Hotspot of *de novo* telomere addition stabilizes linear amplicons in yeast grown in sulfate-limiting conditions

Remington E. Hoerr[1], Alex Eng[2], Celia Payen[2,3], Sara C. DiRienzi[2,4], M. K. Raghuraman[2], Maitreya J. Dunham[2], Bonita J. Brewer[2], Katherine L. Friedman[1]

[1] Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA, Email: remington.hoerr@vanderbilt.edu; katherine.friedman@vanderbilt.edu

[2] Department of Genome Sciences, University of Washington, Seattle, WA, USA, Email: engal@uw.edu; raghu@uw.edu; bbrewer@uw.edu; maitreya@uw.edu;

[3] Current affiliation: IFF, Wilmington, Delaware, USA, Email: celia.e.payen@iff.com

[4] Current affiliation: Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston TX, USA, Email: sara.dirienzi@bcm.edu

Keywords: Copy number variation; gene amplification; evolutionary genomics; *de novo* telomere addition; Cdc13; DNA repair; *Saccharomyces cerevisiae*

Abstract:
Evolution is driven by the accumulation of competing mutations that influence survival. A broad form of genetic variation is the amplification or deletion of DNA (≥50 bp) referred to as copy number variation. In humans, copy number variation may be inconsequential, contribute to minor phenotypic differences, or cause conditions such as birth defects, neurodevelopmental disorders, and cancers. To identify mechanisms that drive copy number variation, we monitored the experimental evolution of *Saccharomyces cerevisiae* populations grown under sulfate-limiting conditions. Cells with increased copy number of the gene *SUL1*, which encodes a primary sulfate transporter, exhibit a fitness advantage. Previously, we reported interstitial inverted triplications of *SUL1* as the dominant rearrangement in a haploid population. Here, in a diploid population, we find instead that small linear fragments containing *SUL1* form and are sustained over several generations. Many of the linear fragments are stabilized by *de novo* telomere addition within a telomere-like sequence near *SUL1* (within the *SNF5* gene). Using an assay that monitors telomerase action following an induced chromosome break, we show that this region acts as a hotspot of *de novo* telomere addition and that required sequences map to a region of <250 base pairs. Consistent with previous work showing that association of the telomere-binding protein Cdc13 with internal sequences stimulates telomerase recruitment, mutation of a four-nucleotide motif predicted to associate with Cdc13 abolishes *de novo* telomere addition. Our study suggests that internal telomere-like sequences that stimulate *de novo* telomere addition can contribute to adaptation by promoting genomic plasticity.

35 **Introduction:**

36 Genomic variation is the basis of evolution. Mutations, including both single-nucleotide
37 polymorphisms and larger genomic rearrangements, can rise in prevalence within a population if they
38 confer a fitness advantage. Copy number variation (CNV) refers to a genomic rearrangement that
39 results in loss or gain of DNA (Alkan et al. 2011; MacDonald et al. 2014). CNV is most likely to have a
40 phenotypic impact when the DNA that is duplicated or lost contains one or more genes that impact
41 cellular fitness. In humans, CNV is associated with a variety of disorders primarily (but not exclusively)
42 related to intellectual disabilities, developmental delays, and predisposition to cancers (reviewed in
43 Zhang et al. 2009; Shaikh 2017; Doran and Pennington 2022; R. H. Song et al. 2022).

44 CNV can arise through a variety of chromosomal rearrangements stemming from errors in
45 replication fork progression, recombination, or in DNA damage repair pathways (Hastings et al. 2009).
46 The frequency of CNVs is often increased in cancer cells due to dysregulation of DNA repair or
47 replication pathways (Doran and Pennington 2022). Gene amplification can present in a variety of
48 configurations including both direct and inverted tandem duplication (additional copies at the original
49 location), dispersed duplication (integration of additional copies elsewhere in the genome), or double
50 minute chromosomes that are small circular extrachromosomal fragments of DNA that lack a
51 centromere (Biedler et al. 1983; Chakraborty and Ay 2018; Doran and Pennington 2022).

52 The origin of CNVs is not well understood, in part because the initiating event may be replaced
53 over time by more beneficial secondary and tertiary events that confer increased growth advantages
54 (Ilić et al. 2022). For example, double minute chromosomes may in some cases serve as
55 recombinational intermediates in the formation of tandem or dispersed duplications (Turner et al. 2017;
56 Rosswog et al. 2021; K. Song et al. 2022). Since most studies of CNV in humans rely on samples from a
57 single point in time, they provide only a snapshot of the most dominant rearrangement (Chakraborty

58 and Ay 2018). This limitation highlights the need for experimental systems in which the temporal
59 relationship between events can be observed.

60 The budding yeast *Saccharomyces cerevisiae* utilizes DNA repair pathways similar to those of
61 human cells but is more easily manipulated and monitored for genomic rearrangements (Krogh and
62 Symington 2004; Wellinger and Zakian 2012). For example, growth under specific environmental
63 conditions selects for amplification of one of several genes (e.g., *HXT6/7*, *CUP1, GAP1,* or *SUL1*) (Brown
64 et al. 1998; Gresham et al. 2008; Kao and Sherlock 2008; Gresham et al. 2010; Zhang et al. 2013).
65 Growth of continuous yeast cultures in a chemostat allows the genetic composition of the population to
66 be monitored over time, providing an attractive system to monitor CNV formation and persistence in a
67 heterogenous population. Previously reported evolution experiments with *S. cerevisiae* describe several
68 different types of chromosomal rearrangements that increase gene copy number. Extrachromosomal
69 circular DNAs, similar to the double minute chromosomes described above, have been reported
70 (Gresham et al. 2010). Additionally, persistent linear fragments have also been described (Dorsey et al.
71 1992; Narayanan et al. 2006). Circular molecules lack DNA ends, but linear fragments must protect their
72 ends from nucleolytic degradation. In most cases, protection of the ends is afforded by the telomere, a
73 repetitive, protein-bound DNA sequence that resists nuclease action and engagement with the DNA
74 repair machinery (Blackburn 1991). Where such linear fragments have been described, they arise from
75 palindromic duplication, resulting in an isochromosomal fragment capped at both ends by copies of the
76 same pre-existing telomere (Dorsey et al. 1992; Narayanan et al. 2006).

77 In haploid laboratory strains grown in sulfate-limiting chemostats, the fitter variants almost
78 always contain an inverted interstitial triplication of the primary sulfate transporter gene (*SUL1*) and
79 variable amounts of the flanking regions in an otherwise unrearranged chromosome (Araya et al. 2010;
80 Payen et al. 2014; Brewer et al. 2015). Here, we report that diploids of the same laboratory strain
81 appear to follow a different path to *SUL1* amplification. Instead of chromosomally stabilized amplicons,

82 we find extrachromosomal linear fragments that contain *SUL1* and the nearby endogenous telomere at
83 one end and a new telomere added to a sequence centromere proximal to *SUL1* at the other end.
84 Similar fragments arose independently multiple times, suggesting that one or more sequences
85 centromere proximal to *SUL1* have a propensity to stimulate new telomere formation. Indeed, we show
86 that a sequence within the *SNF5* gene is sufficient to stimulate *de novo* telomere addition when moved
87 to an ectopic location. The ability of this sequence to stimulate *de novo* telomere addition requires at
88 least two short sequence motifs that are predicted to associate with the telomere binding protein Cdc13
89 (Eldridge et al. 2006) and can be functionally replaced by artificial recruitment of Cdc13.
90
91 **Methods and Materials:**
92 Strains and Plasmids
93 All strains for SiRTA analyses were derived from S288c and are listed in File S1. Plasmid pAB180
94 (pRS414-ADHpromoter-GBD-*CDC13*) and pAB220 (pRS414-ADHpromoter-GBD) were gifts of A. Bianchi.
95 Sulfate-limited chemostat experiments were performed on strains of the S288c background (FY3 and
96 FY5) (Gresham et al. 2008; Miller et al. 2013). Creation of the *ars228* deletion has been described
97 (Brewer et al. 2015). Strains for the HO cleavage assay were generated as described in Ngo et al. 2020.
98 In brief, the HO endonuclease recognition sequence was integrated on chromosome VII by one-step
99 gene replacement using plasmid pHO invC (based on JH2017; gift of J. Haber). The *URA3* gene was
100 integrated by one-step gene replacement using pRS306 (Sikorski and Hieter 1989) as template. Putative
101 SiRTAs were inserted on chromosome VII using CRISPR/Cas9 as described (Anand et al. 2017) using gRNA
102 sequence TGCGGCAAGTTCATCTTCCA. Sequences of different length within *SNF5* were amplified from
103 genomic yeast DNA with appropriate primers (File S1) and additional sequences were added that lie
104 upstream (5'-TTTCTTTGGAAAACGTTGAAAATGAGGTTCTATGATCTAC) and downstream (5'-

