SUPPLEMENTARY DATA

Acentric chromosome ends are prone to fusion with functional chromosome ends through a homology-directed rearrangement

Yuko Ohno[1], Yuki Ogiyama[1,†], Yoshino Kubota[1], Takuya Kubo[2] and Kojiro Ishii[1,3,*]

[1] Graduate School of Frontier Biosciences, Osaka University, Suita, Osaka 565-0871, Japan
[2] Graduate School of Environmental Science, Hokkaido University, Sapporo, Hokkaido 060-0810, Japan
[3] Institute for Academic Initiatives, Osaka University, Suita, Osaka 565-0871, Japan

* To whom correspondence should be addressed. Tel.: +81 6 6879 4660; Fax: +81 6 6879 4660; Email: ishii@fbs.osaka-u.ac.jp

[†] Present address: Institute of Human Genetics, CNRS UPR 1142, 34396 Montpellier, France

Supplementary Methods

Bayesian estimation of strain-dependent effects on survivor emergence

To evaluate the joint probability of telomere-fusion survivor emergence at the *Δcen1* screens in wild-type and mutant strain backgrounds, we developed a Bayesian generalised linear mixed model that allowed us to estimate the posterior distributions of strain-dependent effects on survivor emergence as random effects by fitting to the experimental data. Because the random effect parameter for each screen is estimated based on a subset of all of the data, by definition, hierarchical Bayesian modelling can provide a better inference by choosing a variance parameter that refers to the dispersion of whole strains. Here, $p(X, Y, Z)$ was used as a general notation for the joint probability, where X, Y and Z represent the telomere-fusion, removal of *cen1*-reintegrated false-positive survivors, and initial drug resistant-survivor emergence, respectively (40). The p(X, Y, Z) notation can be decomposed into three probabilities, i.e., $p(X, Y, Z) = p(X \mid Y, Z)\, p(Y \mid Z)\, p(Z)$, where the $p(A \mid B)$ notation refers to the conditional probability that event A occurs under the condition that event B occurs. All of these decomposed probabilities, $p(X \mid Y, Z)$, $p(Y \mid Z)$, and $p(Z)$, were estimated using a common framework for Bayesian logistic regression, of which the linear predictor was the sum of the intercepts, replicates, and strain-dependent effects. Hierarchical prior distributions were specified for the effects of replicates and strains as the Gaussian distribution around mean zero, whereas the intercepts were non-hierarchical priors and the variances were non-informative priors. The relative effect of each strain was evaluated as the difference in log odds of $p(X, Y, Z)$ between mutant and wild-type posteriors.

The posterior distributions of all parameters in the Bayesian statistical model were estimated using the Monte Carlo Markov chain (MCMC) method. Sampling from the marginal posterior distributions was performed using the Gibbs sampling software JAGS 3.4.0 (http://mcmc-jags.sourceforge.net/). The posterior samples were obtained by three independent chains in which 1000 values were sampled with a 5-step interval after 1000 burn-in MCMC steps. The convergences of the MCMC samples were confirmed such that all R-hat indexes for all parameters were close to unity. All statistical significances were checked by evaluating the 95% BCIs of the posterior distributions. The JAGS and R codes that are common for estimation of the probabilities of telomere-fusion, *cen1* reintegration, and initial survivor emergence are shown below.

```
# BUGS code
    for (i in 1:N.sample) {
        Y[i] ~ dbin(p[i], N[i])
        logit(p[i]) <- logit.p[i]
    }
    for (i in 1:N.sample) {
        logit.p[i] ~ dnorm(m[i], tau[1])
        m[i] <- alpha + beta[Group[i]]
    }
    alpha ~ dnorm(0, 1.0E-4)
    for (j in 1:N.group) {
        beta[j] ~ dnorm(0, tau[2])
    }
    for (k in 1:N.tau) {
        tau[k] <- 1.0 / (s[k] * s[k])
```

