716–728 Nucleic Acids Research, 2019, Vol. 47, No. 2
doi: 10.1093/nar/gky1150
Published online 16 November 2018

DNA sequence elements required for partitioning competence of the *Saccharomyces cerevisiae* 2-micron plasmid *STB* locus

Mary E. McQuaid, Elizabeth J. Polvi and Melanie J. Dobson*

Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, NS B3H 4R2, Canada

Received August 03, 2018; Revised October 26, 2018; Editorial Decision October 29, 2018; Accepted October 29, 2018

ABSTRACT

Equal partitioning of the multi-copy yeast 2-micron plasmid requires association of plasmid proteins Rep1 and Rep2 with tandem repeats at the plasmid *STB* locus. To identify sequence elements required for these associations we generated synthetic versions of a 63-bp section of *STB*, encompassing one repeat. A single copy of this sequence was sufficient for Rep protein association *in vivo*, while two directly arrayed copies provided partitioning function to a plasmid lacking all other 2-micron sequences. Partitioning efficiency increased with increasing repeat number, reaching that conferred by the native *STB* repeat array. By altering sequences in synthetic repeats, we identified the TGCA component of a TGCATTTTT motif as critical for Rep protein recognition, with a second TGCA sequence in each repeat also contributing to association. Mutation of TGCATTTTT to TGTATTTT, as found in variant 2-micron *STB* repeats, also allowed Rep protein association, while mutation to TGCATTAAT impaired inheritance without abolishing Rep protein recognition, suggesting an alternate role for the T-tract. Our identification of sequence motifs required for Rep protein recognition provides the basis for understanding higher-order Rep protein arrangements at *STB* that enable the yeast 2-micron plasmid to be efficiently partitioned during host cell division.

INTRODUCTION

The 2μm plasmid is a double-stranded, circular DNA, present at high copy number in the nucleus of most strains of the budding yeast *Saccharomyces cerevisiae*. The plasmid confers no phenotype or selective advantage to the host cell (1,2). Plasmid persistence depends instead on plasmid-encoded mechanisms that ensure equal partitioning at cell division (3,4), and a means of amplifying plasmid copy number if it falls below normal levels (5) (for a review see (6)).

Equal partitioning of the 2μm plasmid during host cell division requires association of the plasmid-encoded Rep1 and Rep2 proteins with the *cis*-acting plasmid partitioning locus, *STB* (3,4,7). Absence of any one of these three components results in a strong maternal bias in inheritance, with the majority of plasmid copies being retained in the mother cell during cell division (8). The nature of these associations is not well understood. *In vivo*, Rep1 association with *STB* does not require the presence of other 2μm proteins, while Rep2 association is greatly reduced in the absence of Rep1 (9). *In vitro*, purified Rep2 displays a preference for binding *STB* DNA in a southwestern assay but neither Rep protein displays DNA-binding activity in gel shift assays (10,11). Rep1 and Rep2 have been found to associate with *STB* DNA using a plasmon resonance assay, but only in the presence of urea-solubilized yeast extracts, suggesting that host proteins may be required to mediate this interaction (10).

In addition to Rep1 and Rep2, 2μm plasmid partitioning is dependent on recruitment of several host factors to *STB*, including the nuclear motor protein Kip1 (12) the RSC2 chromatin remodeling complex (13–15), the centromere-specific histone H3 variant Cse4 (16), and cohesin (17). Although some of the host factors associated with *STB* are also present at chromosomal centromeres, a kinetochore complex is not formed at *STB* (18), making it unclear how the assembly of plasmid and host proteins at *STB* mediates partitioning. The protein complex at *STB* may ensure equal partitioning by tethering the plasmids to chromosomes (17,19,20), a strategy similar to that used by many viral episomes (21). There is also evidence to suggest that the plasmid may interact with the spindle or spindle-associated proteins independently of chromosomes (18,22). Although aspects of the partitioning mechanism remain to be established, association of the Rep proteins and the RSC2 complex with *STB* is a prerequisite for recruitment of cohesin, which provides sister plasmid pairing during mitosis and ensures partitioning of replicated plasmids is coordinated with chromosome segregation upon cohesin disassembly (17,23).

*To whom correspondence should be addressed. Tel: +1 902 494 7182; Fax: +1 902 494 1355; Email: dobson@dal.ca

The *STB* locus can be subdivided at a HpaI restriction site into *STB-proximal* (AvaI to HpaI; *STB-P*) and *STB-distal* (HpaI to PstI; *STB-D*), so named for their positioning relative to the single origin of replication (*ORI*) on the 2μm plasmid (Figure 1A) (3,24). *STB-D* contains a transcriptional silencer and transcription termination signals that protect *STB-P* from disruptive transcription from the adjacent *RAF* gene (24). The 296-bp *STB-P* sequence is necessary and sufficient for Rep protein-mediated partitioning in most sequence contexts, and contains a tandem array of five 62- to 63-bp imperfect direct repeats, with another half repeat on the *ORI* side of the AvaI site (24,25). Sequence identity shared by the repeats ranges from 65% to 98%. Each repeat contains a 25-bp core sequence showing greater conservation flanked by more divergent sequences (Figure 1A) (26). Nuclease sensitivity assays indicate the presence of two nucleosomes in the *STB-P* region although the repetitive nature of the underlying sequence has prevented precise delineation of nucleosome boundaries (13,27,28).

The sequence requirements for Rep protein association with *STB-P* are not known. The sequence TGCATTTTT has been suggested as a potential Rep protein-binding motif based on its presence in the promoter regions of the 2μm plasmid genes and also occurrence five times within *STB-P* (Figure 1A) (29,30). Co-expression of the Rep proteins represses transcription of all four plasmid genes, with repression of *FLP*, and therefore of Flp recombinase-dependent plasmid amplification (5), controlling 2μm plasmid copy number (30–32). On this basis, Rep protein association with the plasmid promoters was expected and has recently been demonstrated (33).

In this study we have created synthetic versions of the *STB* repeat to identify sequence elements required for Rep protein recognition and partitioning function. We found that Rep1 is able to associate with a single copy of the 63-bp synthetic *STB* repeat *in vivo*, while two directly arrayed copies of this sequence were sufficient to confer partitioning function in the absence of any other 2μm sequences. We also showed that the previously identified TGCATTTTT sequence is required for partitioning function, with the TGCA element, being critical for Rep protein association with the synthetic *STB* repeat. These results represent the first experimental exploration of *STB* sequence elements.

MATERIALS AND METHODS

Strains and media

Yeast strains used in this study (Supplementary Table S1) were all derived in a W303 genetic background (*ade2-1 his3-11,15 trp1-1 leu2-3,112 ura3-1*) (34). Strains lacking the 2μm plasmid, designated *cir*^{0}, were derived from strains containing the 2μm plasmid, *cir*$^+$, by expression of a defective Flp recombinase from the plasmid pBIS-GALkFLP-(TRP1) (35).

Yeast were cultured in YPAD (1% yeast extract, 2% Bacto Peptone, 0.003% adenine, 2% glucose), or synthetic defined (SD; 0.67% Difco yeast nitrogen base without amino acids, 2% glucose, 0.003% adenine, 0.002% uracil and all required amino acids) medium at 28°C. For induction of galactose-inducible promoters, 2% glucose was replaced with 2% galactose. For selection of plasmids or gene replacements tagged with nutritional genes, SD medium lacking the appropriate nucleotide base or amino acid was used. For selection of *kanMX4*-tagged plasmids, YPAD supplemented with 200 mg/L geneticin (G418, Sigma) was used. Yeast were transformed using the Li/SSS-DNA/PEG method (36).

Escherichia coli strain DH5α was used for propagation of plasmids. *Escherichia coli* were cultured and manipulated according to standard protocols (37).

Synthetic *STB* repeats

A 63-bp portion of *STB-P* (nucleotides 2986–3049 in the A form of Scp1, the variant of the 2μm plasmid found in most laboratory strains of *S. cerevisiae*, (NCBI GenBank J01347.1) was used as the basis of the synthetic *STB* repeat duplexes used in this study (Figure 1B). A duplex based on a second 63-bp stretch, nucleotides 3006–3069 (WT_{+20}) had similar partitioning efficiency (Supplementary Figure S2) and was not used for subsequent mutational studies. To facilitate creation of *STB* loci with varying numbers of synthetic repeats or altered sequence composition, overhangs matching BamHI and BglII restriction sites were included on the ends of each synthetic repeat.

Pairs of complementary oligonucleotides (Supplementary Table S2) were annealed to form synthetic *STB* single repeat duplexes 68 bp in length (63 bp matching *STB-P* with 5 bp of overhang to generate restriction sites) (Table 1). To create tandem arrays of synthetic *STB* repeats, duplexes were phosphorylated and then incubated with T4 ligase. Head-to-head and tail-to-tail ligation products were eliminated by digestion with BamHI and BglII. Sequences of all synthetic repeats were confirmed by sequencing after introduction into plasmids used to assess their function.

Plasmids

A complete list of plasmids used in this study is found in Supplementary Table S3.

2μm-based plasmids. The 2μm-based plasmid pKan (9) was derived from the partitioning-competent, amplification-defective, *flp*$^-$ *ADE2*-tagged B-form 2μm plasmid, pAS4 (11) by replacing the *ADE2* marker gene (inserted in the *FLP* gene) with the *kanMX4* gene cassette. The *kanMX4*-marked plasmid provides a sensitive assay for plasmid missegregation events; unlike plasmids containing a nutritional marker gene such as *ADE2*, *kanMX4*-marked plasmids do not confer an additional growth advantage when present at higher copy number (9). Selection for cells with more copies of a nutritional marker drives plasmid copy number in the population upward increasing the rate at which some daughter cells will receive plasmid by random diffusion rather than active partitioning. Bacterial vector sequences inserted in the inverted repeat downstream of the *REP1* and *REP2* genes in pKan enable propagation of the plasmid in *E. coli.* A derivative of pKan that lacks the *REP1* gene (pKanΔREP1) has previously been described (9).