105 AGAACATAGAATAAATTTGGTACTGGAACGTTGATTAACT) of the gRNA recognition site. *RAD52* was
106 replaced by one-step gene replacement using pFA6a-TRP1 (Longtine et al. 1998) as template.
107
108 Contour-clamped homogeneous electric field (CHEF) gels
109 Agarose plugs for CHEF gel electrophoresis were generated using the method by Lucas Argueso (Kwan et
110 al. 2016). For separation of whole chromosomes, run conditions in the BioRad CHEF-DRII were 1% LE
111 agarose in 0.5XTBE in 2.3 L 0.5% TBE running buffer at 14˚C. Switch times were either 47" to 170" at
112 165V for 39-62 hours or 300" to 900" at 100V for 68 hr. Linear fragments were separated from
113 chromosomes by CHEF gels with switch times 1.7" to 2.6" at 198 V for 20 hr. Restriction analysis of
114 linear fragments were performed by conventional agarose gel electrophoresis in 0.4% ME agarose gels
115 in 1XTBE at 1V/cm for 19 hr at room temperature. Standard protocols for restriction digestion,
116 Southern blotting and hybridization using ^{32}P-labeled PCR probes were used. Linear fragments were
117 concentrated from plug supernatants using Centricon Filter devices and then ethanol precipitated.
118
119 Sequence analysis of linear fragments
120 Whole genome sequencing of fragments purified from plug supernatants and split read analysis were as
121 described (Payen et al. 2014).
122
123 HO cleavage assay
124 The HO cleavage assay was performed as described in Ngo et al. 2020. In brief, cells were grown in SD-
125 Ura medium (+2% raffinose) overnight to an optical density at 600 nm (OD600) of 0.6-0.9. Cells were
126 plated on yeast extract peptone medium with 2% galactose (YEPGal). Simultaneously, cells were serially
127 diluted in water and plated on yeast extract peptone medium with 2% dextrose (YEPD) to determine
128 total cell number. When plasmid selection was required (pRS414, pAB220, and pAB180), cells were

129 plated on synthetic medium lacking tryptophan and containing either 2% glucose or 2% galactose. After
130 3 days, colonies were counted and at least 100 individual colonies were transferred to plates containing
131 5-fluoroorotic acid (5-FOA) to determine the frequency of *URA3* loss. For each experiment, at least 30 5-
132 FOA resistant clones were grown in YEPD for DNA extraction by MasterPure™ Yeast DNA Purification Kit
133 (Lucigen). The relative location of each chromosome truncation event was determined using multiplex
134 PCR with primers that anneal centromere or telomere proximal to the SiRTA (File S1). *De novo* telomere
135 addition events were amplified using CHECKCHR7GUIDE1F and TelomereRev as primers. The purified
136 PCR product was sequenced using Sanger sequencing to determine the site of *de novo* telomere
137 addition. Full data sets corresponding to Figures 3 and 4 are found in File S2.
138
139 **Results:**
140 **Linear fragments containing *SUL1* accumulate in cells grown in sulfate-limiting conditions**
141 Under sulfate-limiting conditions, *S. cerevisiae* cells with increased capacity for sulfate transport
142 exhibit a selective growth advantage (Gresham et al. 2008). The most common genomic rearrangement
143 observed after prolonged competitive growth of haploid lab strains in low-sulfate conditions is an
144 interstitial triplication of a chromosomal region containing the primary sulfate-transporter gene *SUL1*
145 and its adjacent origin of replication *ARS228*, located ~25 kb from the right telomere of chromosome II
146 (Payen et al. 2014; Brewer et al. 2015). From the structure of these amplicons—the central copy of the
147 triplication is in inverted orientation—Brewer et al. (2011) proposed a model for their generation that
148 relies on the presence of the adjacent origin of replication and novel junctions that occur at pre-existing
149 flanking short interrupted inverted repeats (Origin Dependent Inverted Repeat Amplification; ODIRA).
150 That the process is a byproduct of replication was inferred from the observation that deletion of *ARS228*
151 results in larger amplicons that encompass the next nearest replication origin (Brewer et al. 2015).

152 A key feature of the ODIRA model is the generation of an extrachromosomal intermediate that
153 later recombines with the chromosomal locus (Brewer et al. 2011). In an attempt to gather evidence for
154 the unique extrachromosomal intermediate, we constructed a diploid heterozygous for the *ARS228*
155 deletion (*ARS228*/*ars228Δ*; Brewer et al. 2015) and two diploid control strains (*ARS228/ARS228* and
156 *ars228Δ/ars228Δ*) and carried out independent evolution experiments in sulfate-limited chemostats (n =
157 10, 4, and 4, respectively). We reasoned that if the intermediate arose from the *ARS228* chromosome
158 and subsequently recombined with the *ars228Δ* chromosome, the resulting amplicon would contain
159 copies of both *ARS228 and ars228Δ*.
160 Contrary to expectations, these diploids did not produce the same type of inverted interstitial
161 amplicons of *SUL1*. Array comparative genome hybridization (aCGH) of each of the diploid populations
162 after ~180 generations (33-36 days) under sulfate-limitation revealed consistent amplification of the
163 *SUL1* region (e.g., Figure 1A, Figure S1). However, in each case the amplicon was not interstitial but
164 continued through the right telomere. The amplified sequences were limited to this region of
165 chromosome II and the new junctions were narrowly restricted to a small region of chromosome II near
166 the *SNF5* gene.
167 To determine the genomic location of the amplicons, we analyzed samples taken every few days
168 through the 33 days of chemostat growth and analyzed the yeast karyotype by contour-clamped
169 homogeneous electric field (CHEF) gels. Southern blotting and hybridization with probes specific to
170 *SUL1* detected chromosome II as expected, but also revealed the generation of small (~30 kilobases)
171 molecules in all samples from all three genotypes (e.g., Figure 1B, C; Figures S1 and S2). Because they
172 were generated in diploid cells with both copies of *ARS228* deleted (Figure S1C) we propose that the
173 adjacent origin of replication *ARS229* can support the creation and maintenance of these fragments.
174 The sizes of these molecules estimated from their migration on CHEF gels were consistent with the sizes
175 of the amplicon detected by aCGH suggesting that they were acentric linear fragments.

176 The linear fragments usually appeared early in the course of the selection (~by 50 generations)
177 and often persisted through the end of the experiment. In half of the cases, chromosomally integrated
178 *SUL1* amplicons were the predominant form by the end of the chemostat run. In all but one of these
179 cases, the transient linear fragments disappeared concurrently with the emergence of new, slower-
180 migrating *SUL1*-hybridizing bands (Figures S1 and S2). The size of the major amplicon from the final day
181 in the chemostat culture matched the size of the transient linear fragments [in all but one of the
182 cultures (S702)], suggesting a precursor-product relationship between the linear fragments and the
183 chromosomally integrated form. One of the most intriguing aspects of these free linear fragments is
184 their consistent small size—ranging between ~26 and 42 kb (Figure 1C). Since all of the linear fragments
185 extend through the right telomere of chromosome II, the newly formed left ends of the linear fragments
186 appear to arise near or within *SNF5*, a gene centromere proximal to *SUL1*.
187
188 **Linear fragments are stabilized by addition of a new telomere**
189 While selective pressure can partially compensate for the lack of a centromere, stability of a
190 linear fragment requires protection at both ends from exonucleases. Normal chromosome ends
191 (telomeres) in *S. cerevisiae* consist of a thymine and guanine (TG)-rich strand that extends 10-15 bases
192 past the complementary adenine and cytosine (AC)-rich strand to create a 3' overhang (Blackburn 1991).
193 This TG-rich single-stranded overhang interacts with the enzyme telomerase, which utilizes its reverse
194 transcriptase catalytic subunit and intrinsic AC-rich RNA template to add telomeric repeats to the
195 chromosome end (Singer and Gottschling 1994; Lingner et al. 1997). In yeast, the 3' terminating strand
196 consists of a heterogeneous sequence pattern with a single thymine followed by one to three guanines
197 (abbreviated TG_{1-3}) (Wellinger and Zakian 2012).
198 The aCGH analyses (Figures 1A and S1) suggested that the right end of the linear fragment
199 contained the endogenous right telomere, but the structure of the left end was initially unclear. More

200 detailed analysis of the linear fragments was facilitated by the development of a method to purify these
201 small molecules through passive diffusion from agarose CHEF gel plugs. Over the course of several
202 weeks, the small fragments diffuse out of the agarose plug and into the supernatant where they can be
203 recovered, while the intact chromosomes remain trapped in the plug (Figure S3). Purified linear
204 fragments were digested with *Eco*NI, which cleaves three times in the terminal region of chromosome II
205 (Figure 2A). Following separation by CHEF gel, the samples were Southern blotted and probed
206 sequentially with sequences that hybridize to either *SUL1* or telomeric DNA (Figure 2B). *Eco*NI digestion
207 separated *SUL1* from the right telomere and allowed us to query both ends of the linear fragments. The
208 restriction fragments containing the right telomere hybridized to telomeric DNA ($C_{1-3}A/G_{1-3}T$) and not to
209 *SUL1* (Figure 2B, gold rectangles). However, fragments of multiple sizes (ranging from ~3-25 kb)
210 hybridized to both the *SUL1* and the telomeric probe (Figure 2B), suggesting that the left end of each
211 fragment was also capped by telomeric DNA.
212 To determine the sites of telomeric repeat addition, we determined the sequence of fragments
213 purified from the plug supernatants by Illumina sequencing. The results revealed that nine of the 15
214 fragments contained telomeric repeats clustered within a ~100 bp region of the gene *SNF5*,
215 approximately ~7 kb from *SUL1* (Figure S4). With one exception, there was no evidence of subtelomeric
216 sequences, but rather the telomeric DNA was appended directly to the chromosome II sequence,
217 consistent with addition of a new telomere to internal sequences by telomerase.
218
219 **Sequences in the *SNF5* gene stimulate unusually high levels of *de novo* telomere addition**
220 The observation that many of the telomere addition events associated with extrachromosomal
221 *SUL1* fragments cluster within a small (~100 bp) region is reminiscent of sequences in the yeast genome
222 that are hotspots of *de novo* telomere addition (termed SiRTAs or Sites of Repair-associated Telomere
223 Addition) (Stellwagen et al. 2003; Obodo et al. 2016; Hoerr et al. 2021). SiRTAs contain strand-specific,

224 TG-rich sequence motifs (Obodo et al. 2016; Ngo et al. 2020). These motifs can provide a substrate for
225 the acquisition of telomeric sequences if they become exposed in single-stranded DNA either as a result
226 of 5' end resection following a DSB or at nicks in single-stranded gaps. In either case, the resulting TG-
227 rich 3' overhang can invade a similar sequence and acquire a telomere through break-induced
228 replication or can undergo *de novo* telomere addition by direct action of telomerase (Putnam and
229 Kolodner 2017). Such events are most easily assayed in cases where the new telomere is added to the
230 centromere-proximal side of the break, causing a terminal truncation. In contrast, the putative SiRTA
231 within *SNF5* is oriented in the opposite orientation on the right arm of the chromosome. Rather than
232 stabilizing the centromere-containing region of the chromosome, *de novo* telomere addition stabilizes a
233 terminal acentric fragment containing *SUL1*.