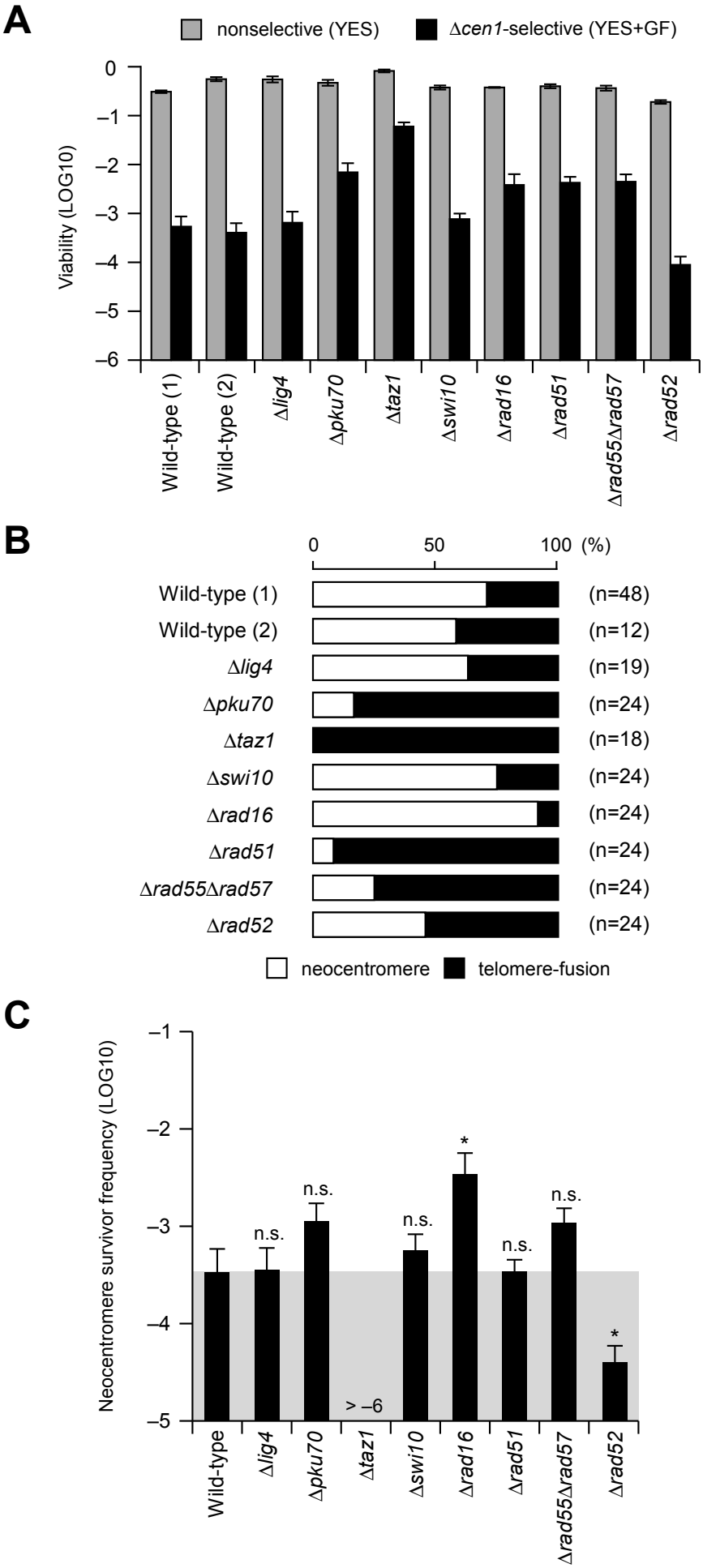
```
                s[k] ~ dunif(0, 1.0E+4)
        }


# R code
library(rjags)

N.sample <- nrow(d)
N.group <- length(levels(d$group))
list.data <- list(
        N.sample = N.sample,
        N.group = N.group,
        N.tau = 2,
        Y = d$y,
        N = d$N,
        Group = d$group_id
)
inits <- list( # parameter initial values
        alpha = 0,
        beta = rnorm(N.group, 0, 0.1),
        s = c(1, 1)
)

n.burnin <- 1000
n.chain <- 3
n.thin <- 5
n.iter <- n.thin * 1000

model <- jags.model(
        file = "model.bug",
        data = list.data,
        inits = inits,
        n.chain = 3 # !!!
)
update(model, n.burnin) # burn in
post.mcmc.list <- coda.samples(
        model,
        c("alpha", "beta", "s"),
        thin = n.thin, n.iter = n.iter
)
```