To create a version of pKan lacking *STB-P* (pKanΔSTB) in which synthetic *STB* repeats could be introduced at

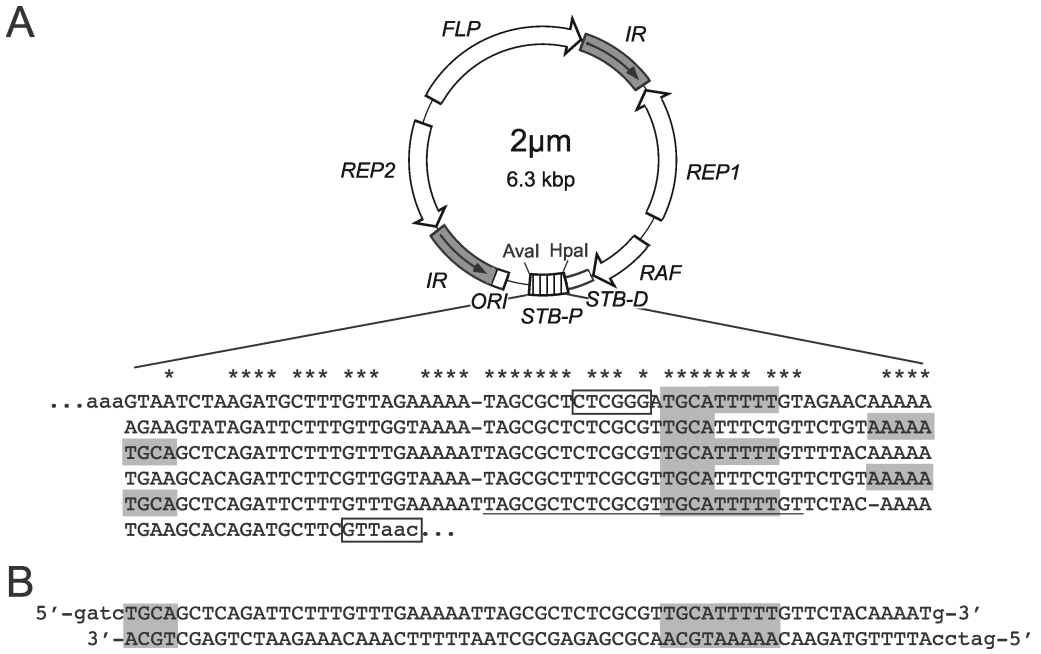

Figure 1. Map of the 2μm plasmid and *STB* repeat sequences. (**A**) Organization of the A form of the 2μm plasmid is shown at top, with positions of plasmid genes (white arrows), inverted repeat (*IR*; gray boxes with arrows showing orientation), origin of replication (*ORI*; white box), *STB-proximal* (*STB-P*); striped box), and *STB-distal* (*STB-D*; white box) indicated. The *STB-P* sequence is aligned beneath to show the degree of similarity between the five-and-a-half tandem *STB* repeats, with sequences flanking the repeats shown in lowercase. Asterisks indicate positions of identical residues in the alignment. AvaI and HpaI sites at the boundaries of *STB-P* are boxed. The 25-bp repeat core consensus sequence is underlined (26). TGCATTTTT elements are shaded in grey. (**B**) The sequence of the wild-type synthetic repeat used in this study is shown. Residues matching the *STB* repeat closest to *STB-D*, on which the synthetic repeat is based, are shown in upper case, while residues added to create BglII and BamHI overhangs are shown in lowercase.

Table 1. Sequence of synthetic *STB* duplexes

STB duplex	Sequence of duplex (5′→3′)[a]
	TGCA TGCATTTTT
WT	gatcTGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTCTACAAAATg
TGTA	gatcTGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTG**T**ATTTTTGTTCTACAAAATg
TGAA	gatcTGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTG**A**ATTTTTGTTCTACAAAATg
$CTAG_m$	gatcTGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGT**CTAG**TTTTTGTTCTACAAAATg
$TCAG_j$	gatct**CAG**GCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCATTTTTGTTCTACAAAATg
CTAG2	gatc**CTAG**GCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGT**CTAG**TTTTTGTTCTACAAAATg
TAATT	gatcTGCAGCTCAGATTCTTTGTTTGAAAAATTAGCGCTCTCGCGTTGCAT**AA**TTGTTCTACAAAATg
WT_{+20}	gatctGGTAAAATAGCGCTTTCGCGTTGCATTTCTGTTCTGTAAAAATGCAGCTCAGATTCTTTGTTg

[a]The sequence of the origin-to-*STB-distal* orientation strand for each duplex is shown. Sequence added for restriction site overhangs is in lower case. Relevant sequence elements are shown above. Residues altered in duplexes are in bold.

the native locus, a BglII site in the *kanMX4* cassette was first removed by digestion of pKan with BglII and SalI, followed by incubation with Klenow fragment to fill in overhangs, and incubation with ligase to close the plasmid. The *STB-P* sequence was then deleted by digestion with HpaI and AvaI, and the plasmid closed with a linker containing a BglII site flanked by XhoI sites (5′-CCCTCGAGATCTCCTCGAGGG-3′) (Figure 2A).

Synthetic *STB* repeats were inserted at the unique BglII site in pKanΔSTB to create the pKan-STB series of plasmids. Plasmids with inserts in the orientation of the native *STB-P* repeats were identified by restriction mapping and used to assay partitioning function.

To create a control plasmid with the native *STB-P* repeat array reinserted in pKanΔSTB flanked by the XhoI linker sequences, the 0.3-kb HpaI/AvaI *STB-P* fragment was made flush with Klenow and cloned at the EcoRV site of the vector pBluescript (Stratagene). An XhoI linker was introduced at the SmaI site in the vector to produce the plasmid pBS-STBX, from which *STB-P* could be excised as a 346-bp XhoI fragment. This XhoI fragment was ligated with XhoI-digested pKanΔSTB, creating pKan-STB-P, which has the 296-bp native *STB-P* sequence reintroduced in the original orientation and position, but separated from the native flanking sequences on the *ORI* and *STB-D* sides by 26 and 32 bp of linker sequence, respectively.

Plasmids for one-hybrid assays. To test for association of Rep proteins with synthetic *STB* sequences *in vivo*, the pSTB series of plasmids was generated (Supplementary Table S3), in which BglII/BamHI-digested *STB* repeat duplexes were cloned at a BglII site flanked by XhoI sites up-

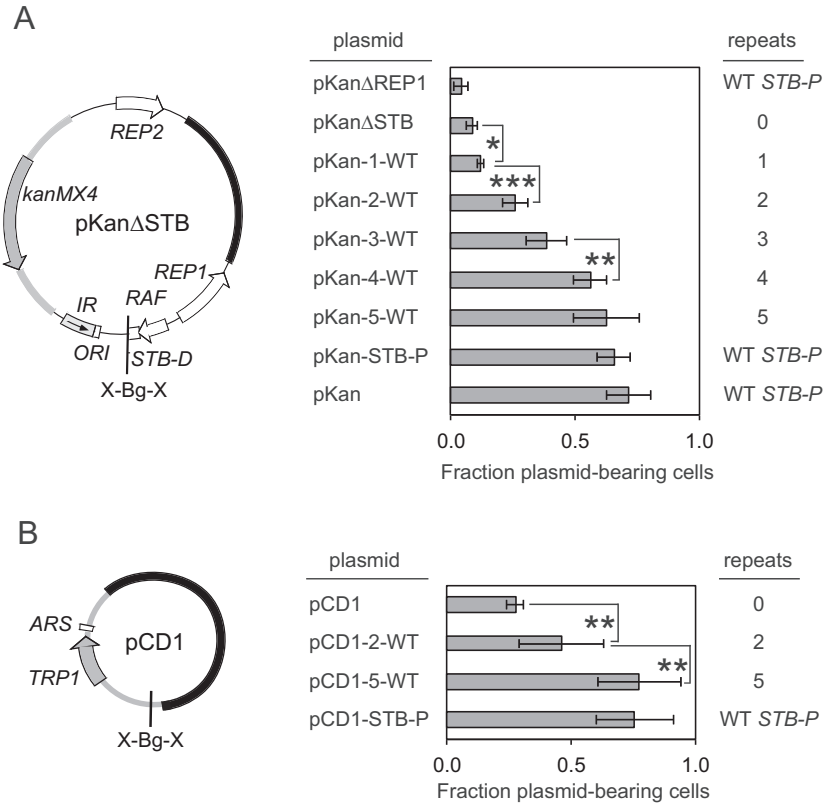

Figure 2. Two or more directly arrayed copies of the synthetic *STB* repeat provide partitioning function. Maps of the parental plasmids, (**A**) *kanMX4*-tagged 2μm-based pKanΔSTB and (**B**) *TRP1*-tagged *ARS*-only pCD1, are shown at left. The BglII site flanked by XhoI sites (X-Bg-X) at which synthetic repeats were inserted is indicated. pKanΔSTB has all 2 μm sequences with the exception of *STB-P*, and these are depicted as in Figure 1; the disrupted *FLP* gene is not shown. For both plasmids, *E. coli* vector sequences (thick black line), yeast chromosomal DNA (thick grey line) and marker genes (grey arrows) are indicated, as is the chromosomal origin of replication (*ARS*) in the pCD1 plasmid. (A) Inheritance of the 2μm-based (pKan) plasmids in a cir^0 yeast strain and (B) of the non-2μm-based (pCD1) plasmids in a cir^+ yeast strain was determined by plating assays. Results in (A) and (B) represent the average (± s.d.) from assaying six and four transformants for each plasmid, respectively. Asterisks indicate significant differences determined by a two-tailed *t*-test (*$P < 0.01$, **$P < 0.005$, ***$P < 0.0005$).

stream of a *lacZ* reporter gene in the *URA3*-tagged one-hybrid vector pJL638 (38). To generate a pSTB plasmid carrying native *STB-P* (pJL638-STB-P), the XhoI fragment from pBS-STBX was inserted in XhoI-digested pJL638.