234 Like telomeres, SiRTAs have an asymmetric sequence pattern in which one strand is relatively
235 rich in T and G nucleotides (Obodo et al. 2016; Ngo et al. 2020). For telomerase to act upon a SiRTA, this
236 TG-rich strand must be exposed in single-stranded DNA. SiRTAs have a bipartite structure consisting of
237 one TG-rich region that is most often the site of telomere addition (the "Core") and a second TG-rich
238 "Stim" sequence located 5' to the Core on the TG-rich strand (Obodo et al. 2016). The Stim sequence
239 itself is rarely targeted for telomere addition but is required to stimulate *de novo* telomere addition in
240 the Core (Obodo et al. 2016). Based on this known structure, we speculated that the ~100 bp region of
241 *SNF5* most frequently observed to undergo telomere addition during generation of the linear fragments
242 represents the Core sequence and that an adjacent sequence would provide Stim function. For a
243 telomere to be added to the left end of the acentric fragment of chromosome II, the TG-rich sequence
244 must be on the bottom or 3' to 5' strand (Fig. 3A). Analysis of this strand showed that a 718 nucleotide
245 (nt) region including and extending 5' to the putative Core sequence (chromosome II coordinates
246 779784 to 780501) is unusually TG-rich. Sequences of 300 nt from the 3' to 5' strand generated as 10
247 base sliding windows extending across this region contain, on average, 66% T or G nucleotides

248 [(T+G)/300 = 0.66 ±0.05 n= 42; Figure S5A). In contrast, 300 nt sequences randomly selected from the
249 yeast genome (n=493) contain, on average, 50% T or G nucleotides [(T+G)/300 = 0.50±0.04], as expected
250 from an unbiased distribution relative to each other and the two strands (Figure S5A). To ensure that
251 our analysis included any potential Stim sequence, we initially tested this entire 718 nt sequence for its
252 ability to stimulate *de novo* telomere addition.

253 To determine the ability of this 718 nt sequence to stimulate telomere addition, we took
254 advantage of an assay in which the sequence of interest is integrated on the left arm of chromosome VII,
255 replacing sequences between chromosome VII coordinates 18086 and 18108 (Ngo et al. 2020). This
256 strain (Lydeard et al. 2010; Obodo et al. 2016) contains a single recognition site for the yeast
257 homothallic switching (HO) endonuclease that allows regulated generation of a double-strand break
258 approximately 2 kb distal to the sequence being tested (Figure 3A). Of note, the *SNF5* sequence is
259 inserted on chromosome VII in the inverted orientation relative to the centromere compared with its
260 orientation on chromosome II. However, because the endogenous location is on the right arm of
261 chromosome II and the insertion site is on the left arm of chromosome VII, the TG-rich sequence that
262 will stimulate *de novo* telomere addition is, in both cases, located on the bottom (3' to 5') strand. In
263 both configurations, exposure of the TG-rich strand in single-stranded DNA (for example, by resection of
264 a DSB occurring to the left of the sequence of interest) followed by telomerase action on that 3'
265 overhang will stabilize the chromosome fragment located to the right of the lesion. On chromosome II,
266 *de novo* telomere addition will create an acentric terminal fragment. On chromosome VII, *de novo*
267 telomere addition will truncate the chromosome. The site at which the putative SiRTA is integrated is
268 distal to the last essential gene on the chromosome arm (*BRR6*), ensuring that cells adding a *de novo*
269 telomere within the sequence of interest remain viable. Because the sequence being tested is present
270 both at its endogenous location and at the integrated site on chromosome VII, we deleted *RAD52* to
271 prevent repair by ectopic recombination.

272 Growth of this strain on media containing galactose induces expression of the HO endonuclease,
273 generating a double-strand break (DSB) distal to the putative SiRTA (Lydeard et al. 2010; Obodo et al.
274 2016). Correct repair restores the cleavage site, leading to a cycle of cleavage and repair that culminates
275 in cell death or mutation/loss of the HO site. Among the cells that survive HO cleavage, those that have
276 lost a *URA3* marker distal to the HO site are identified through selection for growth on media containing
277 5-fluoroorotic acid (5-FOA) and are subsequently screened through a combination of PCR and Sanger
278 sequencing to identify telomere addition events (Obodo et al. 2016). The fraction of 30 analyzed clones
279 that have undergone *de novo* telomere addition within the sequence of interest represents the
280 "relative" frequency of *de novo* telomere addition and is used as a measure of SiRTA function (Obodo et
281 al. 2016). Previously described SiRTAs display relative frequencies of *de novo* telomere addition ranging
282 from 10-35% (Obodo et al. 2016; Ngo et al. 2020).
283 In this case, analysis of 5-FOA resistant clones surviving HO cleavage from four independent
284 assays (30 clones each) showed that 30.8% ±5% had incurred truncation events mapping to the 718 bp
285 sequence (Figure 3B). Each of the 51 5-FOA resistant clones was analyzed by PCR using one primer that
286 anneals to the AC-rich telomeric strand and a second primer that anneals centromere-proximal to the
287 718 bp sequence (Mangahas et al. 2001). In every case, a product of the expected size was produced
288 which, when sequenced, revealed the direct addition of telomeric DNA to the chromosome II sequence
289 (File S2). Taken together, these results indicate that this region functions as a SiRTA (subsequently
290 referred to as the "*SNF5*" SiRTA due to its location within the *SNF5* open reading frame). 50 of the 51
291 telomere addition events mapped to ~180 bases at the 3' end of the 718 base region, relative to the TG-
292 rich strand (Figure 3 and Figure S6). This ~180 base region includes the ~100 base region in which
293 telomere addition was observed in the linear fragments and 69% of the HO-induced events map within 3
294 bases of a telomere-addition event observed on the linear molecules. However, six of nine telomere
295 additions in the linear fragments lie 3' to the vast majority of telomere additions observed after HO

296 induction (Figure S6). It is unclear why events observed in the linear fragments from the sulfate limited
297 chemostats are distributed differently than in the HO assay.
298
299 ***De novo* telomere addition requires specific motifs associated with binding of Cdc13**
300 To refine the location of sequences required for *de novo* telomere addition, we tested a series of
301 fragments that included the region within which telomere addition occurred but were progressively
302 truncated at the opposite (5') end of the TG-rich strand. Fragments of 456 and 226 bp stimulated *de*
303 *novo* telomere addition as strongly as the full 718 bp sequence (Figure 3B). In contrast, no telomere
304 addition events were observed in a fragment of 175 bp, despite retention of virtually all sequences that
305 were targeted by telomerase in the context of the larger fragment (Figure 3A, B). We conclude that the
306 *SNF5* SiRTA has a structure similar to other characterized SiRTAs, with stimulatory sequences that map
307 to the ~50 bp region delimited by the 226 and 175 bp fragments. We refer to this 50 bp region as the
308 *SNF5* SiRTA Stim. Although the fraction of T and G nucleotides is elevated throughout the 718 bp region
309 (Figure S5B, left axis), the ratio of G to T is notably higher in the centromere proximal 226 nucleotide
310 region that is sufficient for *de novo* telomere addition than in the remainder of the region (Figure S5B,
311 right axis). Because telomeres conform to a pattern of TG_{1-3}, the higher G to T ratio is consistent with a
312 sequence that is more "telomere-like."
313 In *S. cerevisiae*, telomerase is primarily recruited to the telomere by the single-stranded binding
314 protein Cdc13 (Pennock et al. 2001). Binding of Cdc13 to TG-rich Stim sequences rendered single-
315 stranded as a result of 5' end resection following a DSB has been implicated in the high frequency of *de*
316 *novo* telomere addition within a SiRTA (Obodo et al. 2016; Epum et al. 2020). Therefore, we
317 hypothesized that the 50 bp *SNF5* SiRTA Stim functions through recruitment of Cdc13. The optimal
318 Cdc13 binding site within telomeric repeats has been characterized (5'-GTGTGGGTGTG) (Anderson et al.
319 2003; Eldridge et al. 2006), but single nucleotide substitutions within most positions are well tolerated.