Supplementary Table S1. Yeast strain list

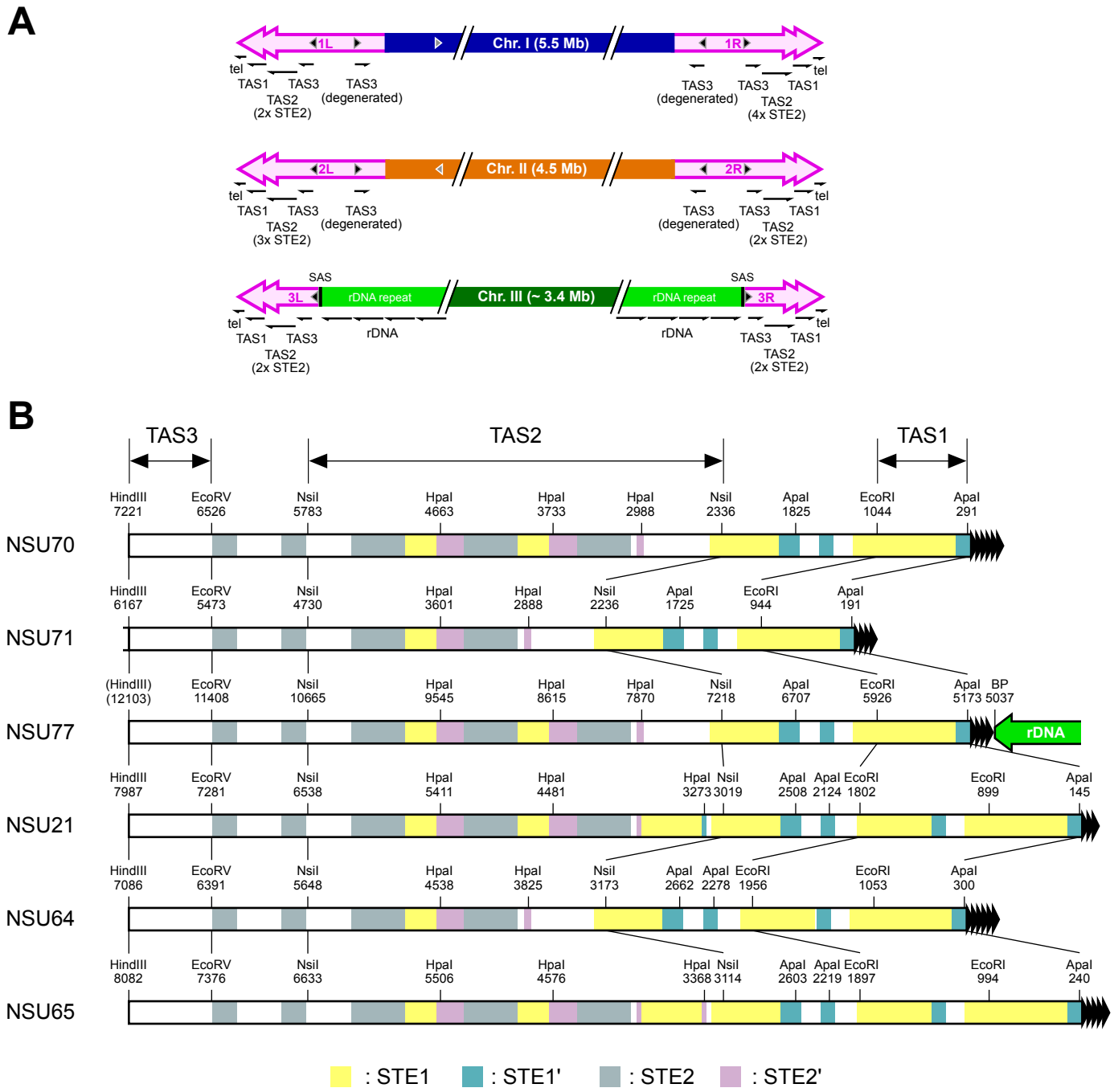
Strain name	Genotype
KYP33	*h⁻ leu1 ura4*
KYP378	*h⁻ leu1 ura4 cen1L::*P^{adh1}*-loxP cen1R::ura4*$^{+}$*-loxP-kanR*^{ORF}
KYP810	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-31 (1R;2L fusion)*
KYP812	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-33 (1L;3R fusion)*
KYP815	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-36 (1L;2R fusion)*
KYP821	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-42 (1L;2R fusion)*
KYP826	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-47 (1L;2L fusion)*
KYP828	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-49 (1L;3R fusion)*
KYP830	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-51 (1R;2R fusion)*
KYP832	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-53 (1R;3L fusion)*
KYP838	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-59 (1L;2R fusion)*
KYP386	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-63 (1R;3L fusion)*
KYP387	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-64 (1R;2R fusion)*
KYP388	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-65 (1L;2R fusion)*
KYP1272	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-73 (1L;2L fusion)*
KYP1273	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-77 (1R;2R fusion)*
KYP4478	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-96 (1R;2R fusion)*
KYP4479	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-97 (1L;2L fusion)*
KYP4480	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-98 (1L;2L fusion)*
KYP4481	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-99 (1R;2L fusion)*
KYP4482	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-100 (1R;2R fusion)*
KYP3948	*h⁻ leu1 ura4 cen1L::*P^{adh1}*-loxP cen1R::ura4*$^{+}$*-loxP-kanR*^{ORF} *rad52-GFP::nat taz1-mCherry::kanR*
KYP2481	*h⁻ leu1 ura4 cen1L::*P^{adh1}*-loxP cen1R::ura4*$^{+}$*-loxP-kanR*^{ORF} *Δpku70::hph*