LEU2-marked plasmids used for expressing Rep1 or Rep2 fused to the transcriptional activation domain of Gal4 ($Gal4_{AD}$; pGAD-REP1 and pGAD-REP2) (11) and *TRP1*-marked plasmids used for expressing untagged Rep1 or Rep2 under the control of the galactose-inducible *GAL1* promoter (pGAL-TRP-REP1 and pGAL-TRP-REP2) have been previously described (9).

Non-2μm-based plasmids for assaying partitioning function. To assess partitioning function conferred by synthetic *STB* repeats in a non-2μm context, the pCD1 series of plasmids was generated (Figure 2B). pJL638-STB-P and pSTB one-hybrid reporter plasmids were digested with BamHI and ScaI, and the fragment containing the inserted *STB* sequence with the same non-2μm flanking sequence for each was introduced into BamHI/NruI-digested YRp7, an *ARS*-only plasmid that consists of a 1.45-kb yeast genomic EcoRI fragment encoding *TRP1* and the adjacent origin of replication cloned at the EcoRI site of *E. coli* vector pBR322 (39).

One-hybrid assays

One-hybrid reporter plasmids (pJL638 and the pSTB series) were linearized by digestion with StuI and used to transform a *ura3-1* cir^0 yeast strain (MD83/1c) to uracil prototrophy, generating a series of yeast strains with no *STB* sequences, or with synthetic *STB* repeats integrated in the genome at the *URA3* locus upstream of a *lacZ* reporter gene (Supplemental Table S1).

For one-hybrid assays, cir^0 strains with integrated reporter genes were co-transformed to leucine and tryptophan prototrophy with pGAD-REP1 and either pGAL-TRP (empty vector) or pGAL-TRP-REP2, or co-transformed with pGAD-REP2 and either pGAL-TRP or pGAL-TRP-REP1. Transformed yeast were cultured until early stationary phase in liquid SD medium lacking tryptophan and leucine (SD-trp-leu) to maintain both plasmids, with galactose as the carbon source to induce expression of the untagged Rep protein (Rep2 for $Gal4_{AD}$-Rep1 assays and Rep1 for $Gal4_{AD}$-Rep2 assays). Activation of the reporter gene, taken as a measure of association of the $Gal4_{AD}$-fusion protein with the inserted *STB* sequence, was determined by measuring β-galactosidase (β-gal) activity using a permeabilized-cell assay (40).

Plasmid inheritance assay

Plasmid inheritance, as a measure of partitioning function, was monitored by determining the fraction of plasmid-bearing cells as previously described (9). Inheritance of 2μm-based pKan plasmids was assayed in *cir*^{0} yeast (AG8/5). Inheritance of the non-2μm-based pCD1 plasmids was assayed in *cir*$^+$ yeast (JP48/2b), where the presence of the native 2μm plasmid provides the Rep proteins *in trans*.

To determine the fraction of plasmid-bearing cells, transformed yeast were cultured for 16 to 24 h (6–8 generations) in selective medium (YPAD+G418 for pKan plasmid transformants or SD-trp for pCD1 transformants). Appropriate dilutions of each culture were plated on solid medium and the fraction of cells containing the pKan or pCD1 plasmid was determined by calculating the ratio of colonies formed on selective (YPAD+G418 or SD-trp) versus non-selective (YPAD or SD) medium.

Chromatin immunoprecipitation (ChIP)

One-hybrid reporter plasmids containing synthetic *STB* repeats were integrated at the *URA3* locus in the genome of a *ura3-1 cir*$^+$ yeast strain (JP48/2b) as described for one-hybrid assay strains to create yeast strains for ChIP analysis (Supplementary Table S1). Yeast strains were cultured in YPAD to a density of ∼1.5 × 10^7 cells/ml. Cells were formaldehyde-fixed and ChIP was performed as previously described (9) using rabbit polyclonal antibodies specific for Rep1 or Rep2 (11) or LexA (Pierce) and Protein A Sepharose CL-4B (GE Healthcare). Enrichment of DNA loci of interest in immunoprecipitates was evaluated by semi-quantitative polymerase chain reaction (PCR) using Platinum Taq (Invitrogen) as recommended by the supplier. Primers used are listed in Supplementary Table S4. DNA templates were diluted to give products within the linear range of 30 cycles of PCR. PCR products were separated by agarose gel electrophoresis, stained with ethidium bromide, imaged on a VersaDoc 4000 MP with Quantity One software and quantified with ImageLab software (BioRad). Yield of PCR products from immunoprecipitated DNA was compared to that from input DNA to give locus immunoprecipitation as a percentage of input DNA.

RESULTS

Two copies of a synthetic *STB* repeat are sufficient to confer 2μm plasmid partitioning function

The *STB-P* locus of the 2μm plasmid variant found in most laboratory strains of *Saccharomyces cerevisiae* (Scp1) consists of five tandemly-arrayed repeats (Figure 1A); however, deletion studies and the sequences of other 2μm plasmid variants suggest that three or fewer *STB* repeats are sufficient to confer Rep protein-dependent partitioning function (41–43). To determine whether differences in repeat number required might reflect sequence variation between repeats, we generated a synthetic *STB* repeat based on a single 63-bp portion of the *STB-P* sequence (Figure 1B), and used it to build arrays of one to five *STB* repeats. The sequence was chosen based on containing the perfect match to a core consensus present in all five Scp1 repeats (26). These repeat arrays differed from the native *STB-P* in two respects: first, all repeats within the synthetic array were identical, and second, the linker sequence (GGATC) added between the repeats to facilitate ligation resulted in arrays in which sequences were spread 5 bp farther apart than in the native *STB-P*.

The synthetic *STB* arrays were introduced into a *kanMX4*-tagged 2μm-based plasmid from which the native *STB-P* sequence had been removed and replaced with a 21-bp linker sequence (pKanΔSTB) (Figure 2A). This generated a series of plasmids with one to five copies of the synthetic *STB* repeat directly arrayed in the position and orientation of the native *STB-P* (pKan-1-WT through pKan-5-WT). As a control for potential disruptive effect from the linkers, the native *STB-P* sequence, with approximately 30 bp of flanking sequence on either end, was also inserted at the linker sequence in pKanΔSTB (pKan-STB-P). A pKan-based plasmid with an unperturbed *STB-P* locus but lacking the *REP1* coding region (pKanΔREP1) was used to establish the defect expected for a plasmid that lacks all Rep protein-dependent partitioning, in this case, due to lack of Rep1 protein expression. Inheritance of a pKan plasmid with an unmodified *STB* locus (pKan), which we have previously shown to be efficiently partitioned in *cir*^{0} yeast (9), was included for comparison.

Plasmids were introduced into a strain lacking native 2μm plasmid (*cir*^{0}), and the fraction of cells containing plasmid determined using a plating assay. The inheritance of the 2μm-based plasmid in which the native *STB-P* had been removed and re-inserted (pKan-STB-P) did not differ significantly from the pKan plasmid where the native *STB-P locus* had not been disrupted ($P = 0.24$) (Figure 2A). Thus the linker sequences used to facilitate the insertion of synthetic and native *STB-P* sequences were not detrimental to *STB* partitioning efficiency.

The fraction of cells holding the plasmid lacking the native *STB-P* array (pKanΔSTB) was slightly higher than observed for the plasmid with an intact *STB* locus but lacking the *REP1* coding region (pKanΔREP1) ($P < 0.05$) suggesting that 2μm plasmid sequences other than those in *STB-P* were providing some limited Rep protein-dependent partitioning function. However, the fraction of cells that were able to form colonies on selective medium was very low for both plasmids, indicating that, as expected, neither was being efficiently partitioned.

Plasmids carrying only a single copy of the synthetic *STB* repeat (pKan-1-WT) showed a slight improvement in inheritance over the plasmid lacking *STB-P* (pKanΔSTB) (0.12 vs. 0.08 for fraction of plasmid-bearing cells, $P < 0.01$) suggesting one repeat might be conferring some limited partitioning function. With the introduction of each additional synthetic repeat up to five copies (pKan-2-WT through pKan-5-WT), further increases in the fraction of cells containing plasmid were observed with the most striking improvement being when repeat number was increased from one (pKan-1-WT) to two copies (pKan-2-WT) (0.12 versus 0.26 for fraction of plasmid-bearing cells, $P < 0.0005$). Above five repeats, this trend of improvement did not appear to continue, as a plasmid containing eight synthetic *STB* repeats had similar inheritance to that of a plasmid

containing five repeats ($P = 0.49$) (Supplementary Figure S1). These observations suggest that increased repeat copy number improves partitioning function, but that this improvement may plateau above five repeats.

Inheritance of the plasmid with the array of five synthetic *STB* repeats, (pKan-5-WT) did not differ significantly from the two plasmids with native STB-P repeat arrays, pKan ($P = 0.21$) and pKan-STB-P ($P = 0.42$) (Figure 2A). These results imply that the synthetic *STB* repeats are functionally equivalent to the native repeats in this context, and that arrays of synthetic *STB* repeats are therefore valid tools for testing the effects of *STB* repeat number and sequence composition on function. This result also suggests that slight alterations in the spacing of the repeats do not significantly impact function, and that the 63-bp region of *STB-P* used as the basis for the synthetic repeat contains all sequence elements required for partitioning.

