.9	30.8
imr1 partial	0.7	28.9
imr2 partial	0.8-0.9	29.8
imr3 partial	1.1	33.3

Supplementary Table 8: S. pombe centromere repeat coordinates

chromosome	feature	start*	end	size	strand	type
pom-chr1	tRNA-PheGAA	3695781	3695854	690	-	tRNA
pom-chr1	dh	3696860	3702324	5464	-	repeat
pom-chr1	dg	3702332	3706599	4267	+	repeat
pom-chr1	dh partial	3706600	3706981	381	+	repeat
pom-chr1	imr1L	3706975	3712076	5101	+	repeat
pom-chr1	tRNA-AlaAGC	3708122	3708196	571	-	tRNA
pom-chr1	tRNA-GluCTC	3708521	3708588	321	-	tRNA
pom-chr1	tRNA-IleAAT	3711299	3711373	650	+	tRNA
pom-chr1	cc1	3712076	3716110	4034	+	central core
pom-chr1	cc3 homolgy (TM element)	3712554	3715778	3224	+	repeat
pom-chr1	imr1R	3716110	3721232	5122	-	repeat
pom-chr1	tRNA-IleAAT	3716809	3716883	650	-	tRNA
pom-chr1	tRNA-GluCTC	3719603	3719673	476	+	tRNA
pom-chr1	tRNA-AlaAGC	3720001	3720075	571	+	tRNA
pom-chr1	dh partial	3721226	3721614	388	-	repeat
pom-chr1	dg	3721614	3725702	4088	-	repeat
pom-chr1	dh	3725725	3732297	6572	+	repeat
pom-chr2	tRNA-AsnGTT	1595753	1595826	552	-	tRNA
pom-chr2	tRNA-MetCAT	1595830	1595910	573	-	tRNA
pom-chr2	tRNA-TyrGTA	1596265	1596349	623	+	tRNA
pom-chr2	chrII-repeat1	1596266	1597808	1542	+	repeat
pom-chr2	tRNA-LeuCAA	1596495	1596595	409	-	tRNA
pom-chr2	tRNA-GlyGCC	1597078	1597147	506	-	tRNA
pom-chr2	tRNA-LysCTT	1597436	1597519	747	+	tRNA
pom-chr2	imr2 partial	1597808	1598739	931	+	repeat
pom-chr2	tRNA-IleAAT	1597810	1597884	650	+	tRNA
pom-chr2	tRNA-AlaAGC	1597949	1598023	571	-	tRNA
pom-chr2	tRNA-ValAAC	1598648	1598731	585	-	tRNA
pom-chr2	tRNA-chrII-repeat2	1598739	1599360	621	+	repeat
pom-chr2	tRNA-GluTTC	1599158	1599228	364	-	tRNA
pom-chr2	imr3 partial	1599844	1600951	1107	-	repeat
pom-chr2	tRNA-ArgACG	1599844	1599916	590	+	tRNA
pom-chr2	tRNA-AspGTC	1599995	1600065	273	-	tRNA
pom-chr2	imr1 partial	1601337	1602081	744	-	repeat
pom-chr2	dh partial	1602075	1602427	352	-	repeat
pom-chr2	dg	1602346	1606637	4291	-	repeat
pom-chr2	dh	1606638	1613364	6726	-	repeat
pom-chr2	imr2L	1613364	1617481	4117	+	repeat
pom-chr2	tRNA-IleAAT	1615062	1615136	650	+	tRNA
pom-chr2	tRNA-AlaAGC	1615201	1615275	571	-	tRNA
pom-chr2	tRNA-ValAAC	1615890	1615973	585	-	tRNA
pom-chr2	cc2	1617481	1624115	6634	+	central core
pom-chr2	imr2R	1624115	1628237	4122	-	repeat

pom-chr2	tRNA-ValAAC	1625620	1625701	405	+	tRNA
pom-chr2	tRNA-AlaAGC	1626323	1626397	571	+	tRNA
pom-chr2	tRNA-IleAAT	1626462	1626536	650	-	tRNA
pom-chr2	dh	1628237	1634576	6339	+	repeat
pom-chr2	dg	1634577	1639026	4449	+	repeat
pom-chr2	chrII-repeat1	1639026	1641350	2324	+	repeat
pom-chr2	tRNA-TyrGTA	1639824	1639908	623	+	tRNA
pom-chr2	tRNA-LeuCAA	1640052	1640152	413	-	tRNA
pom-chr2	tRNA-GlyGCC	1640634	1640704	550	-	tRNA
pom-chr2	tRNA-LysCTT	1640990	1641073	629	+	tRNA
pom-chr2	imr2 partial	1641350	1642261	911	+	repeat
pom-chr2	tRNA-IleAAT	1641352	1641426	650	+	tRNA
pom-chr2	tRNA-AlaAGC	1641491	1641565	571	-	tRNA
pom-chr2	tRNA-ValAAC	1642170	1642253	585	-	tRNA
pom-chr2	chrII-repeat2	1642261	1642863	602	+	repeat
pom-chr2	tRNA-ArgACG	1643100	1643171	528	-	tRNA
pom-chr3	tRNA-AlaAGC	1032183	1032257	571	-	tRNA
pom-chr3	imr2 partial	1032184	1032971	787	+	repeat
pom-chr3	tRNA-ValAAC	1032883	1032964	405	-	tRNA
pom-chr3	tRNA-SerAGA	1033721	1033803	797	+	tRNA
pom-chr3	tRNA-ArgTCG	1034177	1034252	483	-	tRNA
pom-chr3	tRNA-AspGTC	1035103	1035174	413	+	tRNA
pom-chr3	tRNA-ArgACG	1035248	1035320	590	-	tRNA
pom-chr3	tRNA-LeuAAG	1036213	1036291	277	+	tRNA
pom-chr3	chrIII-repeat(overlap with IRC3-I	1036925	1039161	2236	+	repeat
pom-chr3	tRNA-LysCTT	1038106	1038189	747	+	tRNA
pom-chr3	dh partial	1039161	1039574	413	-	repeat
pom-chr3	imr1 partial	1039652	1040291	639	-	repeat
pom-chr3	dh partial	1040292	1040643	351	-	repeat
pom-chr3	dg	1040643	1045143	4500	-	repeat
pom-chr3	dh	1045144	1047186	2042	-	repeat
pom-chr3	dg	1047186	1051687	4501	-	repeat
pom-chr3	dh	1051688	1053743	2055	-	repeat
pom-chr3	dg	1053743	1058244	4501	-	repeat
pom-chr3	dh	1058245	1060286	2041	-	repeat
pom-chr3	dg	1060286	1064777	4491	-	repeat
pom-chr3	dh	1064778	1068911	4133	-	repeat
pom-chr3	imr3L	1068911	1074778	5867	+	repeat
pom-chr3	tRNA-AspGTC	1070268	1070337	380	+	tRNA
pom-chr3	tRNA-ArgACG	1070417	1070489	590	-	tRNA
pom-chr3	tRNA-ValAAC	1070750	1070831	405	+	tRNA
pom-chr3	tRNA-ThrAGT	1070893	1070965	796	-	tRNA
pom-chr3	tRNA-LeuCAA	1073808	1073908	409	-	tRNA
pom-chr3	cc1 homology (TM element)	1074469	1077686	3217	-	repeat

pom-chr3	cc3	1074778	1079503	4725	+	central core
pom-chr3	tRNA-GluCTC	1079024	1079093	440	+	tRNA
pom-chr3	imr3R	1079503	1085367	5864	-	repeat
pom-chr3	tRNA-LeuCAA	1080370	1080471	565	+	tRNA
pom-chr3	tRNA-ThrAGT	1083315	1083387	796	+	tRNA
pom-chr3	tRNA-ValAAC	1083449	1083530	405	-	tRNA
pom-chr3	tRNA-ArgACG	1083791	1083864	714	+	tRNA
pom-chr3	tRNA-AspGTC	1083943	1084014	360	-	tRNA
pom-chr3	dh	1085367	1089513	4146	+	repeat
pom-chr3	dg	1089514	1094030	4516	+	repeat
pom-chr3	dh	1094030	1096077	2047	+	repeat
pom-chr3	dg	1096078	1100579	4501	+	repeat
pom-chr3	dh	1100579	1102625	2046	+	repeat
pom-chr3	dg	1102626	1107138	4512	+	repeat
pom-chr3	dh	1107138	1109183	2045	+	repeat
pom-chr3	dg	1109184	1113683	4499	+	repeat
pom-chr3	dh	1113683	1115718	2035	+	repeat
pom-chr3	dg	1115719	1120196	4477	+	repeat
pom-chr3	dh	1120196	1122227	2031	+	repeat
pom-chr3	dg	1122228	1126724	4496	+	repeat
pom-chr3	dh	1126724	1128769	2045	+	repeat
pom-chr3	dg	1128770	1133273	4503	+	repeat
pom-chr3	dh	1133273	1135316	2043	+	repeat
pom-chr3	dg	1135317	1139845	4528	+	repeat
pom-chr3	dh	1139845	1141890	2045	+	repeat
pom-chr3	dg	1141891	1146392	4501	+	repeat
pom-chr3	dh partial	1146392	1146745	353	+	repeat
pom-chr3	imr1 partial	1146745	1147384	639	+	repeat
pom-chr3	dh partial	1147461	1147875	414	+	repeat
pom-chr3	chrIII-repeat(overlap with IRC3-I	1147875	1150110	2235	-	repeat
pom-chr3	tRNA-LysCTT	1148846	1148929	747	-	tRNA
pom-chr3	tRNA-PheGAA	1151459	1151532	690	-	tRNA
* co-ordinates are for genome assembled from nanopor						