KYP2428	*h⁻ leu1 ura4 cen1L::*P^{adh1}*-loxP cen1R::ura4*$^{+}$*-loxP-kanR*^{ORF} *Δlig4::hph*
KYP3949	*h⁻ leu1 ura4 cen1L::*P^{adh1}*-loxP cen1R::ura4*$^{+}$*-loxP-kanR*^{ORF} *Δrad16::nat*
KYP2339	*h⁻ leu1 ura4 cen1L::*P^{adh1}*-loxP cen1R::ura4*$^{+}$*-loxP-kanR*^{ORF} *Δswi10::hph*
KYP674	*h⁻ leu1 ura4 cen1L::*P^{adh1}*-loxP cen1R::ura4*$^{+}$*-loxP-kanR*^{ORF} *Δaz1::hph*
KYP2428	*h⁻ smt-0 leu1 ura4 cen1L::*P^{adh1}*-loxP cen1R::ura4*$^{+}$*-loxP-kanR*^{ORF} *Δrad51::hph*
KYP2429	*h⁻ smt-0 leu1 ura4 cen1L::*P^{adh1}*-loxP cen1R::ura4*$^{+}$*-loxP-kanR*^{ORF} *Δrad52::hph*
KYP2554	*h⁻ smt-0 leu1 ura4 cen1L::*P^{adh1}*-loxP cen1R::ura4*$^{+}$*-loxP-kanR*^{ORF} *Δrad55::hph Δrad57::nat*
KYP2131	*h⁻ leu1 ura4 cen3L::*P^{adh1}*-loxP cen3R::ura4*$^{+}$*-loxP-kanR*^{ORF}
KYP4885	*h⁻ leu1 ura4 cen3L::*P^{adh1}*-loxP cen3R::ura4*$^{+}$*-loxP-kanR*^{ORF} *SAS(3L)::LEU2*
KYP4889	*h⁻ leu1 ura4 cen3L::*P^{adh1}*-loxP cen3R::ura4*$^{+}$*-loxP-kanR*^{ORF} *SAS(3R)::LEU2*
KYP4890	*h⁻ leu1 ura4 cen3L::*P^{adh1}*-loxP cen3R::ura4*$^{+}$*-loxP-kanR*^{ORF} *SAS(3L)::LEU2 SAS(3R)::LEU2*
KYP811	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-32*
KYP814	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-35*
KYP816	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-37*
KYP817	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-38*
KYP818	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-39*
KYP819	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-40*
KYP820	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-41*
KYP822	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-43*
KYP823	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-44*
KYP824	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-45*
KYP825	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-46*
KYP827	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-48*
KYP829	*h⁻ leu1 ura4 Δcen1::*P^{adh1}*-loxP-kanR cd1-50*