A second 63-bp synthetic *STB* repeat, with a starting position shifted 20 nt toward the *ORI* relative to the original synthetic *STB* repeat (Table 1), was also tested. Two copies of the shifted repeat were seen to provide a similar degree of plasmid inheritance to the two synthetic *STB* repeats shown here ($P = 0.61$) (Supplementary Figure S2); therefore, the partitioning function conferred by two synthetic repeats was not unique to the original repeat frame chosen.

Synthetic *STB* sequences confer partitioning in the absence of any other 2μm plasmid sequences

To determine whether the non-*STB-P* 2μm sequences responsible for the slight improvement in plasmid partitioning function observed for pKanΔSTB over pKanΔREP1 (Figure 2A) were required for the function of synthetic *STB* sequences, synthetic and native *STB* sequences were placed on a *TRP1*-tagged *ARS*-only plasmid (pCD1) containing no other 2μm sequences (Figure 2B). The partitioning function of this pCD1-based series of plasmids was assessed in a strain containing native 2μm plasmid (*cir*$^+$), which supplied the Rep1 and Rep2 proteins *in trans*. The inheritance of a pCD1 plasmid carrying two synthetic *STB* repeats was significantly improved over that of a pCD1 plasmid carrying no repeats (Figure 2B). This observation demonstrates that two copies of the synthetic *STB* repeat can mediate plasmid partitioning in the absence of other 2μm plasmid sequences.

No significant difference in inheritance was observed between pCD1 plasmids carrying five arrayed copies of the synthetic *STB* repeat or native *STB-P* ($P = 0.30$), with both plasmids giving similar fractions of plasmid-bearing cells (Figure 2B). The efficient inheritance of both supports prior observations that the *STB-P* repeats are sufficient for Rep-mediated partitioning with *STB-D* serving an ancillary role (24) and demonstrates that our synthetic *STB* sequences are no more reliant on a 2μm sequence context for function than is the native *STB-P* sequence. Therefore, despite potential small contributions from non-*STB-P* 2μm sequences, the remainder of the plasmid inheritance assays in this study were conducted using pKan-based plasmids, to avoid the need for native 2μm or a second plasmid to supply the Rep proteins, and to eliminate the selective advantage associated with plasmid-borne nutritional markers that can mask differences in plasmid partitioning efficiency.

Sequence elements within the *STB* repeat required for partitioning function

Having established that two copies of the synthetic *STB* repeat were sufficient to confer partitioning function, we wanted to assess the significance of a TGCATTTTT motif, identified bioinformatically in two previous studies as a putative Rep protein binding motif (29,30). The TGCATTTTT sequence occurs in the middle of three of the five native *STB* repeats, and also in the inverse orientation at two of the junctions between native repeats (Figure 1A). The 63-bp region of *STB-P* used as the basis for our synthetic *STB* repeat contains one of the middle TGCATTTTT sequences and the TGCA component from the junction sequence (Figure 1B).

To determine whether the middle TGCATTTTT or junction TGCA was required for the partitioning function of our synthetic *STB* repeat, a series of 2μm-based plasmids was created. Each of these contained two directly arrayed copies of the synthetic *STB* repeat, in which either a portion of the TGCATTTTT sequence or the TGCA sequence was mutated in both copies of the repeat. The abilities of these plasmids to be partitioned were then assessed using an inheritance assay.

A single base change in the TGCA portion of the middle TGCATTTTT to TGTA (pKan-2-TGTA) had no effect on plasmid inheritance, while a change to TGAA (pKan-2-TGAA) led to a decrease in the fraction of plasmid-bearing cells (Figure 3). Scrambling of the nucleotides in this TGCA sequence to CTAG (pKan-2-CTAGm) led to a greater decrease in plasmid inheritance, indicating the importance of the TGCA element in the middle of the *STB* repeat for partitioning function, and demonstrating that a pyrimidine but not a purine substitution is tolerated at the third position of this sequence.

To investigate whether the TGCA sequence found at the junction of the *STB* repeats might also contribute to partitioning function, the effect of scrambling this element to TCAG (pKan-2-TCAGj) was assessed. The TCAGj mutation led to a decrease in the fraction of plasmid-bearing cells similar to that observed when the middle TGCA was scrambled. Scrambling both the middle and junction TGCA sequences simultaneously (pKan-2-CTAG2) led to decrease in the fraction of plasmid-bearing cells greater than that resulting from mutation of either TGCA sequence alone, implying that the middle and junction TGCA provide non-redundant contributions to *STB* partitioning function. The fraction of plasmid-bearing cells when both TGCA sequences were scrambled (pKan-2-CTAG2) was lower than that observed when no native or synthetic *STB-P* sequences were present (pKanΔSTB) ($P < 0.005$), and similar to that observed when the *REP1* gene was deleted (pKanΔREP1) ($P = 0.89$), suggesting that replacement of native *STB-P* with sequences lacking any TGCA motifs may eliminate the Rep protein-dependent partitioning function conferred by non-*STB-P* 2μm plasmid sequences.

The importance of the T-tract in the middle TGCATTTTT sequence was also investigated. Replacement of

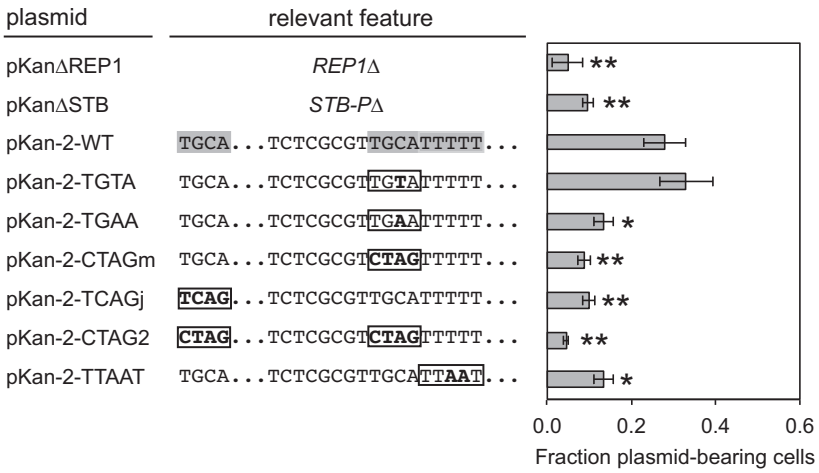

Figure 3. Mutations in TGCA sequences and a TGCA-adjacent T-tract impair partitioning function provided by two directly arrayed copies of the synthetic *STB* repeat. A *cir*^{0} yeast strain was transformed with the indicated pKan plasmid. Transformants were assayed for partitioning function as described in the legend of Figure 2. Results represent the average (± s.d.) from assaying four independent transformants for each plasmid. Asterisks indicate significant differences compared to wild-type synthetic repeats determined by a two-tailed *t*-test (*$P < 0.005$, **$P < 0.0005$). Sequence elements targeted for analysis are shaded in gray, and mutated elements are boxed, with altered residues in bold. Full sequences of WT and mutant synthetic *STB* repeats are given in Table 1.

the third and fourth thymines in the tract with adenines (pKan-2-TTAAT) led to a decrease in the fraction of plasmid-bearing cells (Figure 3), although this decrease was not as great as that observed when either of the TGCA sequences were scrambled. This finding demonstrates that the T-tract of the middle TGCATTTTT also contributes to efficient plasmid partitioning.

Mutation of TGCA sequence impairs association of Rep proteins with one and two tandemly-arrayed *STB* repeats

Having established the importance of the TGCATTTTT element for partitioning function, we wanted to determine whether this motif was required for Rep protein association. We have previously shown that Rep protein associations with *STB-P* detected using a one-hybrid assay reflect the association of native Rep proteins with *STB-P* as detected by chromatin immunoprecipitation (ChIP) (9).

To assay the Rep proteins for interaction with the synthetic *STB* repeats each was expressed fused to the Gal4 transcriptional activation domain (Gal4$_{AD}$) in a *cir*^{0} strain where either no, one, or two copies of wild-type (WT) or mutant synthetic *STB* repeats had been inserted upstream of a chromosomally-integrated *lacZ* reporter gene. Expression of Gal4$_{AD}$-Rep1 in the absence of Rep2 activated the reporter in strains with one or two copies of the WT synthetic repeat (Figure 4A). This activation was significantly above the level observed when Gal4$_{AD}$ alone was expressed (Supplementary Figure S3) or when no *STB* sequence was present (Figure 4A), demonstrating that a single copy of the WT *STB* repeat is sufficient for Rep1 protein association and that Rep1 does not require Rep2 for this interaction. Gal4$_{AD}$-Rep2 expression in the absence of Rep1 failed to activate the reporters (Figure 4A), consistent with previous reports that the presence of Rep1 is required for robust association of native Rep2 with the 2 μm *STB* locus and for Gal4$_{AD}$-Rep2 association with the native *STB-P* sequence to be detected in a one-hybrid assay (9).

When the Gal4$_{AD}$-Rep fusions were assayed for interaction with the synthetic *STB* repeats in cells where the partner Rep protein was co-expressed from a galactose-inducible promoter on a second plasmid both Gal4$_{AD}$-Rep1 and Gal4$_{AD}$-Rep2 activated reporters downstream of one or two copies of the WT repeat significantly above the level observed when no *STB* sequence was present (Figure 4B). Activation of the reporters by Gal4$_{AD}$-Rep1 was increased above the level observed in the absence of Rep2 which could reflect a higher steady-state level for the Rep1 fusion protein; Rep1 protein levels are lower in the absence of Rep2 (9). The results show that both Rep proteins can associate with a single copy of the 63-bp *STB-P* sequence, despite one repeat not being adequate for efficient partitioning when inserted in place of the native *STB-P* on the pKan plasmid (Figure 2). Reporter activity increased approximately two-fold for both Gal4$_{AD}$-Rep1 and Gal4$_{AD}$-Rep2 when two copies of the repeat were present upstream of the reporter rather than one, consistent with the number of Rep protein association sites being doubled.