Supplementary Table 9: S. japonicus retrotransposons

Retrotransposon	Present at putative centromeres?	Heterochromatin	CENP-A
Tj1	rare	low	-
Tj2	common	High	-
Tj3*	common	High	-
Tj4	possibly telomere specific	Intermediate	-
Tj5*	common	High	-
Tj6	common	Low-intermediate	Intermediate
Tj7	common	Low-intermediate	High Partials have low levels
Tj8*	common	High	-
Tj9*	common	High	-
Tj10	common	High	-
Tj11**	common	High Partials have low levels	-

*Tj3, Tj5, Tj8 and Tj9 have high homology.
**Newly defined putative retrotransposon. 5501 bp. Present in Schizosaccharomyces_japonicus.GCA_000149845.2 supercontig 5.5: 40010-45479 (Rhind, 2011)

Supplementary Table 10: S. cryophilus rDNA annotation

chromosome	feature	start	end	strand	
cry-chr1	18S rRNA	1	1831	-	subtelomeric
cry-chr1	28S rRNA	6775	10311	-	subtelomeric
cry-chr1	5.8S rRNA	10563	10806	-	subtelomeric
cry-chr1	18S rRNA	11047	12926	-	subtelomeric
cry-chr1	5S rRNA	16694	16808	-	subtelomeric
cry-chr1	5S rRNA	92021	92138	-	
cry-chr1	5S rRNA	687354	687468	-	
cry-chr1	5S rRNA	960699	960813	-	
cry-chr1	5S rRNA	1113484	1113598	-	
cry-chr1	5S rRNA	1182255	1182369	-	
cry-chr1	5S rRNA	1184319	1184433	-	
cry-chr1	5S rRNA	1227221	1227335	-	
cry-chr1	5S rRNA	1404739	1404853	+	cen1
cry-chr1	5S rRNA	1409012	1409126	+	cen1
cry-chr1	5S rRNA	1412834	1412948	+	cen1
cry-chr1	5S rRNA	1446449	1446563	-	cen1
cry-chr1	5S rRNA	1450273	1450387	-	cen1
cry-chr1	5S rRNA	1454097	1454211	-	cen1
cry-chr1	5S rRNA	1457921	1458035	-	cen1
cry-chr1	5S rRNA	1462188	1462302	-	cen1
cry-chr1	5S rRNA	1466454	1466568	-	cen1
cry-chr1	5S rRNA	1658235	1658349	+	
cry-chr1	5S rRNA	2717711	2717825	+	
cry-chr1	5S rRNA	2719711	2719831	+	
cry-chr1	5S rRNA	2975140	2975254	-	
cry-chr1	5S rRNA	3015049	3015163	+	
cry-chr1	5S rRNA	3024039	3024153	+	
cry-chr1	5S rRNA	3028464	3028578	+	
cry-chr1	5S rRNA	3032071	3032185	-	
cry-chr1	5S rRNA	3226822	3226936	+	
cry-chr1	5S rRNA	3263147	3263261	+	
cry-chr1	5S rRNA	3706807	3706921	+	
cry-chr1	5S rRNA	3734750	3734864	+	
cry-chr1	5S rRNA	3818916	3819030	+	
cry-chr1	5S rRNA	3821677	3821797	+	
cry-chr1	5S rRNA	4064005	4064119	-	
cry-chr1	5S rRNA	4068996	4069110	-	
cry-chr1	5S rRNA	4191287	4191407	-	
cry-chr1	5S rRNA	4193289	4193403	-	
cry-chr1	5S rRNA	4418514	4418634	+	
cry-chr1	5S rRNA	4752278	4752392	-	
cry-chr1	5S rRNA	4758226	4758340	-	
cry-chr1	5S rRNA	4792982	4793096	-	
cry-chr1	5S rRNA	4813584	4813698	-	
cry-chr1	5S rRNA	4826098	4826218	-	
cry-chr2	5S rRNA	47502	47622	+	
cry-chr2	5S rRNA	54858	54978	+	
cry-chr2	5S rRNA	65078	65198	-	
cry-chr2	5S rRNA	417166	417280	+	
cry-chr2	5S rRNA	993493	993607	+	
cry-chr2	5S rRNA	995540	995661	+	
cry-chr2	5S rRNA	999362	999481	+	
cry-chr2	5S rRNA	1307353	1307467	-	
cry-chr2	5S rRNA	1476899	1477013	+	
cry-chr2	5S rRNA	1657788	1657902	+	
cry-chr2	5S rRNA	1685149	1685263	-	
cry-chr2	5S rRNA	1875167	1875281	-	
cry-chr2	5S rRNA	2388795	2388917	-	
cry-chr2	5S rRNA	2390742	2390856	-	
cry-chr2	5S rRNA	2565353	2565467	+	
cry-chr2	5S rRNA	2736294	2736408	-	cen2
cry-chr2	5S rRNA	2740138	2740252	-	cen2
cry-chr2	5S rRNA	2744421	2744535	-	cen2
cry-chr2	5S rRNA	2787221	2787335	+	cen2
cry-chr2	5S rRNA	2800082	2800196	-	cen2
cry-chr2	5S rRNA	2887921	2888035	-	
cry-chr2	5S rRNA	3186877	3186991	-	
cry-chr2	5S rRNA	3605008	3605128	-	
cry-chr2	5S rRNA	3606943	3607057	-	
cry-chr2	5S rRNA	3761550	3761663	+	
cry-chr2	5S rRNA	3815190	3815304	-	
cry-chr2	5S rRNA	3817218	3817332	-	
cry-chr2	5S rRNA	3947351	3947465	-	
cry-chr2	5S rRNA	3969326	3969440	-	
cry-chr2	5S rRNA	3981017	3981131	+	subtelomeric
cry-chr2	18S rRNA	3984894	3986773	+	subtelomeric
cry-chr2	5.8S rRNA	3987015	3987258	+	subtelomeric
cry-chr2	28S rRNA	3987512	3991048	+	subtelomeric
cry-chr3	28S rRNA	4997	8533	-	subtelomeric
cry-chr3	5.8S rRNA	8786	9029	-	subtelomeric
cry-chr3	18S rRNA	9271	11150	-	subtelomeric
cry-chr3	5S rRNA	14922	15036	-	subtelomeric
cry-chr3	5S rRNA	104542	104656	-	
cry-chr3	5S rRNA	195762	195882	+	
cry-chr3	5S rRNA	697342	697462	+	
cry-chr3	5S rRNA	813052	813166	+	
cry-chr3	5S rRNA	814993	815107	+	
cry-chr3	5S rRNA	919373	919487	-	cen3
cry-chr3	5S rRNA	923195	923309	-	cen3
cry-chr3	5S rRNA	927462	927576	-	cen3
cry-chr3	5S rRNA	974003	974117	-	cen3
cry-chr3	5S rRNA	977824	977938	-	cen3
cry-chr3	5S rRNA	982032	982146	-	cen3
cry-chr3	5S rRNA	1034454	1034568	-	
cry-chr3	5S rRNA	1260050	1260163	-	
cry-chr3	5S rRNA	1268567	1268681	+	
cry-chr3	5S rRNA	1486900	1487020	-	
cry-chr3	5S rRNA	2115978	2116099	+	
cry-chr3	5S rRNA	2185939	2186053	-	
cry-chr3	5S rRNA	2549700	2549820	+	
cry-chr3	5S rRNA	2580186	2580300	+	
cry-chr3	5S rRNA	2633580	2633694	+	
cry-chr3	5S rRNA	2649988	2650102	-	
cry-chr3	5S rRNA	2663268	2663382	-	
cry-chr3	5S rRNA	2684675	2684795	-	
cry-chr3	5S rRNA	2740101	2740215	-	