320 The largest decrease in Cdc13 binding affinity is observed when G1, G3, or T4 nucleotide positions are
321 substituted (GTGTGGGTGTG), suggesting that a 5'-GxGT-3' motif may be necessary (but not sufficient)
322 for Cdc13 binding (Anderson et al. 2003; Eldridge et al. 2006). The *SNF5* SiRTA Stim sequence defined
323 above contains two 5'-GxGT-3' motifs on the TG-rich (bottom) strand (Figure 4A). To determine if these
324 motifs contribute to the high frequency of *de novo* telomere addition within the SiRTA, we mutated
325 each 5'-GxGT-3' motif to its complementary CxCA in the context of the 718 bp fragment. Remarkably,
326 mutation of one or both motifs nearly eliminated *de novo* telomere addition (Figure 4A, B), suggesting
327 that both are required for SiRTA function. This result is consistent with previous reports that SiRTA
328 activity likely requires more than one Cdc13 binding site (Strecker et al. 2017). We hypothesize that the
329 SiRTA sequence must reach a threshold of Cdc13 binding to stimulate *de novo* telomere addition and
330 this threshold is not met in the absence of one or both 5'-GxGT-3' motifs.

331 To further interrogate the requirement of Cdc13 binding to stimulate *de novo* telomere
332 addition, we replaced the 50 bp region of the *SNF5* SiRTA Stim sequence with two copies of the Gal4
333 upstream activating sequence (Stim::2xUAS) to allow specific recruitment of the Gal4 DNA binding
334 domain (GBD) (Zhang and Durocher 2010; Obodo et al. 2016). We introduced an empty vector, a vector
335 expressing GBD alone, or a vector expressing a fusion of GBD to full-length Cdc13 into cells with the
336 *SNF5* SiRTA Stim::2xUAS fragment integrated at the chromosome VII test site. As expected, *SNF5* SiRTA
337 Stim::2xUAS did not support high frequency *de novo* telomere addition with empty vector or GBD alone.
338 Consistent with previous reports that artificial recruitment of Cdc13 to the SiRTA supports *de novo*
339 telomere addition (Zhang and Durocher 2010; Obodo et al. 2016), expression of the GBD-Cdc13 fusion
340 protein supported *de novo* telomere addition in 24.4% ±8.3% of clones tested from three independent
341 assays and stimulation was dependent on the presence of the UAS sequences (Figure 4C). These results
342 further support our hypothesis that a threshold of Cdc13 binding must be met to stimulate *de novo*

343 telomere addition and that without the 5'-GxGT-3' motif-containing Stim sequence, this critical
344 threshold is not satisfied.
345
346 **Discussion:**
347 In this work, we demonstrate that diploid laboratory yeast strains grown in sulfate-limiting
348 conditions generate small, extrachromosomal linear DNA fragments containing the primary sulfate
349 transporter *SUL1* that persist because they provide a selective growth advantage. The linear fragments
350 are capped at one end by sequences from the endogenous right telomere and at the other end by *de*
351 *novo* telomere addition events clustered in a region centromere proximal to *SUL1*. Fragments of similar,
352 but not identical, size arise in multiple independent cultures, suggesting that some sequence feature or
353 features in this region are responsible for the creation and/or stabilization of these linear fragments.
354 These results raise three intertwined questions: 1) What features of the region make it a hotspot for
355 telomere addition? 2) What process generates the substrate for telomere addition in the sulfate
356 limiting chemostats (in the absence of HO cleavage)? And 3) Why do the diploids produce a different
357 form of *SUL1* amplicon than their isogenic haploids?
358 <u>Telomere addition:</u> We find multiple sequence features within the *SNF5* gene that contribute to
359 the stabilization of linear fragments through *de novo* telomere addition. We demonstrate that a DNA
360 fragment from the *SNF5* gene that contains the common sites of telomere addition is sufficient to
361 stimulate *de novo* telomere addition at an ectopic location, similar to other characterized hotspots of *de*
362 *novo* telomere addition (SiRTAs) (Obodo et al. 2016; Ngo et al. 2020). Previous work demonstrated that
363 association of Cdc13 with a SiRTA is critical for the stimulation of telomere addition (Obodo et al. 2016;
364 Epum et al. 2020). However, the number and/or affinity of Cdc13 binding sites sufficient to support *de*
365 *novo* telomere formation are unclear. Strecker et al. (2017) placed differing lengths of telomeric repeats
366 adjacent to a double-strand break to determine a threshold length above which telomere addition is

367 stimulated. Telomere addition is strongly inhibited at sequences below that threshold length (Strecker
368 et al. 2017). Hypomorphic mutations in Cdc13 increase this threshold value, suggesting that the net
369 number and/or affinity of Cdc13 binding sites likely determines the threshold at which a sequence acts
370 to stimulate telomere addition (Strecker et al. 2017). Our results are consistent with this hypothesis.
371 Remarkably, mutation of just four nucleotides corresponding to a predicted binding site for Cdc13
372 (Anderson et al. 2003; Eldridge et al. 2006) nearly eliminates detectable *de novo* telomere addition at
373 the *SNF5* SiRTA in the ectopic assay. Our results suggest that there are likely two Cdc13 binding sites in
374 the ~50 bp fragment that together constitute the *SNF5* SiRTA Stim. Mutation of either site lowers the
375 total amount of Cdc13 binding below the threshold required for this sequence to function as a hotspot
376 of *de novo* telomere addition. Consistent with this interpretation, replacement of the 50 bp *SNF5* SiRTA
377 Stim sequence with two copies of the Gal4 upstream activating sequence supports *de novo* telomere
378 addition only when cells express the Gal4 DNA binding domain fused to Cdc13.
379 Substrate for telomere addition: We showed that *SNF5* SiRTA activity is supported by a minimal
380 fragment of 226 bp. This small region comprises approximately ~1% (226 bp/20.2 kb) of the total region
381 of chromosome VII in which a viable chromosome truncation could occur (between the HO site and the
382 most distal essential gene on this chromosome arm), and yet nearly one third of all truncation events
383 were mapped to the *SNF5* SiRTA. This clustering of *de novo* telomere additions in a small region of *SNF5*
384 could result from chromosome fragility in this region; however, observations from the ectopic assay
385 argue against this model. The *SNF5* SiRTA undergoes unusually high levels of *de novo* telomere addition
386 even when the initiating chromosome break is intentionally induced two kilobases distal to the eventual
387 site at which telomerase acts, suggesting that it is a site at which DNA ends are preferentially repaired.
388 We conclude that this region of *SNF5* is not necessarily more prone to double stranded breaks, but
389 rather has the fortuitous combination of features that permit telomere addition.

390 While we have not measured double stranded breaks in chemostat cultures, we have evidence
391 that replication errors, instead of double stranded breaks, may generate aberrant intermediates that are
392 subsequently targeted for repair (Brewer et al. 2015). The discontinuous nature of DNA replication forks
393 provides an opportunity for nascent leading strands to anneal to the lagging strand template during fork
394 reversal at sites of interrupted inverted repeats (described as ODIRA in Brewer et al. 2011). In a scan of
395 the region at the 5' end of the *SNF5* gene we find a ~5 fold increase in the density of short interrupted
396 inverted repeats (Figure S7A) resulting primarily from a sequence in the promoter region (5'-TGTTGTTG-
397 3') that is complementary to the several stretches of CAA triplet repeats within the N-terminal coding
398 region of *SNF5*. When compared to >10,000 regions across the genome, the *SNF5* region is enriched for
399 short interrupted inverted repeats (Fig S7B). The regression of a replication fork proceeding toward the
400 centromere from *ARS228* or another distal origin (e.g. *ARS229*) could provide multiple opportunities for
401 this sequence on the leading strand to find a match with a complementary sequence in the single
402 stranded gaps on the lagging strand and thereby create a closed loop. After completion of replication by
403 a fork established centromere proximal to the *SNF5* region, the hairpin capped linear could be released
404 (Brewer et al. 2011) and suffer one of several fates. Replication in the next cell cycle would generate an
405 isochromosomal palindromic linear fragment. Alternatively, recognition of the single stranded loop by
406 nucleases and repair enzymes could provide the substrate for telomere addition (Figure S7C,D).
407 <u>Ploidy specific amplification pathways:</u> In previous experiments with haploid laboratory strains,
408 we have not observed the generation of linear fragments—essentially all events were interstitial
409 inverted triplications on chromosome II (Payen et al. 2014). In some cases, rearrangements that would
410 be lethal in haploids are tolerated in diploids by the presence of a second, unrearranged chromosome
411 (Zhang et al. 2013). It is unlikely that a similar phenomenon is affecting the nature of *SUL1* amplification
412 in diploid and haploid cells as there are no essential genes beyond *SNF5* on chromosome II, so loss of
413 this terminal fragment should be tolerable in haploids. In point of fact, we do not detect loss of the