To determine whether mutations in the TGCATTTTT sequence that led to loss of partitioning function altered Rep1 or Rep2 association, mutant *STB* repeats were assessed using the one-hybrid assay. For a single copy of the synthetic *STB* repeat, association of Gal4$_{AD}$-Rep1 was impaired by mutation of the middle TGCA (TGAA and CTAGm) and almost completely abolished when the junction TGCA was scrambled, either alone (TCAGj), or in conjunction with the middle TGCA (CTAG2) in both the absence (Figure 4A) and presence of Rep2 (Figure 4B). The results suggest that the junction TGCA sequence is required for Rep1 protein association with a single copy of the *STB-P* sequence and that the middle TGCA sequence can contribute to this association, but is not sufficient to mediate association of Rep1 when the junction TGCA sequence is scrambled.

In contrast to Gal4$_{AD}$-Rep1 association, Gal4$_{AD}$-Rep2 association with a single copy of the synthetic *STB* re-

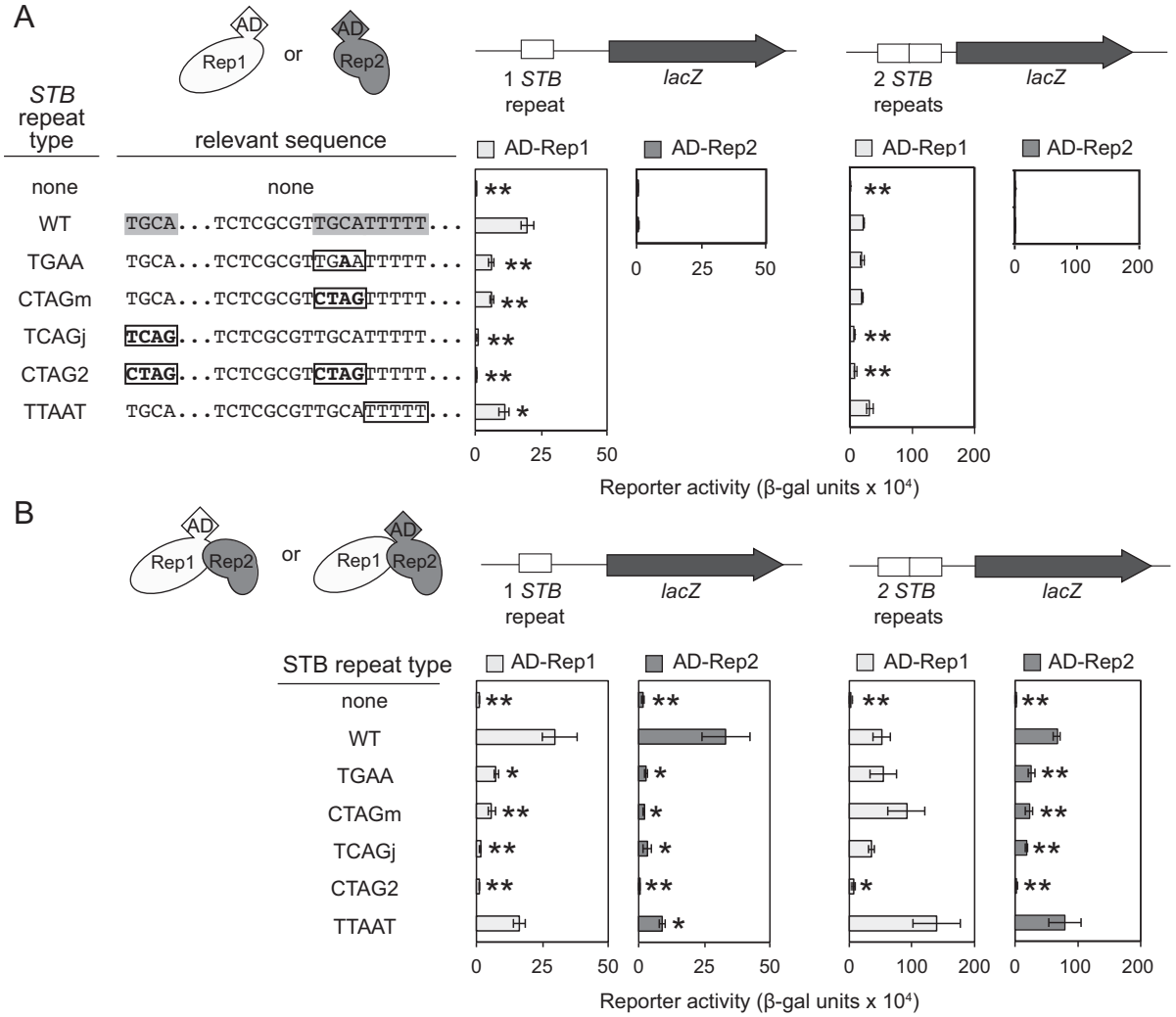

Figure 4. One-hybrid assay for *in vivo* association of Rep1 and Rep2 with synthetic *STB* repeats. Yeast strains lacking native 2μm (*cir*^{0}) and with no, one or two copies of WT or mutant synthetic *STB* repeats integrated in the chromosome upstream of a *lacZ* reporter gene were co-transformed with two plasmids, one expressing either Rep1 or Rep2 as Gal4$_{AD}$-fusion proteins (AD-Rep1 or AD-Rep2) and the second (**A**) an empty vector, or (**B**) a plasmid expressing the respective partner Rep protein untagged. Co-transformants were assayed for activity of the *lacZ* gene product, ß-gal, which indicates association of the Gal4$_{AD}$-Rep fusion protein with the synthetic *STB* repeat of interest. Results represent the average (±s.d.) from assaying four independent co-transformants for each strain. Asterisks indicate significant differences compared to wild-type synthetic repeats determined by a two-tailed *t*-test (*$P <$ 0.005, **$P < 0.0005$). Sequence elements targeted for analysis are shaded in gray, and mutated elements are boxed, with altered residues in bold. Full sequences of WT and mutant synthetic *STB* repeats are given in Table 1.

peat was equally impaired when either the middle TGCA was mutated to TGAA, or the middle or junction TGCA sequences were scrambled, although as for Gal4$_{AD}$-Rep1, all association was lost when both TGCA elements were scrambled. These results suggest that Rep2 may have more stringent sequence requirements than Rep1 for association with a single *STB* repeat.

Upon expression of Gal4$_{AD}$-Rep1 in strains with two directly arrayed copies of the synthetic *STB* repeat upstream of the reporter, reporter activity was not significantly decreased when the middle TGCA sequence in both repeats was mutated, irrespective of the presence or absence of Rep2, or when the junction TGCA was scrambled, provided Rep2 was present. Association was significantly impaired when these elements were scrambled simultaneously (CTAG2) (Figure 4A and B). These results suggest that the presence of two repeats may allow Rep1 to associate with DNA in a cooperative manner that bypasses the requirement for two intact TGCA sequences in each repeat. The Rep2-dependent improvement in Rep1 association when the junction TGCA elements were mutated in both copies might suggest a direct role for Rep2 in promoting Rep1 association with the middle element or could be a consequence of increased Rep1 fusion protein levels in the presence of Rep2.

Upon Gal4$_{AD}$-Rep2 expression in strains with two copies of the synthetic *STB* repeat, reporter activity was reduced when the repeats had mutations in either the middle or junction TGCA sequence, but not to the extent observed when no *STB* sequences were present, suggesting that, un-

like a single copy of these mutant *STB* repeats, two copies allowed some Rep2 association. However, when both the core and junction TGCA motifs were scrambled simultaneously, $Gal4_{AD}$-Rep2 failed to activate reporter gene expression above the level observed when no *STB* repeats were present. These results suggest that both the middle and junction TGCA motifs contribute similarly to Rep2 association with *STB*, and that Rep2 requires a minimum of two intact TGCA motifs for association with *STB* to be detected in this assay system. These results are consistent with those obtained with a single copy of the *STB* repeat, implying that sequence requirements for Rep2 association with *STB* are stricter than those of Rep1.

Mutation of the T-tract in the TGCATTTTT motif to TTAAT did not significantly reduce $Gal4_{AD}$-Rep1 or $Gal4_{AD}$-Rep2 association with two copies of the repeat in the presence of their partner proteins although association of $Gal4_{AD}$-Rep2 with a single copy was reduced (Figure 4B) as was that of $Gal4_{AD}$-Rep1 with the solo repeat in the absence of Rep2 (Figure 4A). The improved association with two mutant copies relative to one could be due to cooperative interactions between Rep proteins bound at adjacent repeats. As substantial improvements in plasmid inheritance were only conferred by synthetic *STB* repeats in arrays of two or more, this result suggests that the partitioning defect associated with the TTAAT mutation, unlike those associated with mutations of the TGCA sequences, must be due to some effect other than loss of Rep protein association with these repeats. The T-tract may be required to allow the appropriate configuration of the Rep proteins, such as higher-order interactions that are not measured in this assay. The slight but consistent increase in activation of the reporter gene when $Gal4_{AD}$-Rep1 and $Gal4_{AD}$-Rep2 were expressed in strains with two copies of the TTAAT mutant repeat could indicate that Rep proteins associate with *STB* more stably or more abundantly when the T-tract is mutated.