Supplementary Table 11: S. octosporus rDNA annotation

chromosome	feature	start	end	strand
oct-chr1	5S rRNA	33235	33349	+
oct-chr1	5S rRNA	413920	414039	+
oct-chr1	5S rRNA	880727	880841	+
oct-chr1	5S rRNA	1078204	1078318	-
oct-chr1	5S rRNA	1080784	1080894	-
oct-chr1	5S rRNA	1264920	1265034	-
oct-chr1	5S rRNA	1270014	1270128	-
oct-chr1	5S rRNA	1276840	1276954	-
oct-chr1	5S rRNA	1787532	1787646	-
oct-chr1	5S rRNA	1794347	1794461	-
oct-chr1	5S rRNA	2031940	2032054	+
oct-chr1	5S rRNA	2829548	2829662	-
oct-chr1	5S rRNA	2964472	2964586	-
oct-chr1	5S rRNA	3029426	3029540	-
oct-chr1	5S rRNA	3258531	3258645	+
oct-chr1	5S rRNA	3310547	3310661	+
oct-chr1	5S rRNA	3313022	3313136	+
oct-chr1	5S rRNA	3347434	3347548	+
oct-chr1	5S rRNA	3349892	3350006	+
oct-chr1	5S rRNA	3364338	3364452	-
oct-chr1	5S rRNA	3366808	3366922	-
oct-chr1	5S rRNA	3373028	3373142	+
oct-chr1	5S rRNA	3456387	3456501	-
oct-chr1	5S rRNA	3753939	3754086	-
oct-chr1	5S rRNA	3756499	3756613	-
oct-chr1	5S rRNA	4164520	4164643	-
oct-chr1	5S rRNA	4167108	4167222	-
oct-chr1	5S rRNA	4361472	4361586	+
oct-chr1	18S rRNA	4514514	4516392	+
oct-chr1	5.8S rRNA	4516668	4516914	+
oct-chr1	28S rRNA	4517197	4519984	+
oct-chr2	5S rRNA	12977	13091	+
oct-chr2	5S rRNA	29055	29169	-
oct-chr2	5S rRNA	96705	96824	-
oct-chr2	5S rRNA	101070	101184	-
oct-chr2	5S rRNA	393309	393428	+
oct-chr2	5S rRNA	672545	672659	-
oct-chr2	5S rRNA	1041835	1041949	+
oct-chr2	5S rRNA	1114943	1115057	-
oct-chr2	5S rRNA	1571359	1571473	+
oct-chr2	5S rRNA	1739423	1739537	+
oct-chr2	5S rRNA	1957207	1957323	-
oct-chr2	5S rRNA	1999314	1999433	+
oct-chr2	5S rRNA	2249382	2249496	-
oct-chr2	5S rRNA	2400626	2400740	-
oct-chr2	5S rRNA	2578022	2578136	+
oct-chr2	5S rRNA	2580580	2580694	+
oct-chr2	5S rRNA	2583145	2583259	+
oct-chr2	5S rRNA	2585728	2585842	+
oct-chr2	5S rRNA	2588289	2588403	+
oct-chr2	5S rRNA	2590787	2590901	+
oct-chr2	5S rRNA	2593353	2593467	+
oct-chr2	5S rRNA	2595972	2596086	+
oct-chr2	5S rRNA	2598524	2598638	+
oct-chr2	5S rRNA	2601010	2601124	+
oct-chr2	5S rRNA	2629927	2630041	-
oct-chr2	5S rRNA	2632404	2632518	-
oct-chr2	5S rRNA	2634905	2635019	-
oct-chr2	5S rRNA	2896020	2896134	-
oct-chr2	5S rRNA	2963574	2963688	-
oct-chr2	5S rRNA	3120708	3120822	-
oct-chr2	5S rRNA	3607716	3607830	-
oct-chr2	5S rRNA	3610298	3610412	-
oct-chr2	5S rRNA	3618746	3618860	-
oct-chr2	18S rRNA	3789499	3791378	+
oct-chr2	5.8S rRNA	3791653	3791899	+
oct-chr2	28S rRNA	3792181	3794048	+
oct-chr3	5S rRNA	8156	8270	+
oct-chr3	5S rRNA	29384	29500	-
oct-chr3	5S rRNA	104712	104826	-
oct-chr3	5S rRNA	132256	132372	-
oct-chr3	5S rRNA	140190	140304	-
oct-chr3	5S rRNA	161992	162106	-
oct-chr3	5S rRNA	642970	643084	+
oct-chr3	5S rRNA	766517	766631	+
oct-chr3	5S rRNA	772371	772485	+
oct-chr3	5S rRNA	774887	775006	+
oct-chr3	5S rRNA	1005175	1005289	-
oct-chr3	5S rRNA	1083257	1083371	-
oct-chr3	5S rRNA	1107657	1107771	-
oct-chr3	5S rRNA	1114356	1114470	-
oct-chr3	5S rRNA	1560721	1560835	-
oct-chr3	5S rRNA	1600041	1600155	-
oct-chr3	5S rRNA	1623740	1623854	+
oct-chr3	5S rRNA	1810276	1810390	-
oct-chr3	5S rRNA	1812703	1812817	-
oct-chr3	5S rRNA	1829943	1830057	+
oct-chr3	5S rRNA	1832407	1832521	+
oct-chr3	5S rRNA	1834869	1834983	+
oct-chr3	5S rRNA	1952114	1952228	-
oct-chr3	5S rRNA	2029273	2029387	+
oct-chr3	5S rRNA	2894060	2894174	+
oct-chr3	5S rRNA	2952779	2952893	+
oct-chr3	5S rRNA	2955384	2955508	+
oct-chr3	18S rRNA	2966653	2968532	+
oct-chr3	5.8S rRNA	2968807	2969053	-
oct-chr3	28S rRNA	2969336	2971974	+

Supplementary Table 12: Homology between FSAR repeats within S. cryophilus and S. octosporus

S.cry FSAR identity %							
	cFSAR-1-1			cFSAR-2-1			cFSAR-3-1
cFSAR-1-1	100		cFSAR-2-1	100		cFSAR-3-1	100
cFSAR-1-2	99.95		cFSAR-2-2	99.67		cFSAR-3-2	98.7
cFSAR-1-3	99.95		cFSAR-2-3	99.6		cFSAR-3-3	93.95
cFSAR-1-4	99.95		cFSAR-2-4	99.44		cFSAR-3-4	95.13
cFSAR-1-7	97.57		cFSAR-2-5	96.83		cFSAR-3-5	99.51
cFSAR-1-5	95.82		cFSAR-2-6	98.78		cFSAR-3-6	95.26
cFSAR-1-6	95.66					cFSAR-3p-7	94
S.oct FSAR identity %							
	oFSAR-1-1			oFSAR-2-1			oFSAR-3-1
oFSAR-1-1	100		oFSAR-2-1	100		oFSAR-3-1	100
oFSAR-1-2	99.07		oFSAR-2-2	97.47		oFSAR-3-2	99.83
oFSAR-1-3	98.12		oFSAR-2-3	97.05		oFSAR-3-3	98.98
oFSAR-1-4	98.37		oFSAR-2-4	97		oFSAR-3-4	98.98
oFSAR-1p-5	97.25		oFSAR-2-5	98		oFSAR-3-5	98.98
oFSAR-1v-6	96.57		oFSAR-2-6	97.78		oFSAR-3-6	96.38
oFSAR-1v-7	95.75		oFSAR-2-7	95.38		oFSAR-3-7	97.66
			oFSAR-2-8	96.16			
			oFSAR-2-9	95.78			
			oFSAR-2-10	97.52			
			oFSAR-2-11	95.4			
			oFSAR-2-12	98.2			
			oFSAR-2-12	96.21			
			oFSAR-2-13	96.16			
			oFSAR-2-14	97.99			
			oFSAR-2-15	97.73			
			oFSAR-2-16	96.75			

Supplementary Table 13: S. cryophilus and S. octosporus Hsp16 gene annotation and coordinates

chromosome	hsp16 ORF	start	end	strand
cry-chr1	cry-ORF1-cFSAR-3-1	1403407	1403798	-
cry-chr1	cry-ORF2-cFSAR-2-1	1405538	1405959	-
cry-chr1	cry-ORF3-cFSAR-2-2	1461082	1461503	+
cry-chr1	cry-ORF4-cFSAR-2-3	1465348	1465769	+
cry-chr1	cry-ORF5-cFSAR-3-2	1467506	1467895	+
cry-chr1	cry-ORF6	2716245	2716666	-
cry-chr1	cry-ORF7	3738213	3738633	-
cry-chr1	cry-ORF8	4065105	4065526	+
cry-chr1	cry-ORF9	4787816	4788541	-
cry-chr1	cry-ORF10	4819528	4819949	+
cry-chr1	cry-ORF11	4876023	4876444	+
cry-chr2	cry-ORF12-cFSAR-2-4	2743315	2743736	+
cry-chr2	cry-ORF13-cFSAR-3p-3	2745491	2745883	+
cry-chr2	cry-ORF14-cFSAR-3-4	2785889	2786281	-
cry-chr2	cry-ORF15-cFSAR-3-5	2801168	2801560	+
cry-chr3	cry-ORF16-cFSAR-2-5	926356	926777	+
cry-chr3	cry-ORF17-cFSAR-3-6	928536	928925	+
cry-chr3	cry-ORF18-cFSAR-2-6	980964	981385	+
cry-chr3	cry-ORF19-cFSAR-3-7	983108	983500	+
cry-chr3	cry-ORF20	2539982	2540707	-
cry-chr3	cry-ORF21	2555876	2556297	-
cry-chr3	cry-ORF22	2706305	2706726	-
cry-chr3	cry-ORF23	2734174	2734595	+
cry-chr3	cry-ORF24	2754385	2754806	+
chromosome	**ORF**	**start**	**end**	**strand**
oct-chr1	oct-ORF1	32330	32773	-
oct-chr2	oct-ORF2	12072	12515	-
oct-chr2	oct-ORF3	101646	102089	+
oct-chr3	oct-ORF4	7253	7696	-
oct-chr3	oct-ORF5	66337	66780	-
oct-chr3	oct-ORF6	140764	141207	+
oct-chr3	oct-ORF7	162568	163011	+
oct-chr3	oct-ORF8	2892999	2893652	-

Supplementary Table 14: S. cryophilus retrotansposon and LTR annotation and coordinates

chromosome	start	end	Element	strand	size
cry-chr1	26090	28101	Tcry1 partial retrotransposon	-	2011
cry-chr1	30927	31247	Tcry1-type LTR	-	320
cry-chr1	30927	31288	Tcry1-type LTR	-	361
cry-chr1	965190	965585	Tcry1 partial retrotransposon	+	395
cry-chr1	965223	965585	Tcry1-type LTR	+	362
cry-chr1	2024308	2024682	Tcry1-type LTR	-	374
cry-chr1	2024308	2024714	Tcry1-type LTR	-	406
cry-chr1	4844387	4844681	Tcry1-type LTR	-	294
cry-chr1	4844387	4844714	Tcry1-type LTR	-	327
cry-chr2	1431930	1434972	Tcry1-type LTR	-	3042
cry-chr2	1434981	1436403	Tcry1-type LTR	-	1422
cry-chr3	24898	25200	Tcry1-type LTR	-	302
cry-chr3	24898	25238	Tcry1-type LTR	-	340
cry-chr3	943270	943583	Tcry1-type LTR	+	313
cry-chr3	943312	943583	Tcry1-type LTR	+	271
cry-chr3	956300	956571	Tcry1-type LTR	-	271
cry-chr3	956300	956613	Tcry1-type LTR	-	313
cry-chr3	2743314	2748368	Tcry1-1	-	5054

Supplementary Table 15: S. octosporus retrotransposon remnat annotation

	homology with	Tcry1			transposon remnant S.oct-mat			oTAR-14ex		
chromosome	**locus**	start	end	size	start	end	size	start	end	size
S.oct-chr1	**mat locus**	834081	835897	1816	834236	836137	1901	833988	835179	1191
					821484	821925	441	835889	836945	1056
S.oct-chr2	**retrotransposon homology**							761733	762238	505
S.oct-chr3	**retrotransposon homology**	134707	135239	532	134071	135804	1733	132948	134195	1247
								134834	136047	1213
	centromere: oTAR-14ex	1801780	1802233	453	1801805	1802703	898	1797792	1804814	7022
	centromere: oTAR-14ex	1842409	1843493	1084	1842570	1843468	898	1840459	1847481	7022