414 distal markers on chromosome II in the diploids. The aCGH profiles indicated that there were no regions
415 at the right end of chromosome II that were present at fewer than two copies per diploid genome and
416 CHEF gels did not reveal any truncated variants of chromosome II (Figures 1, S1, and S2). While these
417 results are consistent with repair of a broken chromosome II by homology directed repair off the
418 homolog or sister chromatid (perhaps by break-induced replication), it is also possible that a double
419 stranded break was not the initiating event. We would like to propose that the same type of ODIRA
420 event that can explain the events that occur in haploid cells is also occurring in the diploids, but that the
421 resolution of the intermediate differs between haploids and diploids. Differences in the regulation of
422 repair pathways between haploid and diploid cells may contribute to the different outcomes (Li and Tye
423 2011). There may be a race between replication of the hairpin intermediate and the processing of the
424 single stranded loop that is influenced by the levels of repair enzymes. In support of this notion, we
425 have recently observed that in some haploid wild strains with varied genetic backgrounds, linear
426 amplicons similar to those described here, also arising from the vicinity of the *SNF5* gene, are the
427 preferred amplicon (manuscript in preparation). Analysis of the transcriptomes and proteomes of these
428 strains could identify co-varying RNA/proteins that influence the amplicon outcome. Regardless of
429 mechanism, our observations in yeast may provide a useful model for the types of rearrangements that
430 occur in humans and other organisms.
431 Finally, results presented here suggest that interstitial TG-rich hotspots of *de novo* telomere
432 addition that are oppositely oriented relative to the chromosome arm can stabilize an acentric fragment
433 to provide, in some circumstances, a selective advantage. These observations highlight a possible
434 evolutionary consequence of SiRTA sequences in *S. cerevisiae*, regardless of orientation, by increasing
435 adaptation and genome plasticity.
436
437

438 **Data Availability**

439 Strains and plasmids are available upon request. Source data for Figures 3, 4, and S6 are provided in File
440 S2. Sequencing data will be available from NIH Sequence Read Archive (SRA). aCGH data will be
441 available from NIH Gene Expression Omnibus (GEO).

442

443 **Acknowledgements**

444 We thank James Haber and Alessandro Bianchi for generously providing us with strains and plasmids
445 and Natalie Guzman for initial exploratory studies.

446

447 **Funding**

448 This work was supported by National Institutes of Health awards R01GM123292 to KLF, R35GM122497
449 to BJB, and R01GM094306 to MJD and by National Science Foundation award 1516330 to MJD. REH was
450 supported by National Institutes of Health award T32GM137793. The funders had no role in study
451 design, data collection and analysis, decision to publish, or preparation of the manuscript.

452

453 **Figure legends:**

454 Figure 1. Amplification of *SUL1* chromosomal fragments in sulfate-limited chemostats. A) Array
455 Comparative Genomic Hybridization (aCGH) of the population of a diploid yeast strain 33 days after
456 continuous growth in a sulfate limited chemostat. The right end of chromosome II, containing the gene
457 for the primary sulfur transporter *SUL1*, is present in 2-3 copies in excess of the genomic copies
458 (normalized to 1). The presence of two discontinuities at 782 kb and 788 kb suggests two discrete
459 amplicons within the cell population. B) CHEF gel electrophoresis of population samples across the 33
460 days of continuous growth in sulfate limited medium. The ethidium bromide stained gel displays the
461 total yeast karyotype. The CHEF gel was Southern blotted and probed sequentially for *SUL1* and *CEN2* as

462 indicated. Two extrachromosomal bands of low molecular weight are detected by the *SUL1* probe from
463 days 17 onward. C) CHEF gel electrophoresis, under conditions to determine the sizes of
464 extrachromosomal amplicons, from specific days of 15 individual chemostat runs. The ethidium
465 bromide stained gel and Southern blot hybridized with *SUL1* reveal multiple small extrachromosomal
466 fragments. The sizes of extrachromosomal amplicons vary from 26 kb to 40 kb with a recurring event
467 producing a fragment of approximately 32 kb (red arrow).
468
469 Figure 2. Linear fragments are capped by new telomeric sequences. A) Positions of the *Eco*NI restriction
470 sites at the right end of chromosome II (coordinates 770 to 813 kb) and of two probes [$C_{(1-3)}A/G_{(1-3)}T$ and
471 *SUL1* (indicated by the horizontal bars)]. Chromosome II-R telomeric sequences are indicated with gold
472 triangles. Proposed telomeric sequences at variable positions at the left margin of the amplicons are
473 indicated by the orange triangles. B) Linear fragments purified from the plug supernatant (see Figure
474 S3) for five cultures (S708-S712) were digested with *Eco*NI and separated on a CHEF gel. Southern blots
475 were sequentially hybridized with a *SUL1* probe and a telomere specific probe ($C_{(1-3)}A/G_{(1-3)}T$). Cultures
476 S710 and S712 contain two species of extrachromosomal molecules. Orange triangles indicate the
477 fragments with *de novo* telomeres. Fragments boxed in gold correspond to the native right telomere.
478
479 Figure 3. The *SNF5* SiRTA is sufficient to stimulate *de novo* telomere addition at an ectopic location. A)
480 Top schematic: the endogenous right arm of chromosome II; distance between the endogenous *SNF5*
481 sequence and *SUL1* is shown. Triangles represent the endogenous telomere. Middle schematic: lines
482 representing the relative size of each fragment from *SNF5* that was integrated and tested (actual size in
483 bp shown). Black arrows above the 718 bp fragment depict approximate locations of *de novo* telomere
484 addition events stimulated by HO cleavage. Blue arrows below the 718 bp fragment depict approximate
485 locations of spontaneous *de novo* telomere addition events from the chemostat experiments. Arrow

486 size is proportional to the number of independent events at that location (see Figure S6 for exact
487 numbers and locations). Bottom schematic: modified left arm of chromosome VII. Distances between
488 the integrated *SNF5* sequence and the HO cleavage site or the most distal essential gene (*BRR6*) on this
489 chromosome arm are shown. B) Chromosome truncation events following HO cleavage preferentially
490 involve *de novo* telomere addition in the SiRTA. At least 30 clones that survived HO cleavage and lost
491 the *URA3* marker were assayed for each experiment. The percentage of clones that incurred telomere
492 addition within the sequence of interest is shown. Bar height is the average of three to four
493 independent experiments; error bars represent standard deviations. Significant difference is indicated
494 (* $p<0.05$) by unpaired Student's *t-test* comparing 226 bp to 175 bp. See File S2 for data used to
495 generate graph.
496
497 Figure 4. The *SNF5* SiRTA requires Cdc13 binding to stimulate *de novo* telomere addition. A) Top: lines
498 representing 718, 175 bp, and *SNF5* SiRTA Stim::2XUAS fragments integrated and tested on
499 chromosome VII. Sites of *de novo* telomere addition are depicted as in Figure 3. Bottom: Twenty-nine
500 nucleotides of the TG-rich strand of the *SNF5* SiRTA Stim sequence are shown from 3' to 5'. Note that a
501 double-strand break that occurs to the left of this sequence will expose the TG-rich sequence in single-
502 stranded DNA following 5' end resection. 5'-GxGT-3' motifs and corresponding mutations are
503 highlighted in gray. B) Percentage of 5-FOA resistant clones that incurred *de novo* telomere addition
504 within the 718 bp *SNF5* SiRTA in the presence of the indicated mutation. Wild-type (WT) samples are
505 repeated from Figure 3 for comparison. Samples with statistically significant values by one-way ANOVA
506 with *post hoc* Dunnett's are indicated (**, $p< 0.01$). C) Percentage of 5-FOA resistant clones that
507 incurred *de novo* telomere addition within the 175 bp *SNF5* fragment when 2 copies of the Gal4 UAS are
508 added (Stim::2xUAS) or when the 175 bp fragment is present alone. Cells were transformed with
509 pRS414 (empty vector) or pRS414 expressing either Gal4 DNA binding domain (GBD) or GBD fused to the

510 N-terminus of Cdc13 as indicated. For B and C, at least 30 colonies were assayed for each experiment.
511 Bar height is the average of three to four independent experiments; error bars represent standard
512 deviations. Samples with statistically significant values by one-way ANOVA with *post hoc* Tukey's are
513 indicated (*, p< 0.01; ***, p<0.001). See File S2 for data used to generate graph.
514

515 **Literature Cited**

516 Alkan C, Coe BP, Eichler EE. 2011. Genome structural variation discovery and genotyping. Nat Rev Genet.
517 12(5):363–376. doi:10.1038/nrg2958.

518 Anand R, Memisoglu G, Haber J. 2017. Cas9-mediated gene editing in *Saccharomyces cerevisiae*. Protoc
519 Exch. doi:10.1038/protex.2017.021a.

520 Anderson EM, Halsey WA, Wuttke DS. 2003. Site-directed mutagenesis reveals the thermodynamic
521 requirements for single-stranded DNA recognition by the telomere-binding protein Cdc13. Biochemistry.
522 42(13):3751–3758. doi:10.1021/bi027047c.

523 Araya CL, Payen C, Dunham MJ, Fields S. 2010. Whole-genome sequencing of a laboratory-evolved yeast
524 strain. BMC Genomics. 11(1):88. doi:10.1186/1471-2164-11-88.

525 Biedler JL, Meyers MB, Spengler BA. 1983. Homogeneously staining regions and double minute
526 chromosomes, prevalent cytogenetic abnormalities of human neuroblastoma cells. Adv Cell Neurobiol.
527 4:267–307. doi:10.1016/B978-0-12-008304-6.50015-4.