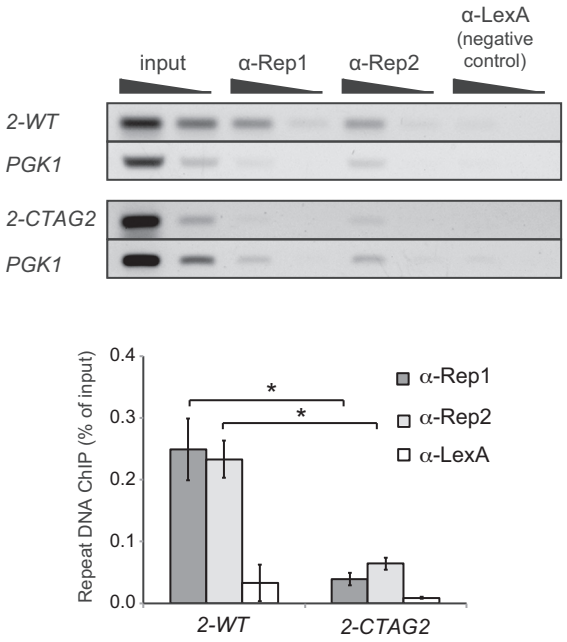

Figure 5. Native Rep protein association with two directly-arrayed copies of the synthetic *STB* repeat is impaired when the TGCA elements are scrambled. Chromatin was immunoprecipitated from yeast strains containing native 2μm (*cir*$^+$) and with an array of two WT or two CTAG2-mutant synthetic *STB* repeats integrated in the genome. Antibodies specific for native Rep1 or Rep2 and, as a negative control, anti-LexA were used. The precipitated DNA was analyzed by semi-quantitative PCR with primers that would amplify the region containing the integrated repeats and, as a negative control, a chromosomal locus (*PGK1*). Inverse images of ethidium bromide-stained agarose gels of PCR products obtained from serial 1:10 dilutions of each template from a representative assay are shown (top). Template DNA amplified in 'input' PCR reactions is 2.5% of that amplified in 'ChIP' PCR reactions. The bar graph (below) indicates ChIP efficiency as the percent of input DNA immunoprecipitated; results are average (± s.d.) from four assays. Asterisks indicate significant differences as determined by a two-tailed *t*-test (*$P < 0.0005$).

Scrambling of both TGCA elements impairs association of native Rep proteins with two tandemly-arrayed *STB* repeats

To determine whether native Rep proteins expressed at endogenous levels would associate with the synthetic *STB* repeats in a manner similar to that observed for the $Gal4_{AD}$-Rep fusion proteins in the one-hybrid assay, tandem arrays of two copies of the synthetic repeat, 2-WT or 2-CTAG2, were integrated in the genome of a *cir*$^+$ yeast strain where the native Rep proteins would be expressed from the resident 2μm plasmid. Association of the Rep proteins with the repeat-containing loci was assessed by ChIP (Figure 5). The WT *STB* repeats were immunoprecipitated significantly more efficiently than the mutant repeats with antibodies specific for Rep1 and Rep2, supporting the one-hybrid assay results indicating that the Rep proteins associate with the synthetic *STB* repeats and that scrambling of the TGCA elements in the repeats negatively impacts this association.

DISCUSSION

Equal partitioning of the yeast 2μm plasmid during host cell division is dependent on the association of the plasmid-encoded Rep1 and Rep2 proteins with a region of repeated sequence at the plasmid *STB* locus (*STB-P*). In this study, we created synthetic versions of the *STB* repeats and used these to identify DNA sequence elements required for association of Rep1 and Rep2 and for plasmid inheritance. We showed that a pair of directly repeated copies of a single 63-bp stretch of *STB-P* is sufficient to confer partitioning function in the absence of any other 2μm plasmid sequences, and that mutation of either of two TGCA sequences or of a TGCA-adjacent T-tract impairs partitioning. We also found that mutation of either TGCA sequence, but not of the T-tract, decreased Rep2 protein association with a pair of direct *STB* repeats while scrambling of both TGCA elements significantly diminished recruitment of both Rep1 and Rep2 to the repeats. Overall, these findings improve our understanding of the interaction between Rep proteins and *STB-P* DNA, and provide insight into the DNA sequence requirements for formation of a functional *STB-P* partitioning complex.

Effect of *STB* repeat copy number on partitioning function

Previous deletion analyses and surveys of natural 2μm plasmid variants have provided clear evidence that fewer than five tandemly-arrayed native *STB* repeats are sufficient for plasmid partitioning (42,44). Our analysis showed that even a single *STB* repeat could confer a slight improvement in plasmid inheritance, while an array of two repeats was markedly more efficient. The length provided by two repeats (131-bp) would be sufficient to fully accommodate a nucleosome containing the centromere-specific histone H3 variant Cse4 (45). Cse4 recruitment to the native 2μm *STB* repeat array is dependent on the Rep proteins and is a prerequisite for RSC2-mediated remodeling of the *STB* chromatin into a partitioning-competent form (16,28). Two synthetic *STB* repeats might allow a nucleosomal configuration that facilitates replacement of the regular histone H3 with Cse4. A further increase in inheritance efficiency was observed when synthetic *STB* repeat number was increased from three to four (Figure 2A). This might indicate an additional benefit from the array accommodating two nucleosomes, the number that nuclease-sensitivity assays suggest occupy the five-repeat native *STB* array (13,27). Notably, adding repeats above this number did not confer any further improvement.

Further support for the importance of nucleosomal organization at *STB* is the high rate of 2μm plasmid missegregation when yeast lack a functional version of RSC2 (13), one of the two RSC chromatin remodelers found in yeast (46). RSC complexes use the energy from ATP hydrolysis to actively move or destabilize nucleosomes (47,48) (for review see (49)), and have been implicated as key determinants of chromatin structure at gene promoters, maintaining nucleosome arrangements that can restrict or facilitate access of trans-acting factors to DNA binding sites (49–52). Appropriate positioning of *STB* repeat sequence elements within an RSC2 complex-positioned nucleosomal context may contribute to their recognition by the Rep proteins and host factors.

Two copies of the repeat tested here were able to supply partitioning function even in absence of other 2μm flanking sequences (Figure 2B). Three or more repeats are found at the *STB* loci of most 2μm plasmid variants although some have two (41–43). Fewer repeats may be required if flanking regions, in addition to encoding silencing and termination signals that protect the *STB* locus from transcriptional interference (24), contain sequence elements that contribute to appropriate nucleosomal organization of the repeats or to Rep protein or host factor association.

Sequence variation between repeats in native *STB* arrays may also determine the number needed for formation of a functional partitioning complex. 2μm plasmid variants may require three or more repeats in their *STB* arrays to supply the level of partitioning proficiency needed for maintenance in the natural context. The synthetic *STB* repeat used in this study was based on the most origin-distal repeat of the Scp1-variant 2μm plasmid *STB* locus (Figure 1). Previous deletion analysis of the Scp1 *STB* array showed that removal of all but the two origin-proximal repeats led to loss of partitioning function and failure to recruit Cse4, while retaining the three origin-proximal repeats did not, suggesting a minimum of three repeats was required for *STB* function (25,28). However, the most origin-proximal repeat in Scp1 is the most divergent of the repeats in the array (Figure 1A) and may function less efficiently than the other repeats. Testing synthetic repeats based on the sequence of other *STB* repeats will allow this possibility to be assessed.

The notable improvement in plasmid inheritance observed for two copies relative to a single *STB* repeat could also reflect the impact of interactions between the Rep proteins and their self-associations (53–55) that would allow them to participate in higher-order interactions upon recruitment to repeated recognition sites in adjacent repeats at the *STB* locus. A single copy of the *STB* repeat, while sufficient for recognition by the Rep proteins, may not be enough to nucleate the formation of stable higher-order protein complexes due to insufficient repetition of the Rep protein recognition sites.

The partitioning loci of many bacterial and viral plasmids, like *STB-P*, are characterized by multiple repeats of their partitioning protein binding sites, which suggests that higher-order associations may be a common strategy for establishing a partitioning locus. The partitioning system of the plasmid prophage of P1 bacteriophage, like that of the 2μm plasmid, consists of two partitioning proteins, ParA and ParB, that associate with a plasmid locus, called *parS*. The *parS* site contains two motifs, called BoxA and BoxB, and two copies of each motif nucleate association of multiple ParB proteins and confer partitioning function to P1 (reviewed in (56)). Similarly, the partitioning proteins of several viral episomes, including EBNA1 of Epstein-Barr virus, LANA1 of Kaposi's sarcoma-associated herpesvirus, and E1 and E2 of bovine papilloma virus, also recognize repeated sequences, which allow them to associate cooperatively with their partitioning loci (reviewed in (21)).

Rep protein association with TGYA motif

The sequence TGCATTTTT was originally identified bioinformatically as a potential Rep protein binding motif (29,30). Our results, which are the first experimental test of this hypothesis, suggest that the TGCA portion of this sequence is more critical than the T-tract for Rep protein association with the *STB* repeats (Figure 4). Despite impairing plasmid inheritance (Figure 3), mutation of the T-tract to TTAAT only seemed to significantly affect Rep2 association with a single copy of the *STB* repeat and indeed seemed to enhance association of both Rep proteins with two copies of the repeat. In contrast, when either of the two TGCA sequences in the synthetic *STB* repeat were scrambled, Rep1 association with one repeat, Rep2 association with one and two repeats, and partitioning function were impaired.