Supplementary Table 16: S. pombe, S. octosporus and S. cruophilus tDNA frequencies within centromeres and genome wide

species	% tDNAs at centromeres	mean frequency centromeres	mean frequency rest of genome	fold frequency centromeres vs genome
S. pombe	32 (55/171)	3.8 kb	105.6 kb	27.8
S. octosporus	32 (96/298)	2.3 kb	54.8 kb	23.6
S. cryophilus	32 (95/294)	2.6 kb	57.5 kb	22.5

Supplementary Table 17: S. cryophilus tDNA coordinates

Chromosome	tDNA-anticodon*	start	end	strand	cen located?
cry-chr1	GlyGCC	91817	91887	-	
cry-chr1	PheGAA	130564	130636	+	
cry-chr1	PheGAA	132151	132223	+	
cry-chr1	ArgACG	312388	312460	+	
cry-chr1	GlyGCC	505819	505889	-	
cry-chr1	GlyGCC	508234	508304	+	
cry-chr1	GlyGCC	509140	509210	-	
cry-chr1	SerGCT	535112	535206	-	
cry-chr1	ThrAGT	679941	680012	+	
cry-chr1	GlyGCC	687139	687209	-	
cry-chr1	LysTTT	690360	690434	+	
cry-chr1	IleTAT	731048	731146	+	
cry-chr1	MetCAT	916227	916298	-	
cry-chr1	SerTGA	916303	916399	-	
cry-chr1	LysCTT	949167	949249	-	
cry-chr1	ProAGG	949343	949414	-	
cry-chr1	SerAGA	959364	959445	-	
cry-chr1	MetCAT	983804	983884	+	
cry-chr1	AsnGTT	983891	983964	+	
cry-chr1	LysCTT	1000051	1000133	+	
cry-chr1	ProAGG	1000466	1000537	+	
cry-chr1	HisGTG	1034546	1034617	+	
cry-chr1	ThrTGT	1130756	1130827	-	
cry-chr1	ArgTCG	1161587	1161659	+	
cry-chr1	LysTTT	1222481	1222555	+	
cry-chr1	GlnTTG	1226243	1226314	-	
cry-chr1	HisGTG	1245191	1245262	-	
cry-chr1	HisGTG	1246203	1246274	+	
cry-chr1	ArgACG	1398356	1398428	-	cen1
cry-chr1	IleAAT	1398802	1398875	+	cen1
cry-chr1	AlaAGC	1399216	1399289	-	cen1
cry-chr1	ValAAC	1399733	1399815	+	cen1
cry-chr1	AspGTC	1400121	1400191	-	cen1
cry-chr1	AlaAGC	1401996	1402069	-	cen1
cry-chr1	ValAAC	1416701	1416783	+	cen1
cry-chr1	AspGTC	1417064	1417134	-	cen1
cry-chr1	GluTTC	1422294	1422365	-	cen1
cry-chr1	LeuCAA	1429970	1430074	-	cen1
cry-chr1	LysCTT	1430651	1430736	+	cen1
cry-chr1	LysCTT	1441670	1441755	-	cen1
cry-chr1	LeuCAA	1442332	1442436	+	cen1
cry-chr1	GluCTC	1442934	1443005	+	cen1
cry-chr1	PheGAA	1443462	1443534	-	cen1
cry-chr1	ArgTCG	1472926	1472998	+	cen1
cry-chr1	ProAGG	1474860	1474931	-	cen1
cry-chr1	ArgACG	1475230	1475302	-	cen1
cry-chr1	IleAAT	1475694	1475767	+	cen1
cry-chr1	AlaAGC	1476104	1476177	-	cen1
cry-chr1	ValAAC	1476621	1476703	+	cen1
cry-chr1	AspGTC	1477020	1477090	-	cen1
cry-chr1	LeuAAG	1478388	1478466	-	cen1
cry-chr1	GluTTC	1479575	1479646	-	cen1
cry-chr1	AspGTC	1482589	1482659	+	cen1
cry-chr1	ValAAC	1482934	1483016	-	cen1
cry-chr1	AlaAGC	1483401	1483474	+	cen1
cry-chr1	ThrTGT	1566588	1566659	+	
cry-chr1	TrpCCA	1600889	1600961	+	
cry-chr1	LeuTAA	1662939	1663038	+	
cry-chr1	AlaAGC	1699268	1699341	+	
cry-chr1	IleAAT	1770660	1770733	+	
cry-chr1	ArgCCT	1879809	1879913	+	
cry-chr1	PheGAA	1925186	1925258	+	
cry-chr1	HisGTG	2071565	2071636	-	
cry-chr1	LeuCAA	2193818	2193922	-	
cry-chr1	GlnTTG	2275238	2275309	-	
cry-chr1	ValTAC	2328129	2328201	-	
cry-chr1	AsnGTT	2475236	2475309	-	
cry-chr1	GlyGCC	2475535	2475605	-	
cry-chr1	SerAGA	2477035	2477116	+	
cry-chr1	GluTTC	2542169	2542240	+	
cry-chr1	ProAGG	2970237	2970308	-	
cry-chr1	ProAGG	2970837	2970908	+	
cry-chr1	SerAGA	3014773	3014854	-	
cry-chr1	GlyGCC	3028768	3028838	+	
cry-chr1	SerAGA	3226546	3226627	-	
cry-chr1	TyrGTA	3562905	3562988	+	
cry-chr1	SerCGA	3586114	3586210	+	
cry-chr1	MetCAT	3586221	3586292	+	
cry-chr1	ValAAC	3630054	3630135	+	
cry-chr1	LeuAAG	3659732	3659810	-	
cry-chr1	SerGCT	3661564	3661658	+	
cry-chr1	AsnGTT	3780368	3780441	+	
cry-chr1	GluCTC	3851396	3851467	+	
cry-chr1	LysCTT	3958742	3958824	-	
cry-chr1	LysTTT	3973358	3973432	-	
cry-chr1	TyrGTA	3984988	3985071	+	
cry-chr1	IleAAT	4010738	4010811	+	
cry-chr1	GlnTTG	4060210	4060281	-	
cry-chr1	ThrCGT	4108849	4108920	-	
cry-chr1	LysCTT	4109634	4109716	-	
cry-chr1	AlaTGC	4115466	4115537	-	
cry-chr1	HisGTG	4193751	4193822	-	
cry-chr1	SerAGA	4227149	4227230	+	
cry-chr1	GlnTTG	4383616	4383687	+	
cry-chr1	ValCAC	4383894	4383965	+	
cry-chr1	ThrAGT	4394451	4394522	-	
cry-chr1	LeuTAA	4458580	4458679	+	
cry-chr1	GlnTTG	4669217	4669288	-	
cry-chr1	LysCTT	4669566	4669648	+	
cry-chr1	ArgTCT	4690337	4690409	+	
cry-chr1	GlyTCC	4690555	4690625	-	
cry-chr1	MetCAT	4747095	4747166	+	
cry-chr1	LeuAAG	4747410	4747488	-	