Supplementary Table S2. Oligonucleotide list

caPCR primers	
7926	5'-GGGTTGCAAAGTATGATTGTGGTAA-3'
14206	5'-GCATAAAGATGGTACTTCAA-3'
31268	5'-TGTTGAATGTCAGAACCAACTGTTGCAT-3'
Asp(1L)	5'-TCCGTACGCAATTATTCGCA-3'
Asp(2L)	5'-TAACGGTTGCGTTTTCTTCC-3'
TAS1_TAIL-II	5'-GGTGAAATTCACTAAGTGTAATACAGTAGTGCAGTG-3'
TAS3_TAS2-F1	5'-GGAGTAAGTAGTGAGTCAG-3'
TAS1_TAS2-F1	5'-TTAAACGATTTTGGAGAGAG-3'
subtelrD-R4	5'-CCTCTAAGCCAGAATCCG-3'
60241	5'-TTGAAGTACCATCTTTATGC-3'
c-rDNA-t	5'-GAAGTTTGTCAATGGAAGGG-3'
SAS-Rv1	5'-CGGGATCCGCGGCCGCAGGCTGGCTACTGTTTTAC-3'
TAS1-Rv1	5'-CGGGATCCGCGGCCGCACCACGTAACCTTGTAACC-3'
TAS2 ligation-independent cloning (LIC) primers	
TAS2cen-infusion-KS-R	5'-TTAACCATTCCACGATGCATTTGATATCGAATTCCTGCAGC-3'
TAS2tel-infusion-KS-F	5'-TATTGGAAAGTCAGATGCATGCTTATCGATACCGTCGAC-3'
qPCR primers	
act1-56F	5'- ATCCAACCGTGAGAAGATGACT-3'
act1-56R	5'- AAACAGCTTGAATAGCAACATAAAAG-3'
STE2-1F	5'- GGAGAACAAAGAAGTAAGTAAAGTAAGAAA-3'
STE2-98R	5'- AAAGTACACCGCATGTTCCTATTAT-3'
Telomere probe	
tel	5'-CGTGTAACCACGTAACCTTGTAACCCGATC-3'


A
nonselective (YES)
Δcen1-selective (YES+GF)
Viability (LOG10)
0
–1
–2
–3
–4
–5
–6
Wild-type (1)
Wild-type (2)
Δlig4
Δpku70
Δtaz1
Δswi10
Δrad16
Δrad51
Δrad55Δrad57
Δrad52
B
0
50
100 (%)
Wild-type (1) (n=48)
Wild-type (2) (n=12)
Δlig4 (n=19)
Δpku70 (n=24)
Δtaz1 (n=18)
Δswi10 (n=24)
Δrad16 (n=24)
Δrad51 (n=24)
Δrad55Δrad57 (n=24)
Δrad52 (n=24)
neocentromere
telomere-fusion
C
Neocentromere survivor frequency (LOG10)
–1
–2
–3
–4
–5
n.s.
n.s.
n.s.
n.s.
n.s.
*
*
> –6
Wild-type
Δlig4
Δpku70
Δtaz1
Δswi10
Δrad16
Δrad51
Δrad55Δrad57
Δrad52

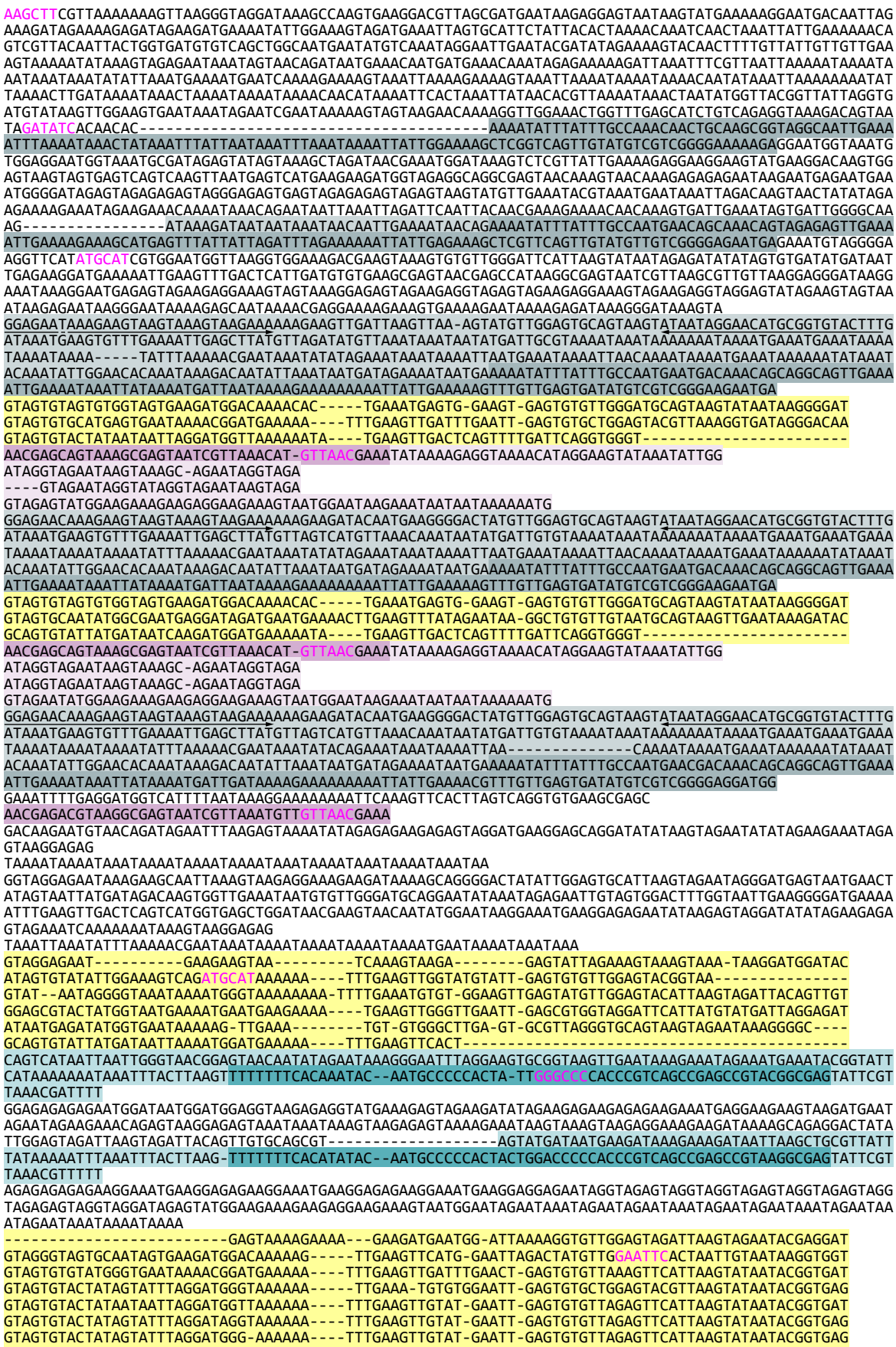
Supplementary Figure S1. Profile of the centromere deletion screen and characterisation of the resulting survivors. (A) The viabilities of total cells (grey) and *cen1*-deleted cells (black), as determined by their colony formation frequency on nonselective YES plates and *Δcen1*-selective G418- and 5-fluoroorotic acid-containing YES (YES+GF) plates. Data are represented as the mean ± SEMs of n = 6 replicates performed simultaneously. (B) The ratios of neocentromere formation and telomere-fusion in the survivors obtained in each screen. The sample size (n) varied depending on the frequency of telomere-fusion events. (C) The frequencies of neocentromere survivor generation upon *cen1* deletion in the indicated mutant backgrounds. Data are represented as the mean ± SEM of n = 6 replicates. *, the 95% BCI does not include zero; n.s., the 95% BCI includes zero (not significant).