528 Blackburn EH. 1991. Structure and function of telomeres. Nature. 350(6319):569–573.
529 doi:10.1038/350569A0.

530 Brewer BJ, Payen C, DiRienzi SC, Higgins MM, Ong G, Dunham MJ, Raghuraman MK. 2015. Origin-
531 dependent inverted-repeat amplification: tests of a model for inverted DNA amplification. PLoS Genet.
532 11(12):e1005699. doi:10.1371/journal.pgen.1005699.

533 Brewer BJ, Payen C, Raghuraman MK, Dunham MJ. 2011. Origin-dependent inverted-repeat
534 amplification: A replication-based model for generating palindromic amplicons. PLoS Genet.
535 7(3):e1002016. doi:10.1371/journal.pgen.1002016.

536 Brown CJ, Todd KM, Rosenzweig RF. 1998. Multiple duplications of yeast hexose transport genes in
537 response to selection in a glucose-limited environment. Mol Biol Evol. 15(8):931–942.
538 doi:10.1093/OXFORDJOURNALS.MOLBEV.A026009.

539 Chakraborty A, Ay F. 2018. Identification of copy number variations and translocations in cancer cells
540 from Hi-C data. Bioinformatics. 34(2):338–345. doi:10.1093/BIOINFORMATICS/BTX664.

541 Doran CG, Pennington SR. 2022. Copy number alteration signatures as biomarkers in cancer: a review.
542 Biomark Med. 16(5):371–386. doi:10.2217/BMM-2021-0476.

543 Dorsey M, Peterson C, Bray K, Paquin CE. 1992. Spontaneous Amplification of the ADH4 gene in
544 *Saccharomyces cerevisiae*. Genetics. 132(4):943–50. doi:10.1093/genetics/132.4.943.

545 Eldridge AM, Halsey WA, Wuttke DS. 2006. Identification of the determinants for the specific recognition
546 of single-strand telomeric DNA by Cdc13. Biochemistry. 45(3):871–879. doi:10.1021/bi0512703.

547 Epum EA, Mohan MJ, Ruppe NP, Friedman KL. 2020. Interaction of yeast Rad51 and Rad52 relieves
548 Rad52-mediated inhibition of *de novo* telomere addition. PLoS Genet. 16(2):e1008608.
549 doi:10.1371/journal.pgen.1008608.

550 Gresham D, Desai MM, Tucker CM, Jenq HT, Pai DA, Ward A, DeSevo CG, Botstein D, Dunham MJ. 2008.
551 The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in
552 yeast. PLoS Genet. 4(12). doi:10.1371/JOURNAL.PGEN.1000303.

553 Gresham D, Usaite R, Germann SM, Lisby M, Botstein D, Regenberg B. 2010. Adaptation to diverse
554 nitrogen-limited environments by deletion or extrachromosomal element formation of the *GAP1* locus.

555 Proc Natl Acad Sci U S A. 107(43):18551. doi:10.1073/PNAS.1014023107.

556 Hastings PJ, Lupski JR, Rosenberg SM, Ira G. 2009. Mechanisms of change in gene copy number. Nat Rev
557 Genet. 10(8):551. doi:10.1038/NRG2593.

558 Hoerr RE, Ngo K, Friedman KL. 2021. When the ends justify the means: regulation of telomere addition
559 at double-strand breaks in yeast. Front Cell Dev Biol. 9(18):655377. doi:10.3389/fcell.2021.655377.

560 Ilić M, Zaalberg IC, Raaijmakers JA, Medema RH. 2022. Life of double minutes: generation, maintenance,
561 and elimination. Chromosoma. epub ahead. doi:10.1007/S00412-022-00773-4.

562 Kao KC, Sherlock G. 2008. Molecular characterization of clonal interference during adaptive evolution in
563 asexual populations of *Saccharomyces cerevisiae*. Nat Genet. 40(12):1499–1504. doi:10.1038/NG.280.

564 Krogh BO, Symington LS. 2004. Recombination proteins in yeast. Annu Rev Genet. 38(1):233–271.
565 doi:10.1146/annurev.genet.38.072902.091500.

566 Kwan EX, Wang XS, Amemiya HM, Brewer BJ, Raghuraman MK. 2016. rDNA copy number variants are
567 frequent passenger mutations in *Saccharomyces cerevisiae* deletion collections and *de novo*
568 transformants. G3 Genes, Genomes, Genet. 6(9):2829–2838. doi:10.1534/g3.116.030296.

569 Li XC, Tye BK. 2011. Ploidy dictates repair pathway choice under DNA replication stress. Genetics.
570 187(4):1031. doi:10.1534/GENETICS.110.125450.

571 Lingner J, Cech TR, Hughes TR, Lundblad V. 1997. Three ever shorter telomere (*EST*) genes are
572 dispensable for *in vitro* yeast telomerase activity. Proc Natl Acad Sci U S A. 94(21):11190–11195.
573 doi:10.1073/pnas.94.21.11190.

574 Longtine MS, McKenzie A, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. 1998.
575 Additional modules for versatile and economical PCR-based gene deletion and modification in
576 *Saccharomyces cerevisiae*. Yeast. 14(10):953–961. doi:10.1002/(SICI)1097-0061(199807)14:10<953::AID-

577 YEA293>3.0.CO;2-U.

578 Lydeard JR, Lipkin-Moore Z, Jain S, Eapen V V., Haber JE. 2010. Sgs1 and Exo1 redundantly inhibit break-
579 induced replication and *de novo* telomere addition at broken chromosome ends. PLoS Genet. 6(5):25.
580 doi:10.1371/journal.pgen.1000973.

581 MacDonald JR, Ziman R, Yuen RKC, Feuk L, Scherer SW. 2014. The database of genomic variants: a
582 curated collection of structural variation in the human genome. Nucleic Acids Res. 42(D1):D986.
583 doi:10.1093/nar/gkt958.

584 Mangahas JL, Alexander MK, Sandell LL, Zakian VA. 2001. Repair of chromosome ends after telomere
585 loss in *Saccharomyces*. Mol Biol Cell. 12(12):4078–4089. doi:10.1091/mbc.12.12.4078.

586 Miller AW, Befort C, Kerr EO, Dunham MJ. 2013. Design and use of multiplexed chemostat arrays. J Vis
587 Exp.(72):1–6. doi:10.3791/50262.

588 Narayanan V, Mieczkowski PA, Kim HM, Petes TD, Lobachev KS. 2006. The pattern of gene amplification
589 is determined by the chromosomal location of hairpin-capped breaks. Cell. 125(7):1283–1296.
590 doi:10.1016/J.CELL.2006.04.042.

591 Ngo K, Epum EA, Friedman KL. 2020. Emerging non-canonical roles for the Rad51–Rad52 interaction in
592 response to double-strand breaks in yeast. Curr Genet. 66(5):917–926. doi:10.1007/s00294-020-01081-
593 z.

594 Obodo UC, Epum EA, Platts MH, Seloff J, Dahlson NA, Velkovsky SM, Paul SR, Friedman KL. 2016.
595 Endogenous hot spots of *de novo* telomere addition in the yeast genome contain proximal enhancers
596 that bind Cdc13. Mol Cell Biol. 36(12):1750–1763. doi:10.1128/mcb.00095-16.

597 Payen C, Di Rienzi SC, Ong GT, Pogachar JL, Sanchez JC, Sunshine AB, Raghuraman MK, Brewer BJ,
598 Dunham MJ. 2014. The dynamics of diverse segmental amplifications in populations of *Saccharomyces*

599 *cerevisiae* adapting to strong selection. G3 Genes, Genomes, Genet. 4(3):399–409.
600 doi:10.1534/g3.113.009365.

601 Pennock E, Buckley K, Lundblad V. 2001. Cdc13 delivers separate complexes to the telomere for end
602 protection and replication. Cell. 104(3):387–396. doi:10.1016/S0092-8674(01)00226-4.

603 Putnam CD, Kolodner RD. 2017. Pathways and mechanisms that prevent genome instability in
604 *Saccharomyces cerevisiae*. Genetics. 206(3):1187–1225. doi:10.1534/genetics.112.145805.

605 Rosswog C, Bartenhagen C, Welte A, Kahlert Y, Hemstedt N, Lorenz W, Cartolano M, Ackermann S,
606 Perner S, Vogel W, et al. 2021. Chromothripsis followed by circular recombination drives oncogene
607 amplification in human cancer. Nat Genet. 53(12):1673–1685. doi:10.1038/S41588-021-00951-7.

608 Shaikh TH. 2017. Copy number variation disorders. Curr Genet Med Rep. 5(4):183. doi:10.1007/S40142-
609 017-0129-2.

610 Sikorski RS, Hieter P. 1989. A system of shuttle vectors and yeast host strains designed for efficient
611 manipulation of DNA in *Saccharomyces cerevisiae*. Genetics. 122(1):19–27.
612 doi:10.1080/00362178385380431.

613 Singer MS, Gottschling DE. 1994. *TLC1*: template RNA component of *Saccharomyces cerevisiae*
614 telomerase. Science. 266(5184):404-9. doi:10.1126/science.7545955.

615 Song K, Minami JK, Huang A, Dehkordi SR, Lomeli SH, Luebeck J, Goodman MH, Moriceau G, Krijgsman
616 O, Dharanipragada P, et al. 2022. Plasticity of extrachromosomal and intrachromosomal BRAF
617 amplifications in overcoming targeted therapy dosage challenges. Cancer Discov. 12(4):1046–1069.
618 doi:10.1158/2159-8290.CD-20-0936.