cry-chr1	ArgTCT	4762521	4762593	+	
cry-chr1	ProAGG	4802986	4803057	+	
cry-chr1	GlyTCC	4803434	4803504	+	
cry-chr1	GlyTCC	4864713	4864783	+	
cry-chr2	AsnGTT	105306	105379	-	
cry-chr2	ProAGG	105557	105628	+	
cry-chr2	AspGTC	215535	215605	+	
cry-chr2	AsnGTT	426497	426570	+	
cry-chr2	CysGCA	551460	551531	-	
cry-chr2	AlaCGC	563920	564002	+	
cry-chr2	ProTGG	617444	617515	+	
cry-chr2	GlyGCC	992813	992883	+	
cry-chr2	GlyGCC	999636	999706	+	
cry-chr2	LysTTT	1024900	1024974	-	
cry-chr2	LeuTAA	1062616	1062714	-	
cry-chr2	MetCAT	1075030	1075110	+	
cry-chr2	TrpCCA	1100377	1100449	+	
cry-chr2	SerAGA	1122870	1122951	-	
cry-chr2	GluCTC	1195689	1195760	+	
cry-chr2	ThrAGT	1206451	1206522	+	
cry-chr2	AsnGTT	1273506	1273579	-	
cry-chr2	MetCAT	1273586	1273668	-	
cry-chr2	GlnCTG	1327870	1327941	+	
cry-chr2	LysCTT	1351024	1351106	+	
cry-chr2	GlnCTG	1351701	1351772	+	
cry-chr2	GlnTTG	1352057	1352128	+	
cry-chr2	GlnTTG	1353859	1353930	+	
cry-chr2	GlyGCC	1398779	1398849	+	
cry-chr2	ThrAGT	1437157	1437228	+	
cry-chr2	GlyCCC	1486779	1486849	-	
cry-chr2	LeuAAG	1581580	1581658	+	
cry-chr2	ValCAC	1594163	1594234	-	
cry-chr2	PheGAA	1674504	1674576	-	
cry-chr2	GlyGCC	1683338	1683408	+	
cry-chr2	GlyGCC	1684016	1684086	-	
cry-chr2	TyrGTA	1793743	1793826	-	
cry-chr2	LeuCAG	1794580	1794673	+	
cry-chr2	MetCAT	1832550	1832630	+	
cry-chr2	GlyTCC	1858872	1858942	+	
cry-chr2	ProAGG	1862800	1862871	+	
cry-chr2	GluTTC	1893328	1893399	-	
cry-chr2	CysGCA	1893890	1893961	-	
cry-chr2	TyrGTA	1991231	1991314	-	
cry-chr2	ProCGG	1991852	1991951	+	
cry-chr2	AlaTGC	2105091	2105162	+	
cry-chr2	SerAGA	2155606	2155687	-	
cry-chr2	MetCAT	2190173	2190244	-	
cry-chr2	SerCGA	2190255	2190351	-	
cry-chr2	AlaAGC	2288920	2288993	+	
cry-chr2	GluTTC	2316793	2316864	-	
cry-chr2	ArgACG	2428841	2428913	-	
cry-chr2	ProTGG	2532325	2532396	-	
cry-chr2	SerAGA	2566313	2566394	+	
cry-chr2	LysCTT	2571329	2571411	+	
cry-chr2	SerGCT	2623190	2623284	-	
cry-chr2	CysGCA	2637527	2637598	+	
cry-chr2	ValTAC	2667854	2667935	+	
cry-chr2	PheGAA	2728985	2729057	+	cen2
cry-chr2	GluCTC	2729514	2729585	-	cen2
cry-chr2	LeuCAA	2730082	2730186	-	cen2
cry-chr2	LysCTT	2730767	2730852	+	cen2
cry-chr2	ArgACG	2731068	2731140	-	cen2
cry-chr2	ArgACG	2733498	2733570	-	cen2
cry-chr2	IleAAT	2733945	2734018	+	cen2
cry-chr2	AlaAGC	2734350	2734423	-	cen2
cry-chr2	ValAAC	2734868	2734950	+	cen2
cry-chr2	AspGTC	2735291	2735361	-	cen2
cry-chr2	AsnGTT	2748139	2748212	-	cen2
cry-chr2	MetCAT	2748219	2748299	-	cen2
cry-chr2	GluTTC	2749357	2749428	-	cen2
cry-chr2	AspGTC	2757271	2757341	+	cen2
cry-chr2	ValAAC	2757616	2757698	-	cen2
cry-chr2	AlaAGC	2758143	2758216	+	cen2
cry-chr2	IleAAT	2758569	2758642	-	cen2
cry-chr2	ArgACG	2759016	2759088	+	cen2
cry-chr2	LysCTT	2759303	2759388	-	cen2
cry-chr2	GluCTC	2760005	2760076	-	cen2
cry-chr2	GluCTC	2771702	2771773	+	cen2
cry-chr2	LysCTT	2772390	2772475	+	cen2
cry-chr2	ArgACG	2772690	2772762	-	cen2
cry-chr2	IleAAT	2773136	2773209	+	cen2
cry-chr2	AlaAGC	2773562	2773635	-	cen2
cry-chr2	ValAAC	2774080	2774162	+	cen2
cry-chr2	AspGTC	2774437	2774507	-	cen2
cry-chr2	GluTTC	2782344	2782415	+	cen2
cry-chr2	MetCAT	2783473	2783553	+	cen2
cry-chr2	AsnGTT	2783560	2783633	+	cen2
cry-chr2	ArgACG	2788705	2788777	-	cen2
cry-chr2	IleAAT	2789151	2789224	+	cen2
cry-chr2	AlaAGC	2789557	2789630	-	cen2
cry-chr2	ValAAC	2790074	2790156	+	cen2
cry-chr2	AspGTC	2790431	2790501	-	cen2
cry-chr2	AspGTC	2796516	2796586	+	cen2
cry-chr2	ValAAC	2796869	2796951	-	cen2
cry-chr2	AlaAGC	2797336	2797409	+	cen2
cry-chr2	IleAAT	2797745	2797818	-	cen2
cry-chr2	GluTTC	2798548	2798619	+	cen2
cry-chr2	GluTTC	2801937	2802008	-	cen2
cry-chr2	LeuAAG	3030776	3030854	-	
cry-chr2	MetCAT	3146936	3147007	+	
cry-chr2	SerTGA	3212860	3212956	+	
cry-chr2	MetCAT	3212961	3213032	+	
cry-chr2	LysTTT	3250962	3251036	-	

cry-chr2	HisGTG	3374376	3374447	+	
cry-chr2	AsnGTT	3453220	3453293	-	
cry-chr2	AspGTC	3473162	3473232	-	
cry-chr2	TrpCCA	3487278	3487350	+	
cry-chr2	GlnTTG	3509685	3509756	+	
cry-chr2	SerGCT	3512052	3512146	-	
cry-chr2	SerGCT	3543510	3543604	+	
cry-chr2	LysCTT	3551555	3551637	+	
cry-chr2	GlyGCC	3603153	3603223	-	
cry-chr2	ProAGG	3610949	3611020	+	
cry-chr2	GlyTCC	3612247	3612317	+	
cry-chr2	GlyTCC	3616931	3617001	+	
cry-chr2	ThrAGT	3762741	3762812	+	
cry-chr2	LeuCAA	3771210	3771314	+	
cry-chr2	ArgACG	3776236	3776308	+	
cry-chr2	HisGTG	3819860	3819931	-	
cry-chr2	GlyGCC	3914569	3914639	+	
cry-chr2	SerAGA	3915190	3915271	+	
cry-chr2	ValAAC	3946555	3946638	+	
cry-chr3	LeuAAG	40194	40272	+	
cry-chr3	LysCTT	83679	83761	-	
cry-chr3	AlaAGC	103678	103751	+	
cry-chr3	GlyGCC	106935	107005	-	
cry-chr3	ThrAGT	186980	187051	-	
cry-chr3	PheGAA	242481	242553	+	
cry-chr3	SerAGA	389010	389091	+	
cry-chr3	SerAGA	413100	413181	-	
cry-chr3	TyrGTA	417806	417889	-	
cry-chr3	TyrGTA	418634	418717	+	
cry-chr3	ArgTCT	486107	486179	+	
cry-chr3	ThrAGT	601152	601223	-	
cry-chr3	TyrGTA	687190	687273	-	
cry-chr3	GlyGCC	697625	697695	+	
cry-chr3	IleAAT	700223	700296	-	
cry-chr3	ProAGG	700471	700542	-	
cry-chr3	GlyGCC	701100	701170	-	
cry-chr3	ArgACG	797292	797364	-	
cry-chr3	AsnGTT	818057	818130	+	
cry-chr3	ThrAGT	821031	821102	-	
cry-chr3	AlaAGC	900948	901021	-	
cry-chr3	GluTTC	913484	913555	+	cen3
cry-chr3	LeuAAG	914748	914826	+	cen3
cry-chr3	AspGTC	916129	916199	+	cen3
cry-chr3	ValAAC	916550	916632	-	cen3
cry-chr3	AlaAGC	917079	917152	+	cen3
cry-chr3	IleAAT	917485	917558	-	cen3
cry-chr3	ArgACG	917939	918011	+	cen3
cry-chr3	ArgACG	929697	929769	+	cen3
cry-chr3	LysCTT	929983	930068	-	cen3
cry-chr3	LeuCAA	930628	930732	+	cen3
cry-chr3	GluTTC	938349	938420	+	cen3
cry-chr3	ThrAGT	947532	947603	+	cen3
cry-chr3	ThrAGT	952280	952351	-	cen3
cry-chr3	LeuAAG	963327	963405	+	cen3
cry-chr3	AspGTC	964708	964778	+	cen3
cry-chr3	ValAAC	965159	965241	-	cen3
cry-chr3	AlaAGC	965690	965763	+	cen3
cry-chr3	IleAAT	966096	966169	-	cen3
cry-chr3	ArgACG	966550	966622	+	cen3
cry-chr3	ArgACG	968956	969028	+	cen3
cry-chr3	LysCTT	969244	969329	-	cen3
cry-chr3	LeuCAA	969889	969993	+	cen3
cry-chr3	GluCTC	970492	970563	+	cen3
cry-chr3	PheGAA	971020	971092	-	cen3
cry-chr3	AlaAGC	984839	984912	+	cen3
cry-chr3	AspGTC	986744	986814	+	cen3
cry-chr3	ValAAC	987140	987222	-	cen3
cry-chr3	ValAAC	1022461	1022543	+	
cry-chr3	PheGAA	1030689	1030761	-	
cry-chr3	CysGCA	1031226	1031297	-	
cry-chr3	TrpCCA	1174528	1174600	+	
cry-chr3	ProAGG	1210736	1210807	-	
cry-chr3	ProAGG	1264287	1264358	+	
cry-chr3	LeuTAG	1525949	1526027	-	
cry-chr3	TrpCCA	1548446	1548518	+	
cry-chr3	AlaTGC	1600309	1600380	-	
cry-chr3	ArgCCG	1654835	1654915	-	
cry-chr3	LysCTT	2030138	2030221	+	
cry-chr3	ProAGG	2059834	2059905	+	
cry-chr3	ThrAGT	2090541	2090612	+	
cry-chr3	LysCTT	2246722	2246804	-	
cry-chr3	IleAAT	2269370	2269443	+	
cry-chr3	IleTAT	2328647	2328745	-	
cry-chr3	GluCTC	2364126	2364197	+	
cry-chr3	CysGCA	2394500	2394571	-	
cry-chr3	GlyTCC	2546749	2546819	-	
cry-chr3	ProAGG	2547208	2547279	-	
cry-chr3	GlyTCC	2688401	2688471	+	
* Centromeric tDNAs are shaded purple					