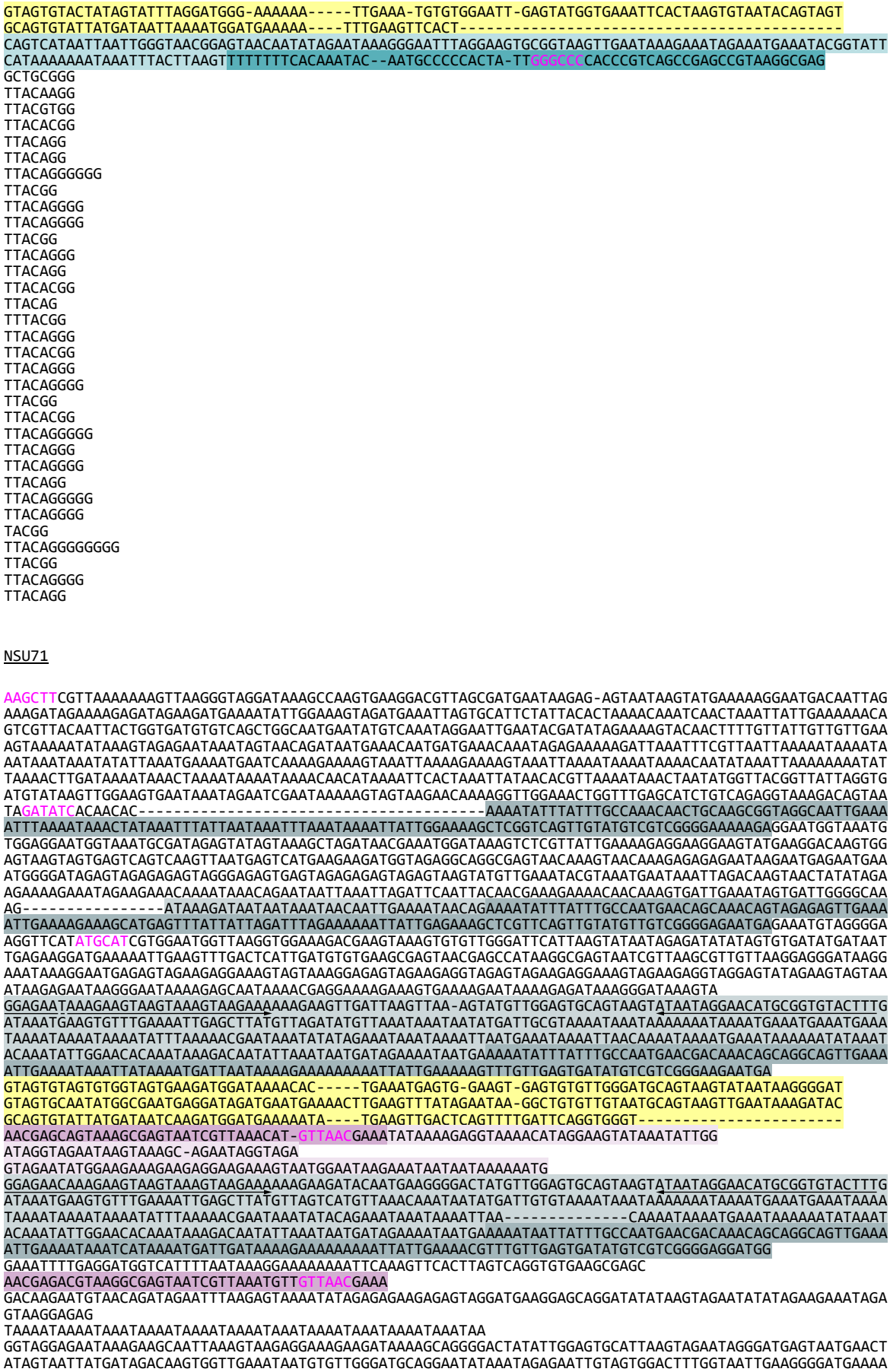



```
4943  GTAGTGTAGTGTGGTAGTGAAGATGGACAAAACAC-----TGAAATGAGTG-GAAGT-GAGTGTGTTGGGATGCAGTAAGTATAATAAGGGGAT  4857
4856  GTAGTGTGCATGAGTGAATAAAACGGATGAAAAA----TTTGAAGTTGATTTGAATT-GAGTGTGCTGGAGTACGTTAAAGGTGATAGGGACAA  4768
4767  GTAGTGTACTATAATAATTAGGATGGTTAAAAAATA----TGAAGTTGACTCAGTTTTGATTCAGGTGGGT----------------------  4701
4017  GTAGTGTAGTGTGGTAGTGAAGATGGACAAAACAC-----TGAAATGAGTG-GAAGT-GAGTGTGTTGGGATGCAGTAAGTATAATAAGGGGAT  3931
3930  GTAGTGCAATATGGCGAATGAGGATAGATGAATGAAAACTTGAAGTTTATAGAATAA-GGCTGTGTTGTAATGCAGTAAGTTGAATAAAGATAC  3838
3837  GCAGTGTATTATGATAATCAAGATGGATGAAAAATA----TGAAGTTGACTCAGTTTTGATTCAGGTGGGT----------------------  3771
2429  GTAGGAGAAT----------GAAGAAGTAATCAAA---------GTAAGAGAGTATTAGA--------AAGTAAAGTAAA-TAAGGATGGATAC  2364
2363  ATAGTGTATATTGGAAAGTCAGATGCATAAAAAA----TTTGAAGTTGGTATGTATT-GAGTGTGTTGGAGTACGGTAA--------------  2290
2289  GTAT--AATAGGGGTAAATAAAATGGGTAAAAAAAA-TTTTGAAATGTGT-GGAAGTTGAGTATGTTGGAGTACATTAAGTAGATTACAGTTGT  2200
2199  GGAGCGTACTATGGTAATGAAAATGAATGAAGAAAA----TGAAGTTGGGTTGAATT-GAGCGTGGTAGGATTCATTATGTATGATTAGGAGAT  2111
2110  ATAATGAGATATGGTGAATAAAAAGTTGAAA---------TGT-GTGGGCTTGA-GT-GCGTTAGGGTGCAGTAAGTAGAATAAAGGGGC----  2033
2032  GCAGTGTATTATGATAATTAAAATGGATGAAAAA----TTTGAAGTTCACT-------------------------------------------  1986
1176  ------------------------GAGTAAAAGAAAA---GAAGATGAAT-GGATTAAAAGGTGTTGGAGTAGATTAAGTAGAATACGAGGAT  1112
1111  GTAGGGTAGTGCAATAGTGAAGATGGACAAAAAG-----TTGAAGTTCATG-GAATTAGACTATGTTGGAATTCACTAATTGTAATAAGGTGGT  1024
1023  GTAGTGTGTATGGGTGAATAAAACGGATGAAAAA----TTTGAAGTTGATTTGAACT-GAGTGTGTTAAAGTTCATTAAGTATAATACGGTGAT   935
 934  GTAGTGTACTATAGTATTTAGGATGGGTAAAAAA-----TTGAAATGTGTG-GAATT-GAGTGTGCTGGAGTACGTTAAGTATAATACGGTGAG   848
 847  GTAGTGTACTATAATAATTAGGATGGTTAAAAAA----TTTGAAGTTGTAT-GAATT-GAGTGTGTTAGAGTTCATTAAGTATAATACGGTGAT   760
 759  GTAGTGTACTATAGTATTTAGGATAGGTAAAAAA----TTTGAAGTTGTAT-GAATT-GAGTGTGTTAGAGTTCATTAAGTATAATACGGTGAG   672
 671  GTAGTGTACTATAGTATTTAGGATGGG-AAAAAA----TTTGAAGTTGTAT-GAATT-GAGTGTGTTAGAGTTCATTAAGTATAATACGGTGAG   585
 584  GTAGTGTACTATAGTATTTAGGATGGG-AAAAAA-----TTGAAATGTGTG-GAATT-GAGTATGGTGAAATTCACTAAGTGTAATACAGTAGT   499
 498  GCAGTGTATTATGATAATTAAAATGGATGAAAAA----TTTGAAGTTCACT-------------------------------------------   452
```