619 Song RH, Gao CQ, Zhao J, Zhang JA. 2022. An update evolving view of copy number variations in
620 autoimmune diseases. Front Genet. 12:794348. doi:10.3389/FGENE.2021.794348.

621 Stellwagen AE, Haimberger ZW, Veatch JR, Gottschling DE. 2003. Ku interacts with telomerase RNA to
622 promote telomere addition at native and broken chromosome ends. Genes Dev. 17(19):2384–2395.
623 doi:10.1101/gad.1125903.

624 Strecker J, Stinus S, Pliego Caballero M, Szilard RK, Chang M, Durocher D. 2017. A sharp Pif1-dependent
625 threshold separates DNA double-strand breaks from critically short telomeres. Elife. 6:e23783.
626 doi:10.7554/eLife.23783.001. https://doi.org/10.7554/eLife.23783.001.

627 Turner KM, Deshpande V, Beyter D, Koga T, Rusert J, Lee C, Li B, Arden K, Ren B, Nathanson DA, et al.
628 2017. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity.
629 Nature. 543(7643):122–125. doi:10.1038/NATURE21356.

630 Wellinger RJ, Zakian VA. 2012. Everything you ever wanted to know about *Saccharomyces cerevisiae*
631 telomeres: beginning to end. Genetics. 191(4):1073–1105. doi:10.1534/genetics.111.137851.

632 Zhang F, Gu W, Hurles ME, Lupski JR. 2009. Copy number variation in human health, disease, and
633 evolution. Annu Rev Genomics Hum Genet. 10:451. doi:10.1146/ANNUREV.GENOM.9.081307.164217.

634 Zhang H, Zeidler AFB, Song W, Puccia CM, Malc E, Greenwell PW, Mieczkowski PA, Petes TD, Argueso JL.
635 2013. Gene copy-number variation in haploid and diploid strains of the yeast *Saccharomyces cerevisiae*.
636 Genetics. 193(3):785–801. doi:10.1534/GENETICS.112.146522.

637 Zhang W, Durocher D. 2010. *De novo* telomere formation is suppressed by the Mec1-dependent
638 inhibition of Cdc13 accumulation at DNA breaks. Genes Dev. 24(5):502–515. doi:10.1101/gad.1869110.

639

Figure 1

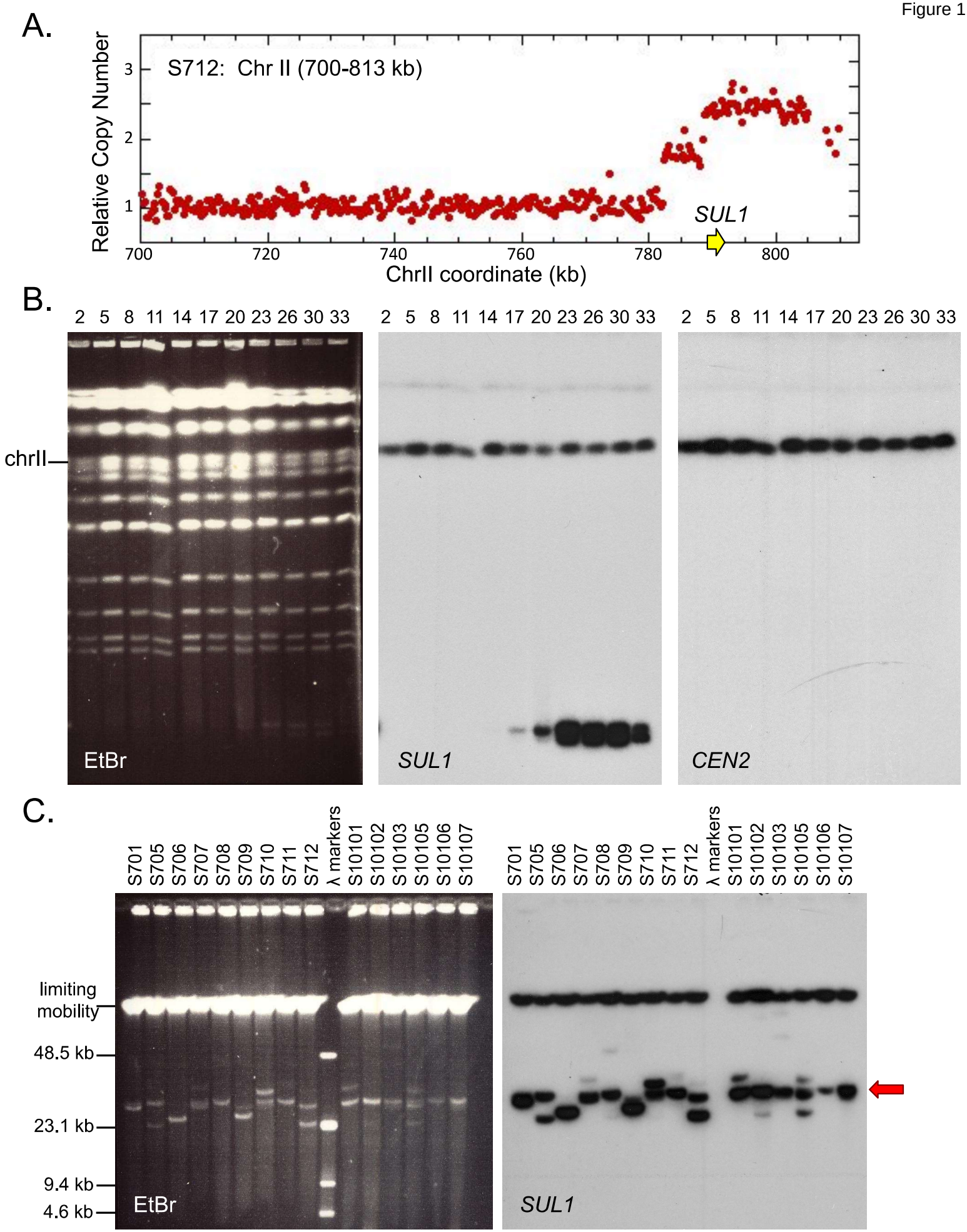

A.
S712: Chr II (700-813 kb)
Relative Copy Number
3
2
1
700
720
740
760
780
800
ChrII coordinate (kb)
SUL1
B.
2 5 8 11 14 17 20 23 26 30 33
chrII
EtBr
SUL1
CEN2
C.
S701 S705 S706 S707 S708 S709 S710 S711 S712 λ markers S10101 S10102 S10103 S10105 S10106 S10107
limiting mobility
48.5 kb
23.1 kb
9.4 kb
4.6 kb
EtBr
SUL1

Figure 2

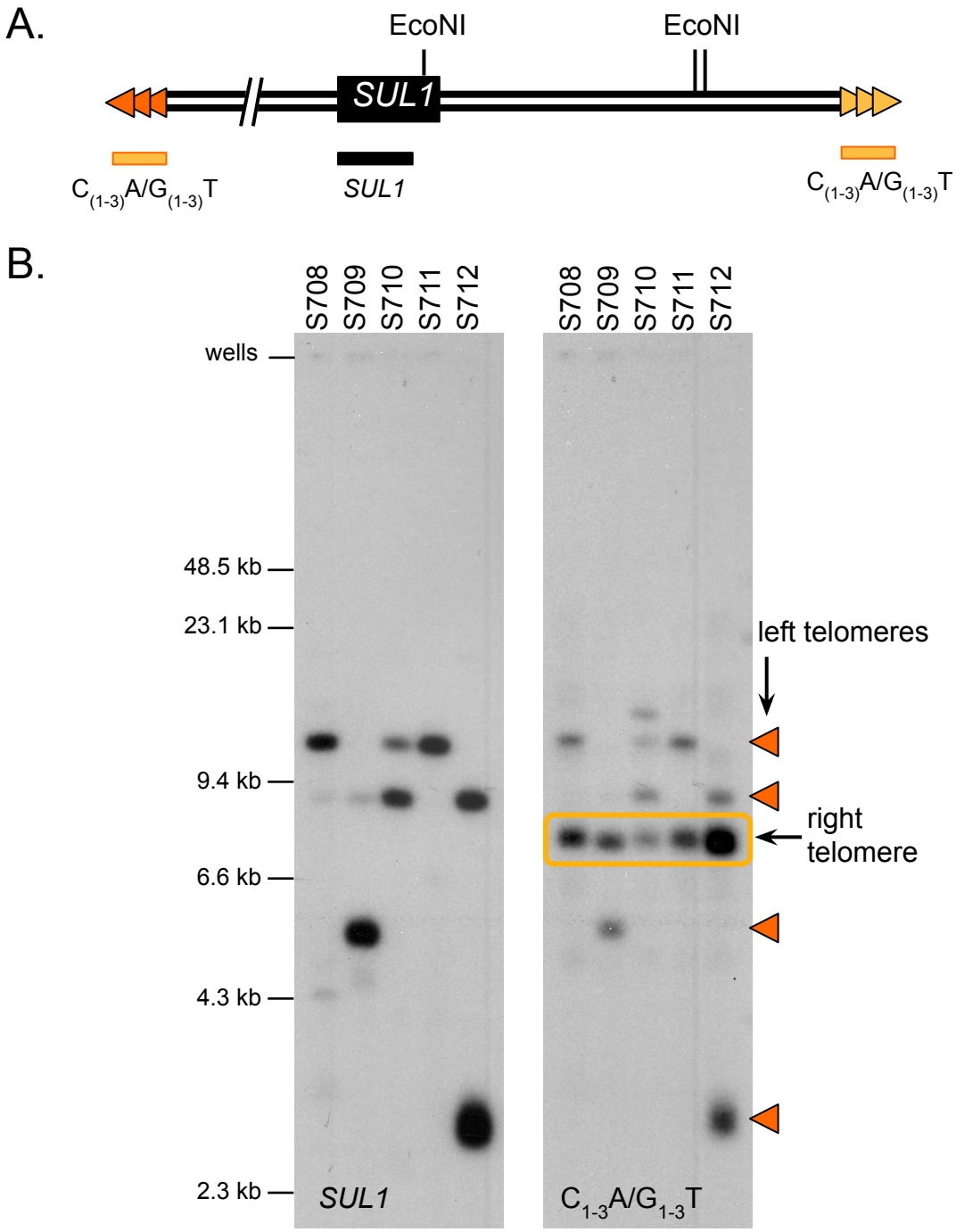

A.
EcoNI
EcoNI
SUL1
C(1-3)A/G(1-3)T
SUL1
C(1-3)A/G(1-3)T
B.
S708
S709
S710
S711
S712
S708
S709
S710
S711
S712
wells
48.5 kb
23.1 kb
9.4 kb
6.6 kb
4.3 kb
2.3 kb
SUL1
C1-3A/G1-3T
left telomeres
right telomere