Supplementary Table 18: S. octosporus tDNA coordinates

Chromosome	tDNA-anticodon*	start	end	strand	cen located?
oct-chr1	ArgTCT	94382	94454	-	
oct-chr1	LeuAAG	99029	99107	+	
oct-chr1	MetCAT	99400	99471	-	
oct-chr1	GlyTCC	144914	144984	+	
oct-chr1	ArgTCT	145161	145233	-	
oct-chr1	LysCTT	166463	166545	-	
oct-chr1	GlnTTG	166744	166815	+	
oct-chr1	LeuTAA	376312	376409	-	
oct-chr1	GlyGCC	418240	418310	+	
oct-chr1	LysTTT	442444	442518	-	
oct-chr1	LeuTAA	474644	474741	-	
oct-chr1	MetCAT	486339	486419	+	
oct-chr1	TrpCCA	511391	511463	+	
oct-chr1	SerAGA	533513	533594	-	
oct-chr1	GluCTC	604477	604548	+	
oct-chr1	ThrAGT	615008	615079	+	
oct-chr1	AsnGTT	681070	681143	-	
oct-chr1	MetCAT	681150	681230	-	
oct-chr1	GlnCTG	731478	731549	+	
oct-chr1	GlnTTG	731749	731820	+	
oct-chr1	LysCTT	754265	754347	+	
oct-chr1	GlnCTG	754551	754622	+	
oct-chr1	GlnTTG	754844	754915	+	
oct-chr1	GlyGCC	800055	800125	+	
oct-chr1	ThrAGT	820103	820174	-	
oct-chr1	GlyCCC	890566	890636	-	
oct-chr1	LeuAAG	979317	979395	+	
oct-chr1	ValCAC	991172	991243	-	
oct-chr1	PheGAA	1066198	1066270	-	
oct-chr1	GlyGCC	1075047	1075117	+	
oct-chr1	GlyGCC	1075498	1075568	-	
oct-chr1	TyrGTA	1182302	1182385	-	
oct-chr1	LeuCAG	1183007	1183100	+	
oct-chr1	MetCAT	1221556	1221636	+	
oct-chr1	GlyTCC	1248902	1248972	+	
oct-chr1	ProAGG	1253052	1253123	+	
oct-chr1	GluTTC	1294479	1294550	-	
oct-chr1	CysGCA	1294927	1294998	-	
oct-chr1	TyrGTA	1394275	1394358	-	
oct-chr1	ProCGG	1394756	1394854	+	
oct-chr1	AlaTGC	1505703	1505774	+	
oct-chr1	SerAGA	1553392	1553473	-	
oct-chr1	MetCAT	1587467	1587538	-	
oct-chr1	SerCGA	1587550	1587646	-	
oct-chr1	AlaAGC	1685167	1685240	+	
oct-chr1	GluTTC	1712857	1712928	-	
oct-chr1	GluTTC	1963237	1963308	-	
oct-chr1	SerAGA	2029390	2029471	-	
oct-chr1	GlyGCC	2033040	2033110	+	
oct-chr1	AsnGTT	2033262	2033335	+	
oct-chr1	ValTAC	2181548	2181629	+	
oct-chr1	GlnTTG	2235140	2235211	+	
oct-chr1	LeuCAA	2313951	2314056	+	
oct-chr1	HisGTG	2434594	2434665	+	
oct-chr1	PheGAA	2566907	2566979	-	
oct-chr1	ArgCCT	2611872	2611975	-	
oct-chr1	IleAAT	2721160	2721233	-	
oct-chr1	AlaAGC	2792246	2792319	-	
oct-chr1	LeuTAA	2828336	2828435	-	
oct-chr1	SerAGA	2832105	2832186	+	
oct-chr1	GlyGCC	3028407	3028477	+	
oct-chr1	GlyGCC	3029025	3029095	-	
oct-chr1	ProAGG	3031819	3031890	+	
oct-chr1	ProAGG	3083017	3083088	+	
oct-chr1	TrpCCA	3119280	3119352	-	
oct-chr1	CysGCA	3262694	3262765	+	
oct-chr1	PheGAA	3263180	3263252	+	
oct-chr1	ValAAC	3271549	3271631	-	
oct-chr1	ArgACG	3295076	3295148	+	
oct-chr1	ValAAC	3307009	3307091	+	cen1
oct-chr1	AspGTC	3307399	3307469	-	cen1
oct-chr1	AlaAGC	3309597	3309670	-	cen1
oct-chr1	ValAAC	3316124	3316206	+	cen1
oct-chr1	AspGTC	3316530	3316600	-	cen1
oct-chr1	AlaAGC	3320302	3320375	+	cen1
oct-chr1	IleAAT	3320686	3320759	-	cen1
oct-chr1	ArgACG	3321277	3321349	+	cen1
oct-chr1	ArgACG	3323522	3323594	+	cen1
oct-chr1	LysCTT	3323783	3323866	-	cen1
oct-chr1	LeuCAA	3324411	3324515	+	cen1
oct-chr1	GluTTC	3329434	3329505	+	cen1
oct-chr1	ThrAGT	3336035	3336106	+	cen1
oct-chr1	ThrAGT	3339535	3339606	-	cen1
oct-chr1	LeuAAG	3352729	3352807	+	cen1
oct-chr1	AspGTC	3355194	3355264	+	cen1
oct-chr1	ValAAC	3355572	3355654	-	cen1
oct-chr1	AlaAGC	3356111	3356184	+	cen1
oct-chr1	IleAAT	3356502	3356575	-	cen1
oct-chr1	ArgACG	3357093	3357165	+	cen1
oct-chr1	ArgACG	3359332	3359404	+	cen1
oct-chr1	LysCTT	3359593	3359676	-	cen1
oct-chr1	LeuCAA	3360225	3360329	+	cen1
oct-chr1	GluCTC	3361058	3361129	+	cen1
oct-chr1	PheGAA	3361584	3361656	-	cen1
oct-chr1	AlaAGC	3367799	3367872	+	cen1
oct-chr1	AspGTC	3369988	3370058	+	cen1
oct-chr1	ValAAC	3370367	3370449	-	cen1
oct-chr1	AlaAGC	3370906	3370979	+	cen1
oct-chr1	IleAAT	3371297	3371370	-	cen1
oct-chr1	GluTTC	3371989	3372060	+	cen1
oct-chr1	GluTTC	3373277	3373348	-	cen1
oct-chr1	LeuAAG	3600727	3600805	-	
oct-chr1	MetCAT	3716580	3716651	+	

oct-chr1	SerTGA	3780535	3780627	+	
oct-chr1	MetCAT	3780632	3780703	+	
oct-chr1	LysTTT	3818008	3818082	-	
oct-chr1	HisGTG	3934735	3934806	+	
oct-chr1	AsnGTT	4012809	4012882	-	
oct-chr1	AspGTC	4032087	4032157	-	
oct-chr1	TrpCCA	4045019	4045091	+	
oct-chr1	GlnTTG	4068273	4068344	+	
oct-chr1	SerGCT	4071356	4071450	-	
oct-chr1	SerGCT	4102633	4102727	+	
oct-chr1	LysCTT	4110405	4110487	+	
oct-chr1	GlyGCC	4159843	4159913	-	
oct-chr1	ProAGG	4173548	4173619	+	
oct-chr1	GlyTCC	4174013	4174083	+	
oct-chr1	GlyTCC	4176699	4176769	+	
oct-chr1	ThrAGT	4311155	4311226	+	
oct-chr1	LeuCAA	4319547	4319652	+	
oct-chr1	ArgACG	4324450	4324522	+	
oct-chr1	HisGTG	4371998	4372069	-	
oct-chr1	GlyGCC	4461913	4461983	+	
oct-chr1	SerAGA	4462213	4462294	+	
oct-chr1	ValAAC	4491370	4491452	+	
oct-chr2	GlyTCC	39382	39452	-	
oct-chr2	ProAGG	39617	39688	-	
oct-chr2	AsnGTT	85822	85895	-	
oct-chr2	ProAGG	86056	86127	+	
oct-chr2	AspGTC	201799	201869	+	
oct-chr2	AsnGTT	405305	405378	+	
oct-chr2	CysGCA	530221	530292	-	
oct-chr2	AlaCGC	542870	542952	+	
oct-chr2	ProTGG	591265	591336	+	
oct-chr2	CysGCA	833576	833647	+	
oct-chr2	GluCTC	864839	864910	-	
oct-chr2	IleTAT	899912	900010	+	
oct-chr2	IleAAT	958450	958523	-	
oct-chr2	LysCTT	980983	981065	+	
oct-chr2	ThrAGT	1139449	1139520	-	
oct-chr2	LysCTT	1198159	1198241	-	
oct-chr2	ArgCCG	1568733	1568813	+	
oct-chr2	AlaTGC	1625369	1625440	+	
oct-chr2	TrpCCA	1676828	1676900	-	
oct-chr2	LeuTAG	1700762	1700840	+	
oct-chr2	SerAGA	1996830	1996911	-	
oct-chr2	ProAGG	1999951	2000022	-	
oct-chr2	ProAGG	2000433	2000504	+	
oct-chr2	ArgACG	2287558	2287630	-	
oct-chr2	ProTGG	2390650	2390721	-	
oct-chr2	TrpCCA	2459797	2459869	-	
oct-chr2	ThrTGT	2493783	2493854	-	
oct-chr2	AlaAGC	2573939	2574012	-	cen2
oct-chr2	ValAAC	2574486	2574568	+	cen2
oct-chr2	AspGTC	2574877	2574947	-	cen2
oct-chr2	AlaAGC	2577056	2577129	-	cen2
oct-chr2	PheGAA	2603477	2603549	+	cen2
oct-chr2	GluCTC	2603977	2604048	-	cen2
oct-chr2	LeuCAA	2604778	2604882	-	cen2
oct-chr2	LysCTT	2605432	2605515	+	cen2
oct-chr2	LysCTT	2616460	2616543	-	cen2
oct-chr2	LeuCAA	2617093	2617197	+	cen2
oct-chr2	GluTTC	2621802	2621873	+	cen2
oct-chr2	AspGTC	2626442	2626512	+	cen2
oct-chr2	ValAAC	2626851	2626933	-	cen2
oct-chr2	ArgTCG	2638281	2638353	+	cen2
oct-chr2	ProAGG	2640583	2640654	-	cen2
oct-chr2	ArgACG	2641923	2641995	-	cen2
oct-chr2	IleAAT	2642544	2642617	+	cen2
oct-chr2	AlaAGC	2642931	2643004	-	cen2
oct-chr2	ValAAC	2643478	2643560	+	cen2
oct-chr2	AspGTC	2643860	2643930	-	cen2
oct-chr2	GluTTC	2646905	2646976	+	cen2
oct-chr2	LeuAAG	2647420	2647498	+	cen2
oct-chr2	AspGTC	2649889	2649959	+	cen2
oct-chr2	ValAAC	2650259	2650341	-	cen2
oct-chr2	AlaAGC	2650815	2650888	+	cen2
oct-chr2	IleAAT	2651208	2651281	-	cen2
oct-chr2	ArgACG	2651815	2651887	+	cen2
oct-chr2	HisGTG	2817295	2817366	+	
oct-chr2	ThrTGT	2915884	2915955	-	
oct-chr2	ArgTCG	2945577	2945649	+	
oct-chr2	LysTTT	3002902	3002976	+	
oct-chr2	GlnTTG	3006348	3006419	-	
oct-chr2	ArgACG	3023201	3023273	+	
oct-chr2	GlyGCC	3116988	3117058	+	
oct-chr2	IleAAT	3119601	3119674	-	
oct-chr2	ProAGG	3119852	3119923	-	
oct-chr2	GlyGCC	3120229	3120299	-	
oct-chr2	TyrGTA	3131404	3131487	+	
oct-chr2	ThrAGT	3212357	3212428	+	
oct-chr2	ArgTCT	3325687	3325759	-	
oct-chr2	TyrGTA	3390505	3390588	-	
oct-chr2	TyrGTA	3391194	3391277	+	
oct-chr2	SerAGA	3395941	3396022	+	
oct-chr2	SerAGA	3418665	3418746	-	
oct-chr2	PheGAA	3561064	3561136	-	
oct-chr2	ThrAGT	3628451	3628522	+	
oct-chr2	GlyGCC	3703319	3703389	+	
oct-chr2	AlaAGC	3704215	3704288	-	
oct-chr2	LysCTT	3724005	3724087	+	
oct-chr2	LeuAAG	3764112	3764190	-	
oct-chr3	GlyTCC	38143	38213	-	
oct-chr3	ProAGG	38378	38449	-	
oct-chr3	ProAGG	60939	61010	-	