We noted a difference between the middle and junction TGCA sequences with respect to the relative effect of mutations in these on the two Rep proteins. For wild-type and most mutant versions of the synthetic *STB* repeat tested in our one-hybrid assays, Rep2 association generally reflected Rep1 association (Figure 4B), consistent with Rep2 association with *STB-P* being bridged by Rep1 (9). However, when the middle TGCA was scrambled or mutated to TGAA, Rep1 association was less severely impaired than when the

junction TGCA was scrambled, while Rep2 association was similarly impaired for both. When two copies of the repeat contained scrambled junction TGCA elements, the presence of Rep2 restored Rep1 association, but did not do so if the middle TGCA element was also scrambled (Figure 4B). The two TGCA sequences were also not functionally redundant, as scrambling both led to a greater defect in both plasmid inheritance (Figure 3) and Rep2 protein association (Figure 4) than scrambling either alone. These observations suggest that Rep2 may have sequence requirements for *STB-P* association that differ from those of Rep1 and render Rep2 more sensitive than Rep1 to changes in the middle TGCA. Mutation of the middle TGCA may indirectly affect Rep2 association by perturbing chromatin conformation. Lack of a functional version of Cse4 has been found to lead to loss of Rep2 association with *STB*, but not that of Rep1 (16). The loss of the middle TGCA might also directly affect Rep2 association. A direct role for Rep2 in recognition of *STB-P*, independent of Rep1 bridging, is supported by *in vitro* assays in which Rep2 has been shown to have DNA-binding ability with a preference for binding *STB-P* (11). Whether recognition of the *STB-P* sequence by Rep1 is direct or dependent on host proteins remains unclear.

It is possible that the apparent greater impact of mutation of the junction TGCA over that of the middle TGCA on Rep1 association is a consequence of our experimental design, as the junction TGCA, like that in the middle of the repeat, is flanked by a T-tract in the context of the native *STB-P*. The linker sequence flanking our synthetic *STB* repeats displaces this T-tract from its position immediately adjacent to the junction TGCA, and further disruption of this region by scrambling the TGCA sequence may be particularly detrimental to Rep protein association. This difference from the native *STB-P* sequence context cannot, however, explain the discrepancies between Rep1 and Rep2 association with the synthetic *STB* repeats in this assay.

A conserved core repeat sequence present at partitioning loci of all 2μm plasmid variants

Natural 2μm plasmid variants can be categorized into one of two groups, based primarily on differences in their *STB* partitioning locus sequence. The laboratory 2μm plasmid, Scp1, and most industrial 2μm plasmid variants are Type-II (44). The Type-II *STB* regions vary primarily in the number of repeats they encode, with the number ranging from three to seven (41–43). Type-I variants have *STB* loci that lack the obvious tandem array structure found in Type-II variants, but have two repeats that are a 70% match to each other and show 70% identity with Type-II *STB-P* repeats (41–43). The Type I repeats also lack TGCATTTTT elements that had been predicted to be the Rep protein recognition motifs.

In our study, mutation of the middle TGCA sequence in our synthetic *STB* repeats to TGTA did not impair partitioning or Rep protein association (Figure 3) (Supplementary Figure S4), while a change to TGAA did, implying that the third position of the TGCA motif need only be a pyrimidine for Rep protein recognition. Alignment of the repeats of a representative Type-I *STB* sequence with those of a Type-II 2μm variant (Scp1) reveals a conserved 23-bp 'core' (Supplementary Figure S5), within the 25-bp region previously identified as being highly conserved in *STB* repeats of Type-II variants (Figure 1) (26). Within this core, the *STB* Type-I variants have the sequence TGTA instead of TGCA, a difference which our results show should not perturb Rep protein recognition or partitioning function. This finding may explain why *STB* loci from Type-I 2μm plasmids function as efficiently as Type-II *STB* loci in the presence of Rep1 and Rep2 proteins encoded by a Type-II 2μm plasmid (44). This TGTA motif is followed by a run of thymines, as seen in the Type II *STB* repeat; however, the run is only three or four thymines in length. Our analysis demonstrated that adenine substitutions in this tract impaired partitioning; therefore, further study will be required to determine whether the T-tract in the Type-I *STB* repeat core is less critical due to other features of the sequence or if a run of three-to-four thymines would also be sufficient for function of the Type-II *STB* core. Taken together, our analyses and the sequence similarities between Type-I and Type-II 2μm plasmid *STB* repeats suggest that Rep protein recognition is dependent on a TGYA element (where Y is a pyrimidine), and that a minimum of two copies of a conserved 23-bp core sequence and at least one TGYA sequence outside the core per repeat may be required to establish a fully functional Rep protein partitioning complex at *STB*. Further experimental studies are required to test these predictions.

Although mutation of the T-tract in the TGCATTTTT motif of the Type II *STB* repeat did not prevent Rep protein association with two directly arrayed synthetic *STB* repeats, this mutation did lead to a reduction in partitioning competency. The T-tracts in the regularly-spaced TGCATTTT motifs in the *STB* repeat array may regulate nucleosome positioning at *STB-P* or confer properties needed for association with the centromere-specific nucleosome (16). A- and T-tracts as short as four nucleotides in length have been shown to influence both translational and rotational positioning of DNA relative to nucleosomes, with AT-rich regions being preferentially associated with nucleosome-free regions (reviewed in (57)). A- and T-tracts have also been shown to activate RSC-mediated displacement of nucleosomes (48) (52). Although tracts of seven residues were more effective, even five-residue runs were reported to stimulate RSC chromatin remodeling *in vitro* (48). The negative impact on plasmid inheritance when the five-residue T-tracts in the synthetic *STB* repeats were changed to TAATT could be due to the RSC2 chromatin remodeler being less active on these sequences, impairing conversion of the *STB* chromatin to a partitioning-competent form (13).

The results presented in this study demonstrate that synthetic *STB* sequences built from directly arrayed repeats can provide valuable insight into the function of the yeast 2μm plasmid partitioning system. We are now poised to exploit the synthetic *STB* repeat approach to identify and characterize sequence elements and protein interactions that contribute to partitioning competence of the 2μm plasmid in its eukaryotic host.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Joyce S.K. Chew for providing technical assistance, Christine Desroches Altamirano for construction of the vector pCD1 and for performing plasmid inheritance assays, Niroshaathevi Arumuggam for performing beta-galactosidase and plasmid inheritance assays, Holly Rasmussen for performing plasmid inheritance assays, Richard Singer and Jordan B. Pinder for comments on the manuscript, and members of the Dalhousie University Yeast Molecular Biology Group for helpful discussions.

FUNDING

Natural Sciences and Engineering Research Council of Canada [Discovery Grant 155268 to M.J.D.]; Killam predoctoral and NSERC CGSM and PGSD scholarships (to M.E.M.); NSERC USRA scholarship (to E.J.P.). Funding for open access charge: Natural Sciences and Engineering Research Council of Canada.

Conflict of interest statement. None declared.