D

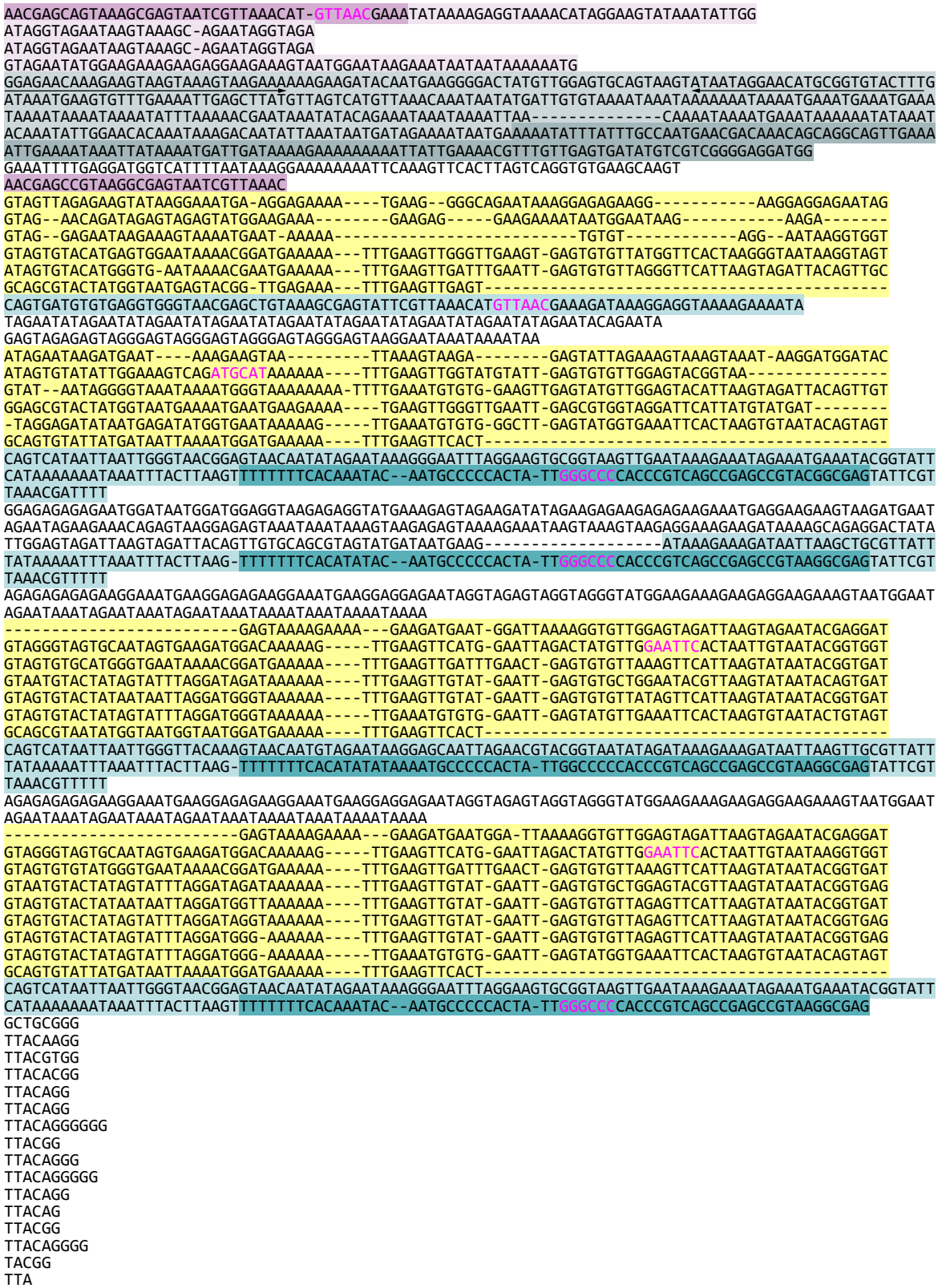
Wang and Baumann (2008) / NSU70




NSU71

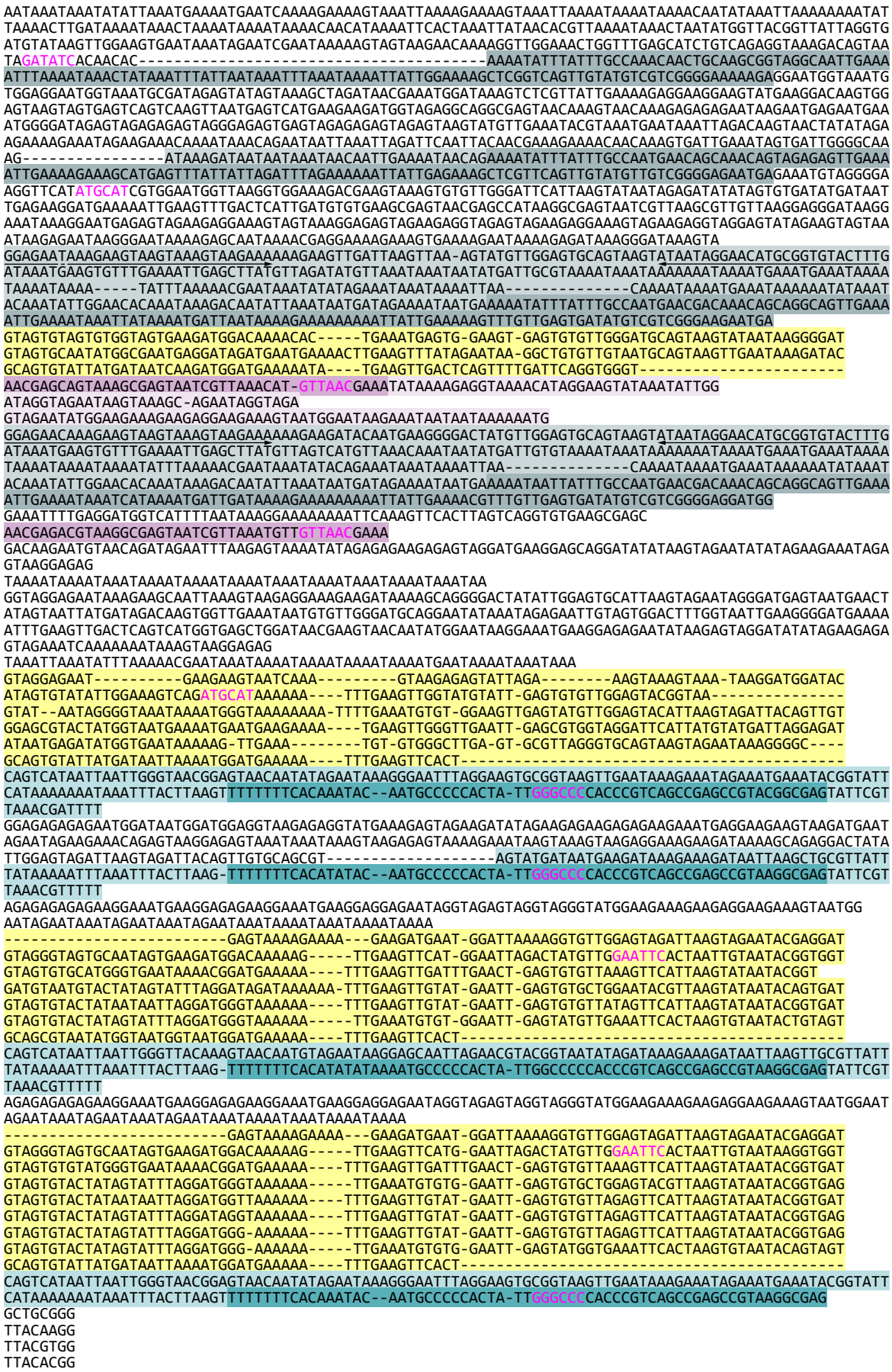
NSU77


NSU21



NSU64

Supplementary Figure S2

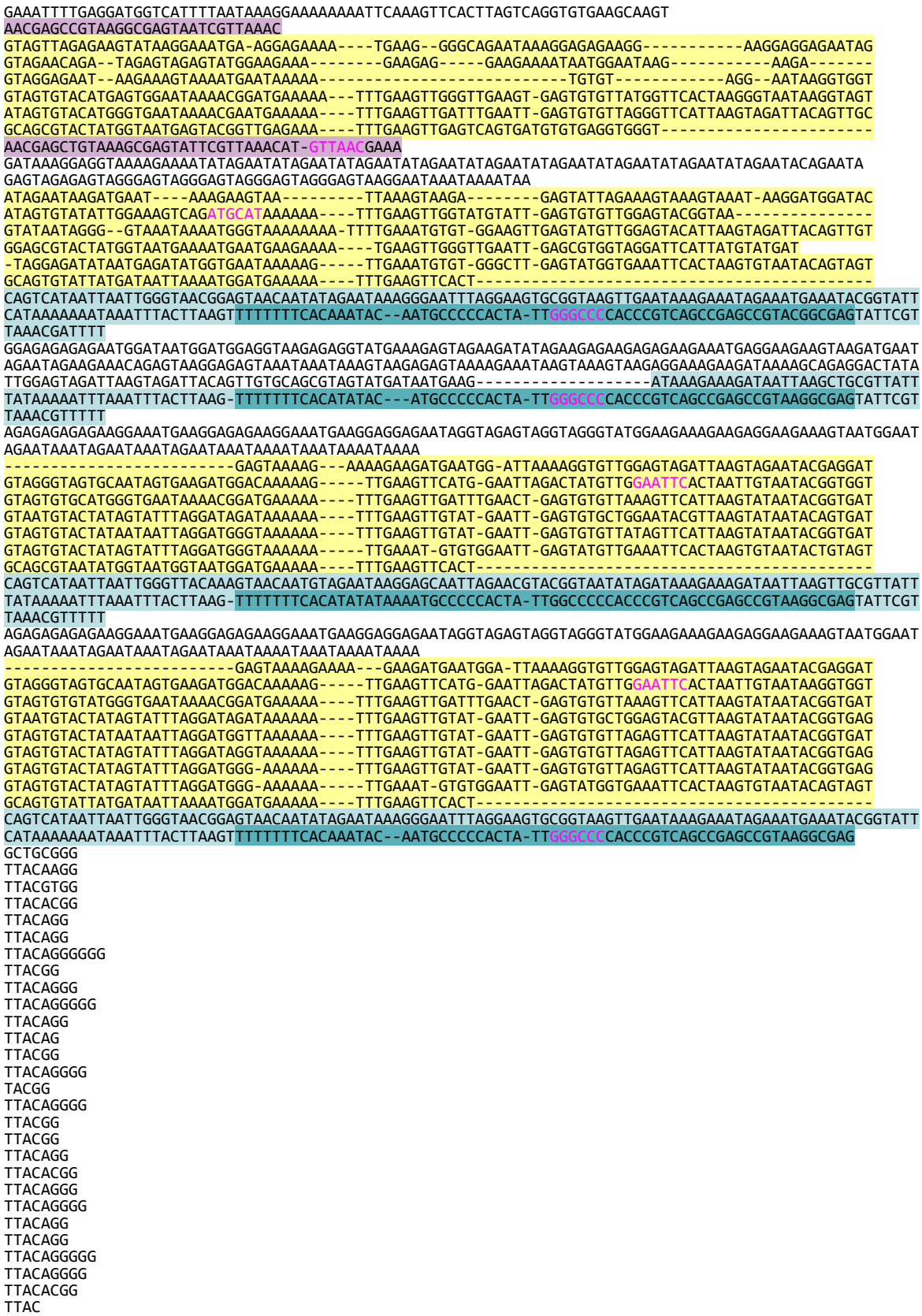


TTACAGG
TTACAGG
TTACAGGGGGG
TTACGG
TTACAGGGG
TTACAGGGG