oct-chr3	GlyGCC	423718	423788	+	
oct-chr3	ThrAGT	441806	441877	+	
oct-chr3	ValCAC	451661	451732	-	
oct-chr3	GlnTTG	451925	451996	-	
oct-chr3	SerAGA	606948	607029	-	
oct-chr3	HisGTG	640295	640366	+	
oct-chr3	AlaTGC	718000	718071	+	
oct-chr3	LysCTT	723496	723578	+	
oct-chr3	ThrCGT	723953	724024	+	
oct-chr3	GluCTC	802545	802616	+	
oct-chr3	LysCTT	905626	905708	-	
oct-chr3	LysTTT	919805	919879	-	
oct-chr3	TyrGTA	931084	931167	+	
oct-chr3	IleAAT	956454	956527	+	
oct-chr3	GlnTTG	1004387	1004458	-	
oct-chr3	AsnGTT	1045547	1045620	-	
oct-chr3	SerGCT	1159113	1159207	-	
oct-chr3	LeuAAG	1160824	1160902	+	
oct-chr3	ValAAC	1189558	1189639	-	
oct-chr3	MetCAT	1233336	1233407	-	
oct-chr3	SerCGA	1233419	1233515	-	
oct-chr3	TyrGTA	1256708	1256791	-	
oct-chr3	SerAGA	1626103	1626184	+	
oct-chr3	LysCTT	1631110	1631192	+	
oct-chr3	SerGCT	1682092	1682186	-	
oct-chr3	CysGCA	1696332	1696403	+	
oct-chr3	ValTAC	1726789	1726870	+	
oct-chr3	PheGAA	1787162	1787234	+	cen3
oct-chr3	GluCTC	1787669	1787740	-	cen3
oct-chr3	LeuCAA	1788461	1788565	-	cen3
oct-chr3	LysCTT	1789138	1789221	+	cen3
oct-chr3	ArgACG	1789410	1789482	-	cen3
oct-chr3	AspGTC	1791072	1791142	+	cen3
oct-chr3	ValAAC	1791451	1791533	-	cen3
oct-chr3	AlaAGC	1792005	1792078	+	cen3
oct-chr3	IleAAT	1792389	1792462	-	cen3
oct-chr3	ArgACG	1792979	1793051	+	cen3
oct-chr3	AsnGTT	1793566	1793639	-	cen3
oct-chr3	MetCAT	1793646	1793726	-	cen3
oct-chr3	GluTTC	1794397	1794468	-	cen3
oct-chr3	AspGTC	1806816	1806886	+	cen3
oct-chr3	ValAAC	1807201	1807283	-	cen3
oct-chr3	AlaAGC	1813684	1813757	+	cen3
oct-chr3	IleAAT	1814086	1814159	-	cen3
oct-chr3	ArgACG	1814690	1814762	+	cen3
oct-chr3	LysCTT	1814946	1815030	-	cen3
oct-chr3	GluCTC	1815189	1815260	-	cen3
oct-chr3	GluCTC	1827491	1827562	+	cen3
oct-chr3	LysCTT	1827721	1827805	+	cen3
oct-chr3	ArgACG	1827989	1828061	-	cen3
oct-chr3	IleAAT	1828592	1828665	+	cen3
oct-chr3	AlaAGC	1828994	1829067	-	cen3
oct-chr3	ValAAC	1837990	1838072	+	cen3
oct-chr3	AspGTC	1838387	1838457	-	cen3
oct-chr3	GluTTC	1850803	1850874	+	cen3
oct-chr3	MetCAT	1851545	1851625	+	cen3
oct-chr3	AsnGTT	1851632	1851705	+	cen3
oct-chr3	ArgACG	1852220	1852292	-	cen3
oct-chr3	IleAAT	1852842	1852915	+	cen3
oct-chr3	AlaAGC	1853242	1853315	-	cen3
oct-chr3	ValAAC	1853788	1853870	+	cen3
oct-chr3	AspGTC	1854215	1854285	-	cen3
oct-chr3	LeuAAG	1856683	1856761	-	cen3
oct-chr3	GluTTC	1857205	1857276	-	cen3
oct-chr3	AlaAGC	1869522	1869595	+	
oct-chr3	ThrAGT	1948053	1948124	+	
oct-chr3	AsnGTT	1951055	1951128	-	
oct-chr3	HisGTG	1970055	1970126	-	
oct-chr3	HisGTG	1970661	1970732	+	
oct-chr3	ProAGG	1986958	1987029	-	
oct-chr3	LysCTT	1987352	1987435	-	
oct-chr3	AsnGTT	2003138	2003211	-	
oct-chr3	MetCAT	2003218	2003298	-	
oct-chr3	SerAGA	2031616	2031697	+	
oct-chr3	ProAGG	2041722	2041793	+	
oct-chr3	LysCTT	2041863	2041945	+	
oct-chr3	SerTGA	2073241	2073333	+	
oct-chr3	MetCAT	2073338	2073409	+	
oct-chr3	IleTAT	2258764	2258862	-	
oct-chr3	LysTTT	2299093	2299167	-	
oct-chr3	GlyGCC	2304024	2304094	+	
oct-chr3	ThrAGT	2311134	2311205	-	
oct-chr3	SerGCT	2451039	2451133	+	
oct-chr3	GlyGCC	2476716	2476786	+	
oct-chr3	GlyGCC	2477292	2477362	-	
oct-chr3	GlyGCC	2478606	2478676	+	
oct-chr3	ArgACG	2669097	2669169	-	
oct-chr3	PheGAA	2846492	2846564	-	
oct-chr3	PheGAA	2847747	2847819	-	
oct-chr3	GlyGCC	2897225	2897295	+	
oct-chr3	ThrAGT	2928563	2928634	-	
oct-chr3	GlyTCC	2956870	2956940	-	
oct-chr3	ProTGG	2957085	2957156	-	
* Centromeric tDNAs are shaded purple					

Supplementary Table 19: Features associated with S. cryophilus and S. octosporus centromere DNA elements

S. cryophilus	*S. octosporus*	Chromatin, features, comments
cCNT-L	oCNT-L	CENP-A chromatin
cCNT-S	oCNT-S	
c-cnt1	o-cnt2	Contain CNT-L and CNT-S elements
c-cnt2	o-cnt3	Contain CNT-L and CNT-S elements
c-cnt3	o-cnt1	Short central core, with long imrs
c-imr3	o-imr1	c-imr3 has LTRs which may act as boundaries
cFSAR-1	oFSAR-1	Heterochromatin. 5S-associated repeats
cFSAR-2	oFSAR-2	
cFSAR-3	oFSAR-3	
cTAR-11	oTAR-11	tDNA-associated repeats (TAR) - elements that are always associated with particular single tDNAs in both species and occur in equivalent positions. cTAR-14 and oTAR-14-ex contain retrotransposon remnants.
cTAR-12	oTAR-12	
cTAR-13	oTAR-13	
cTAR-14	oTAR-14	
cHR-15	oHR-15	Heterochromatin. Not associated with tRNAs but occur in equivalent positions in the two species.
cHR-19	oHR-19	
cTAR-4	oTAR-4	tDNA-associated repeats (TAR) elements that are always associated with multiple specific tDNAs. No/little CENP-A or heterochromatin (except TAR-7s which have heterochromatin on ~2 kb non-tDNA part of repeat). TARs may act as boundaries.
cTAR-5	oTAR-5	
cTAR-6	oTAR-6	
cTAR-7	oTAR-7	
cTAR-8	oTAR-8	
cTAR-9	oTAR-9	
cTAR-10	oTAR-10	
cTAR-11	oTAR-11	

Supplementary Table 20: List of Schiosaccharomyces strains used.