REFERENCES

1. Futcher,A.B. and Cox,B.S. (1983) Maintenance of the 2μm circle plasmid in populations of *Saccharomyces cerevisiae*. *J. Bacteriol.*, **154**, 612–622.
2. Mead,D.J., Gardner,D.C.J. and Oliver,S.G. (1986) The yeast 2μm plasmid: strategies for the survival of a selfish DNA. *Mol. Gen. Genet.*, **205**, 417.
3. Kikuchi,Y. (1983) Yeast plasmid requires a *cis*-acting locus and two plasmid proteins for its stable maintenance. *Cell*, **35**, 487–493.
4. Jayaram,M., Li,Y.-Y. and Broach,J.R. (1983) The yeast plasmid 2μm circle encodes components required for its high copy propagation. *Cell*, **34**, 95–104.
5. Volkert,F.C. and Broach,J.R. (1986) Site-specific recombination promotes plasmid amplification in yeast. *Cell*, **46**, 541–550.
6. Rizvi,S.M.A., Prajapati,H.K. and Ghosh,S.K. (2018) The 2 micron plasmid: a selfish genetic element with an optimized survival strategy within *Saccharomyces cerevisiae*. *Curr. Genet.*, **64**, 25–42.
7. Yang,X.M., Mehta,S., Uzri,D., Jayaram,M. and Velmurugan,S. (2004) Mutations in a partitioning protein and altered chromatin structure at the partitioning locus prevent cohesin recruitment by the *Saccharomyces cerevisiae* plasmid and cause plasmid missegregation. *Mol. Cell Biol.*, **24**, 5290–5303.
8. Murray,A.W. and Szostak,J.W. (1983) Pedigree analysis of plasmid segregation in yeast. *Cell*, **34**, 961–970.
9. Pinder,J.B., McQuaid,M.E. and Dobson,M.J. (2013) Deficient sumoylation of yeast 2-micron plasmid proteins Rep1 and Rep2 associated with their loss from the plasmid-partitioning locus and impaired plasmid inheritance. *PLoS One*, **8**, e60384.
10. Hadfield,C., Mount,R.C. and Cashmore,A.M. (1995) Protein binding interactions at the *STB* locus of the yeast 2μm plasmid. *NucleicAcids Res.*, **23**, 995–1002.
11. Sengupta,A., Blomqvist,K., Pickett,A., Zhang,Y., Chew,J.S.K. and Dobson,M.J. (2001) Functional domains of yeast plasmid-encoded Rep proteins. *J. Bacteriol.*, **183**, 2306–2315.
12. Cui,H., Ghosh,S.K. and Jayaram,M. (2009) The selfish yeast plasmid uses the nuclear motor Kip1p but not Cin8p for its localization and equal segregation. *J. Cell Biol.*, **185**, 251–264.
13. Wong,M.C., Scott-Drew,S.R., Hayes,M.J., Howard,P.J. and Murray,J.A. (2002) RSC2, encoding a component of the RSC nucleosome remodeling complex, is essential for 2 micron plasmid maintenance in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **22**, 4218–4229.
14. Huang,J., Hsu,J.M. and Laurent,B.C. (2004) The RSC nucleosome-remodeling complex is required for cohesin's association with chromosome arms. *Mol. Cell*, **13**, 739–750.
15. Ma,C.H., Cui,H., Hajra,S., Rowley,P.A., Fekete,C., Sarkeshik,A., Ghosh,S.K., Yates,J.R. 3rd and Jayaram,M. (2013) Temporal sequence and cell cycle cues in the assembly of host factors at the yeast 2 micron plasmid partitioning locus. *Nucleic Acids Res.*, **41**, 2340–2353.
16. Hajra,S., Ghosh,S.K. and Jayaram,M. (2006) The centromere-specific histone variant Cse4p (CENP-A) is essential for functional chromatin architecture at the yeast 2-micron circle partitioning locus and promotes equal plasmid segregation. *J. Cell Biol.*, **174**, 779–790.
17. Mehta,S., Yang,X.M., Chan,C.S., Dobson,M.J., Jayaram,M. and Velmurugan,S. (2002) The 2-micron plasmid purloins the yeast cohesin complex: a mechanism for coupling plasmid partitioning and chromosome segregation. *J. Cell Biol.*, **158**, 625–637.
18. Mehta,S., Yang,X., Jayaram,M. and Velmurugan,S. (2005) A novel role for the mitotic spindle during DNA segregation in yeast: Promoting 2-micron plasmid-cohesin association. *Mol. Cell. Biol.*, **25**, 4283–4298.
19. Velmurugan,S., Yang,X., Chan,C., Dobson,M.J. and Jayaram,M. (2000) Partitioning of the 2μm circle plasmid of *Saccharomyces cerevisiae*: functional coordination with chromosome segregation and plasmid-encoded Rep protein distribution. *J. Cell Biol.*, **149**, 553–566.
20. Liu,Y.T., Ma,C.H. and Jayaram,M. (2013) Co-segregation of yeast plasmid sisters under monopolin-directed mitosis suggests association of plasmid sisters with sister chromatids. *Nucleic Acids Res.*, **41**, 4144–4158.
21. Frappier,L. (2004) In: Funnell,BE and Phillips,GE (eds). *Plasmid Biology*. ASM Press, Washington, DC, pp. 325–340.
22. Prajapati,H.K., Rizvi,S.M., Rathore,I. and Ghosh,S.K. (2017) Microtubule-associated proteins, Bik1 and Bim1, are required for faithful partitioning of the endogenous 2 micron plasmids in budding yeast. *Mol. Microbiol.*, **103**, 1046–1064.
23. Ghosh,S.K., Hajra,S. and Jayaram,M. (2007) Faithful segregation of the multicopy yeast plasmid through cohesin-mediated recognition of sisters. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 13034–13039.
24. Murray,J.A.H. and Cesareni,G. (1986) Functional analysis of the yeast plasmid partition locus *STB*. *EMBO J.*, **5**, 3391–3399.
25. Jayaram,M., Sutton,A. and Broach,J.R. (1985) Properties of *REP3*: a *cis*-acting locus required for stable propagation of the *Saccharomyces cerevisiae* plasmid 2μm circle. *Mol. Cell. Biol.*, **5**, 2466–2475.
26. Rank,G.H., Xiao,W. and Pelcher,L.E. (1994) Transpogenes: the transposition-like integration of short sequence DNA into the yeast 2 micron plasmid creates the *STB* locus and plasmid-size polymorphism. *Gene*, **147**, 55–61.
27. Veit,B.E. and Fangman,W.L. (1985) Chromatin organization of the *Saccharomyces cerevisiae* 2μm plasmid depends on plasmid-encoded products. *Mol. Cell. Biol.*, **5**, 2190–2196.
28. Huang,C.C., Hajra,S., Ghosh,S.K. and Jayaram,M. (2011) Cse4 (CenH3) association with the *Saccharomyces cerevisiae* plasmid partitioning locus in its native and chromosomally integrated states: implications in centromere evolution. *Mol. Cell. Biol.*, **31**, 1030–1040.
29. Veit,B.E. and Fangman,W.L. (1988) Copy number and partition of the *Saccharomyces cerevisiae* 2μm plasmid controlled by transcription regulators. *Mol. Cell. Biol.*, **8**, 4949–4957.
30. Som,T., Armstrong,K.A., Volkert,F.C. and Broach,J.R. (1988) Autoregulation of 2μm circle gene expression provides a model for maintenance of stable plasmid copy levels. *Cell*, **52**, 27–37.
31. Murray,J.A.H., Scarpa,M., Rossi,N. and Cesareni,G. (1987) Antagonistic controls regulate copy number of the yeast 2μm plasmid. *EMBO J.*, **6**, 4205–4212.
32. Reynolds,A.E., Murray,A.W. and Szostak,J.W. (1987) Roles of the 2μm gene products in stable maintenance of the 2μm plasmid of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **7**, 3566–3573.
33. McQuaid,M.E., Pinder,J.B., Arumuggam,N., Lacoste,J.S.C., Chew,J.S.K. and Dobson,M.J. (2017) The yeast 2-micron plasmid Raf protein contributes to plasmid inheritance by stabilizing the Rep1 and Rep2 partitioning proteins. *Nucleic Acids Res.*, **45**, 10518–10533.
34. Rothstein,R. (1991) Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. *Methods Enzymol.*, **194**, 281–301.
35. Tsalik,E.L. and Gartenberg,M.R. (1998) Curing *Saccharomyces cerevisiae* of the 2micron plasmid by targeted DNA damage. *Yeast*, **14**, 847–852.
36. Gietz,R.D., Schiestl,R.H., Willems,A.R. and Woods,R.A. (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast*, **11**, 355–360.

37. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning, A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, NY.
38. Li,J.J. and Herskowitz,I. (1993) Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. *Science*, **262**, 1870–1874.
39. Stinchcomb,D.T., Struhl,K. and Davis,R.W. (1979) Isolation and characterization of a yeast chromosomal replicator. *Nature*, **282**, 39–43.
40. Burke,D., Dawson,D. and Stearns,T. (1997) *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual*. Cold Spring Harbor Laboratory Press, NY.
41. Xiao,W., Pelcher,L.E. and Rank,G.H. (1991) Evidence for *cis*- and *trans*-acting element coevolution of the 2-micron circle genome in *Saccharomyces cerevisiae*. *J. Mol. Evol.*, **32**, 145–152.
42. Xiao,W. and Rank,G.H. (1993) Polymorphism of 2-micron plasmids in industrial strains of *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol*, **39**, 80–86.
43. Strope,P.K., Kozmin,S.G., Skelly,D.A., Magwene,P.M., Dietrich,F.S. and McCusker,J.H. (2015) 2μ plasmid in *Saccharomyces* species and in *Saccharomyces cerevisiae*. *FEMS Yeast Res.*, **15**, fov090.
44. Rank,G.H., Xiao,W. and Arndt,G.M. (1994) Evidence for Darwinian selection of the 2μm plasmid *STB* locus in *Saccharomyces cerevisiae*. *Genome*, **37**, 12–18.
45. Furuyama,S. and Biggins,S. (2007) Centromere identity is specified by a single centromeric nucleosome in budding yeast. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 14706–14711.
46. Cairns,B.R., Schlichter,A., Erdjument-Bromage,H., Tempst,P., Kornberg,R.D. and Winston,F. (1999) Two functionally distinct forms of the RSC nucleosome-remodeling complex, containing essential AT hook, BAH, and bromodomains. *Mol. Cell*, **4**, 715–723.
47. Cairns,B.R., Lorch,Y., Li,Y., Zhang,M., Lacomis,L., Erdjument-Bromage,H., Tempst,P., Du,J., Laurent,B. and Kornberg,R.D. (1996) RSC, an essential, abundant chromatin-remodeling complex. *Cell*, **87**, 1249–1260.
48. Lorch,Y., Maier-Davies,B. and Kornberg,R.D. (2014) Role of DNA sequence in chromatin remodeling and the formation of nucleosome-free regions. *Genes Dev.*, **28**, 2492–2497.
49. Kubik,S., Bruzzone,M.J. and Shore,D. (2017) Establishing nucleosome architecture and stability at promoters: Roles of pioneer transcription factors and the RSC chromatin remodeler. *Bioessays*, **39**, 1600237.
50. Parnell,T.J., Huff,J.T. and Cairns,B.R. (2008) RSC regulates nucleosome positioning at Pol II genes and density at Pol III genes. *EMBO J*, **27**, 100–110.
51. Floer,M., Wang,X., Prabhu,V., Berrozpe,G., Narayan,S., Spagna,D., Alvarez,D., Kendall,J., Krasnitz,A., Stepansky,A. *et al.* (2010) A RSC/nucleosome complex determines chromatin architecture and facilitates activator binding. *Cell*, **141**, 407–418.
52. Kubik,S., Bruzzone,M.J., Jacquet,P., Falcone,J.L., Rougemont,J. and Shore,D. (2015) Nucleosome stability distinguishes two different promoter types at all protein-coding genes in yeast. *Mol. Cell*, **60**, 422–434.
53. Scott-Drew,S. and Murray,J.A.H. (1998) Localization and interaction of the protein components of the yeast 2μm circle plasmid partitioning system suggest a mechanism for plasmid inheritance. *J. Cell Sci.*, **111**, 1779–1789.
54. Ahn,Y.-T., Wu,X.-L., Biswal,S., Velmurugan,S., Volkert,F.C. and Jayaram,M. (1997) The 2μm-encoded Rep1 and Rep2 proteins interact with each other and colocalize to the *Saccharomyces cerevisiae* nucleus. *J. Bacteriol.*, **179**, 7497–7506.
55. Velmurugan,S., Ahn,Y., Yang,X., Wu,X. and Jayaram,M. (1998) The 2μm plasmid stability system: analyses of the interactions among plasmid and host encoded components. *Mol. Cell. Biol.*, **18**, 7466–7477.
56. Funnell,B.E. and Slavcev,R.A. (2004) In: Funnell,BE and Phillips,GJ (eds). *Plasmid Biology*. ASM Press, Washington DC, pp. 81–104.
57. Radman-Livaja,M. and Rando,O.J. (2010) Nucleosome positioning: how is it established, and why does it matter? *Dev. Biol.*, **339**, 258–266.