Figure 3

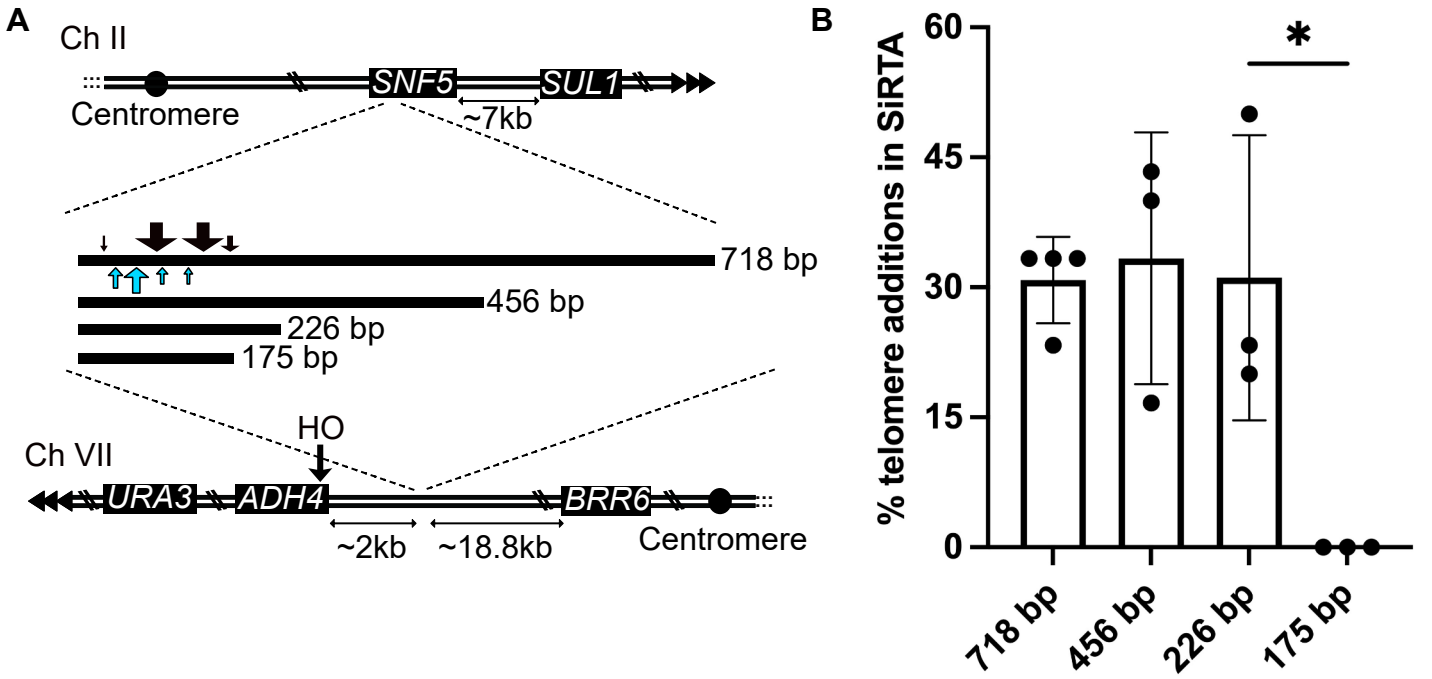
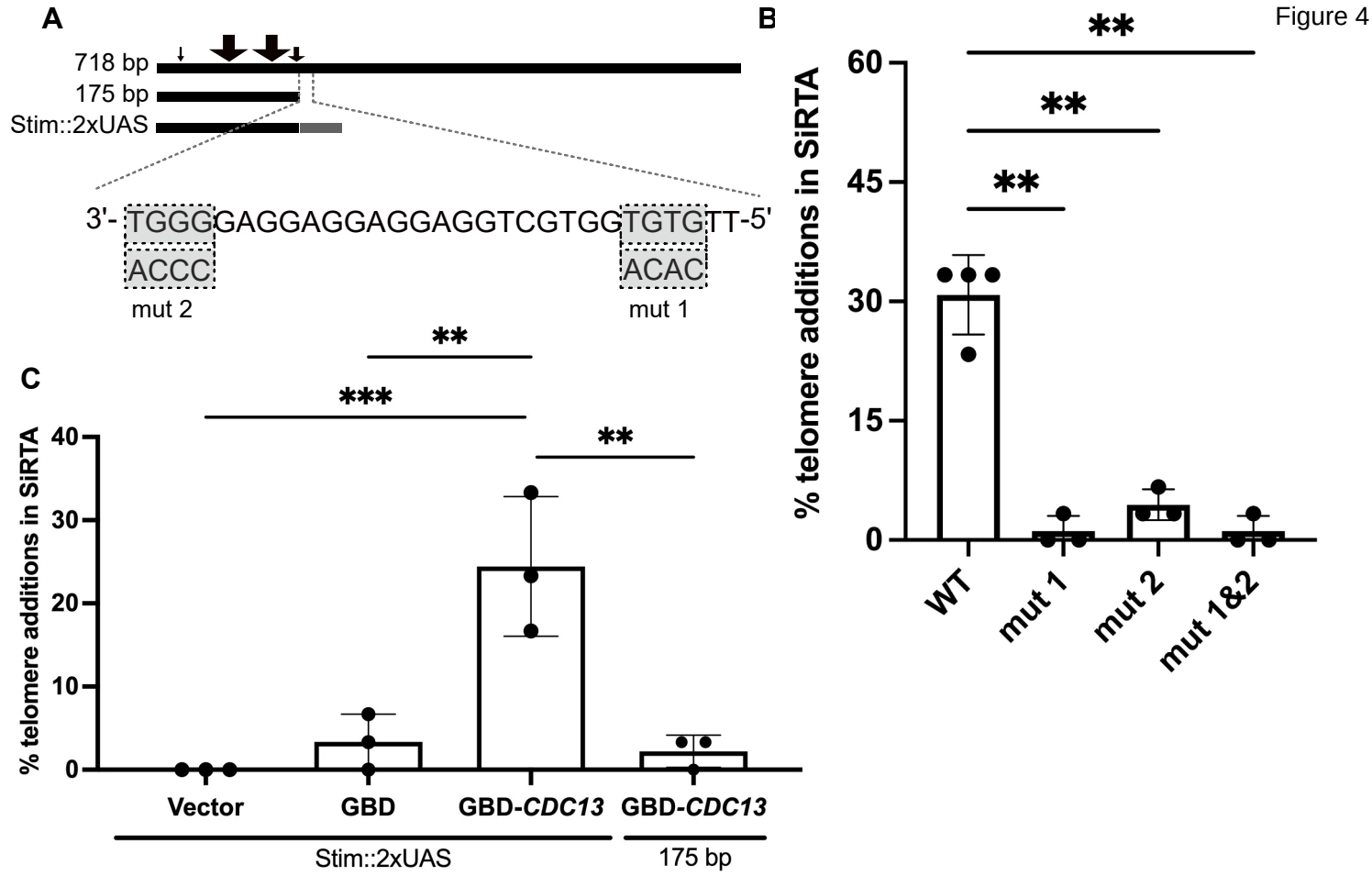

A
Ch II
Centromere
SNF5
SUL1
~7kb
718 bp
456 bp
226 bp
175 bp
HO
Ch VII
URA3
ADH4
BRR6
~2kb
~18.8kb
Centromere
B
% telomere additions in SiRTA
60
45
30
15
0
*
718 bp
456 bp
226 bp
175 bp

Figure 4


A
718 bp
175 bp
Stim::2xUAS
3'- TGGGGAGGAGGAGGAGGTCGTGGTGTGTT-5'
ACCC
ACAC
mut 2
mut 1
B
% telomere additions in SiRTA
60
45
30
15
0
**
**
**
WT
mut 1
mut 2
mut 1&2
C
% telomere additions in SiRTA
40
30
20
10
0
***
**
**
Vector
GBD
GBD-CDC13
GBD-CDC13
Stim::2xUAS
175 bp