Species	Strain ID	Genotype	Used for	Figure	Source Laboratory
S. pombe	A7408	*h- cc2D6kb:cc1 ars1:nmt41-GFP-cnp1-NAT ade6-704-HYGMX6 his3-D1 leu1-32 ura4-DSE/D18 arg3?*	Centromere establishment	Fig 6	Allshire (ref 15)
S. pombe	A7373	*h- ade6-704-HYGMX6 his3-D1 leu1-32 ura4-DSE/D18? arg3? cc2D6kb:cc1*	Centromere establishment	not shown	Allshire (ref 15)
S. pombe	6960	*h- lys1+::cnp1-1 cnp1::ura4+ leu1-32 ura4-*	Complementation	Fig 5	Takahashi (ref 30)
S. pombe	1645	*h+ ade6-210 arg3-D4 his3-D1 leu1-32 ura4-D18*	Localisation	Fig 5	Allshire
S. pombe	968	h^{90}	H3K9me2 ChIP-seq for mating-type region plot (h90)	Supp Fig S1	Fantes / Sawin
S. pombe	972	*h-*	H3K9me2 and CENP-A ChIP-seq	Figs 2, Supp Fig S5	Allshire
S. pombe	AMC501	*h+ ade6-704 ura4-D18 leu1-32 rhn201 cdc6-L591G*	Minion nanopore sequencing, assembly of S. pombe genome	Supp Fig S5	Nieduszynski / Murray
S. pombe	A6372	*h- leu1 ura4 Δcen1::pADH1-loxP-KanR cd39 (tel1L neocentromere)*	CENP-A ChIP-seq for neocentromere regions	Fig 4	Ishii / Takahashi (ref 29)
S. pombe	A6374	*h- leu1 ura4 Δcen1::pADH1-loxP-KanR cd60 (tel1R neocentromere)*	CENP-A ChIP-seq for neocentromere regions	Fig 4	Ishii / Takahashi (ref 29)
S. octosporus	A6969	h^{90}	PacBio sequencing, H3K9me2 and CENP-A ChIP-seq	Figs 1, 2, Supp Fig S2	Rhind (ref 16)
S. cryophilus	A6972	h^{90}	PacBio sequencing, H3K9me2 and CENP-A ChIP-seq	Fig 1, 2, Supp Fig S3	Rhind (ref 16)
S. japonicus	A1856	h^{90}	PacBio sequencing, H3K9me2 and CENP-A ChIP-seq	Supp Fig S5	Rhind (ref 16)

Supplementary Table 21: Summary of sequencing platforms used.

Type of Sequencing	Species/ChIP	Sequencing Facility	Instrument / Chemistry /info etc
PacBio 1	S. pombe S. octosporus S. cryophilus S. japonicus	Biomedical Research Core Facilities, University of Michigan	2 SMRT cells each (no bluepippin) PacBio RSII instrument, P4-XL
PacBio 2 with BluePippin	S. pombe S. octosporus S. cryophilus S. japonicus	CSHL Cancer Center Next Generation Genomics Shared Resource	no BluePippin: 1 SMRT cells each With BluePippin technology: 8 SMRT cells each for Oct and Cry; 15 SMRT cells for Jap PacBio RSII instrument with P4/C3 chemistry
Minion nanopore	S. pombe AMC501	Oxford	MinION nanopore sequencer
ChIP-seq	S. japonicus CENP-A H3K9me2	BGI	Illumina GAII Single end
ChIP-seq	S. octosporus CENP-A H3K9me2	Ark Genomics	HiSeq2000 Paired end
ChIP-seq	S. cryophilus CENP-A H3K9me2	Ark Genomics	HiSeq2000 Paired end
ChIP-seq	S. pombe CENP-A H3K9me2	Ark Genomics	HiSeq2000 Paired end
ChIP-seq	S. pombe 972h- H3K9me2	Edinburgh Genomics	HiSeq2000 Paired end
ChIP-seq	S. pombe h90 H3K9me2	Allshire Lab	Miniseq Paired end

Supplementary Table 22: List of oligonucleotide primers used.

name	sequence	restriction site	anneals	used in
WA638	TACTACacgcgtAATACCAACATAggccatattggccattagtaccagtactagtgtc	MluI, SfiI	S. pombe	plasmid construction
WA644	TACTACctcgagCATGCTTTTAGTGCGGTCATT	XhoI	S. pombe	plasmid construction
WA841	tactacCATATGGCAAAGAAATCTTTAATGGCTGAGCC	NdeI	S. pombe	plasmid construction
WA842	tactacGGATCCTCAAGCACCACGAATCCTCC	BamHI	S. pombe	plasmid construction
WA843	tactacCATATGGCTAAAAAATCGTTGATGGC	NdeI	S. octosporus	plasmid construction
WA844	tactacGGATCCTCAAGCACCACGGATACGACG	BamHI	S. octosporus	plasmid construction
WA845	tactacCATATGGCTAAAAAATCTTTAATGGCAGAACCAGG	NdeI	S. cryophilus	plasmid construction
WA846	tactacGGATCCTCAAGCACCACGAATACGACG	BamHI	S. cryophilus	plasmid construction
WA847	tactacCATATGGCTAAACGCTCTTTTGTTGCGG	NdeI	S. japonicus	plasmid construction
WA848	tactacAGATCTTTAGGATCCTCGAATACGTCG	BglII	S. japonicus	plasmid construction
WB3	CAGACAATCGCATGGTACTATC		S.pom-cnt1, S.pom-cnt3	ChIP-qPCR
WB4	AGGTGAAGCGTAAGTGAGTG		S.pom-cnt1, S.pom-cnt3	ChIP-qPCR
WB11	CATTAAACAAACAACGGCACAC		S.pom-cnt2	ChIP-qPCR
WB12	TAAGCCAGCAAATTCCTTGAG		S.pom-cnt2	ChIP-qPCR
WB388	tactacAGATCTTCCGAATGGACTCATGAAGGG	BglII	cnt1	minichromosome construction
WB389	tactacCCATGGTAAGGCTTACATGAAAGAAATTTTAGTGCTG	NcoI	cnt1	minichromosome construction
WB393	tactacCCATGGCGTAAATATATAGCAGGTTTAACGC	NcoI	cnt2	minichromosome construction
WB395	tactacGGATCCCAGCATGAATTCATAAAGACC	BamHI	cnt2, cnt3	minichromosome construction
WB397	TTCAACAGATCTAGTGAATCCCG	BglII (in sequence)	cnt2, cnt3	minichromosome construction
WB398	tactacGGATCCGTTGAAAATAAAGAGCTGTAACC	BamHI	cnt2, cnt3	minichromosome construction
WB399	tactacGTCGACGAGATACAGAAAAAAGTAAGCC	SalI	cnt2, cnt3	minichromosome construction
WB402	tactacCCATGGCAAGCGGTTAAATAAGTATCAG	NcoI	cnt2	minichromosome construction
WB403	tactacCCATGGGGAGGTATGACCGTATAATTG	NcoI	cnt2	minichromosome construction
WB407	tactacGTCGACCGATCTACTAAGATTTACGATG	SalI	cnt3	minichromosome construction
WB409	tactacGGATCCAATAACAATTTTCGTCAGTTACAGC	BamHI	cnt3	minichromosome construction
WB597	TGGATTGCCTGCTGTTTTGC		S.oct-cnt3	chIP-qPCR
WB598	AAATGCTGGGTTTCGAGGAC		S.oct-cnt3	chIP-qPCR
WB601	TACATCTCCCTTTCGCTGATGC		S.oct-cnt2, S.oct-cnt3	chIP-qPCR
WB602	TAAACGCCGCTATGGTTCTG		S.oct-cnt2, S.oct-cnt3	chIP-qPCR
WB609	TTTGAGTTAGCTGCGGTGAG		S.oct-cnt1	chIP-qPCR
WB610	GACGCGAAAACTGTTTACGG		S.oct-cnt1	chIP-qPCR

Supplementary Table 23: List of plasmids constructed and used.

plasmid / minichromosome	vector	source of insert/plasmid	primer F	primer R
pK(5.6kb)-MCS-ΔBam		S.pombe K repeat		
pKp (pMC91)	pMC1	S.pombe K repeat	WA638	WA644
pK-So-cnt1-3.2kb	pK(5.6kb)-MCS-DBam	S.oct-cen1	WB388	WB389
pK-So-cnt2-6.5kb	pK(5.6kb)-MCS-DBam	S.oct-cen2	WB402	WB397
pK-So-cnt2-4.7kb	pK(5.6kb)-MCS-DBam	S.oct-cen2	WB393	WB395
pK-So-cnt2-10kb	pK-So-cnt2-6.5kb	S.oct-cen2	WB398	WB399
pK-So-cnt3-6.5kb	pK(5.6kb)-MCS-DBam	S.oct-cen3	WB407	WB409
pK-So-cnt3-3.6kb	pK(5.6kb)-MCS-DBam	S.oct-cen3	WB397	WB403
pK-So-cnt3-2.6kb	pK(5.6kb)-MCS-DBam	S.oct-cen3	WB395	WB403
pKp-So-cnt3-6.5kb	pKp	S.oct-cen3	WB407	WB409
pKp-So-cnt3-3.6kb	pKp	S.oct-cen3	WB397	WB403
pREP41-GFP-Sp-Cnp1	pREP41-GFP (Craven et al)	S. pombe gDNA	WA841	WA842
pREP41-GFP-Sc-Cnp1	pREP41-GFP	S. cryophilus genomic DNA	WA843	WA844
pREP41-GFP-So-Cnp1	pREP41-GFP	S. octosporus genomic DNA	WA845	WA846
pREP41-GFP-Sj-Cnp1	pREP41-GFP	S. japonicus genomic DNA	WA847	WA848