TTACGG
TTACAGGG
TTACAGG
TTACACGG
TTACAG
TTTACGG
TTACAGGG
TTACACGG
TTACAG
TTACAGG
TTACAGGG
TTACAGGGG
TTACGG
TTACAGG
TTACAGGG
TTACAGGGGGG
TTACAGGGG
TTACAGG
TTACAGGGG
TTACAGGGGG
TTACAGGG
TTACACGG
TTACAGGG
TTACAGGGGG
TT

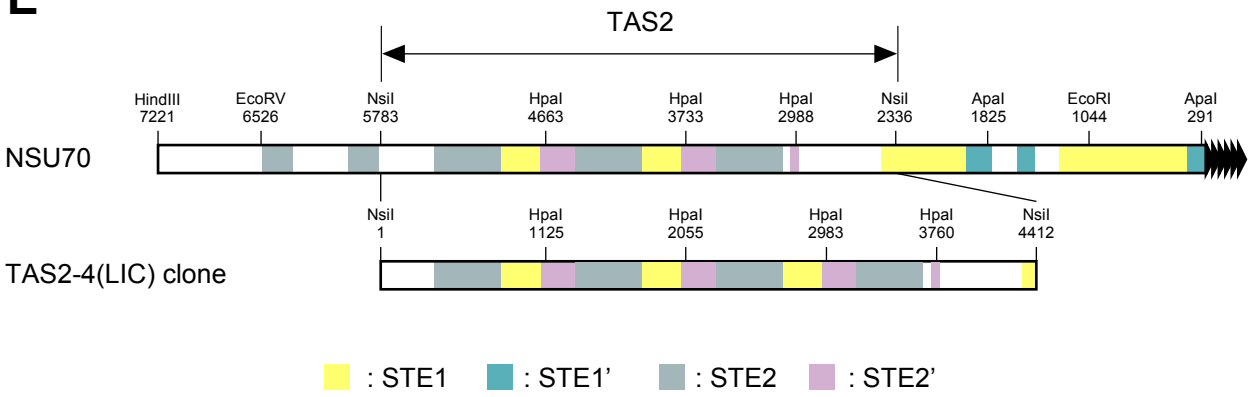
NSU65

AAGCTTCGTTAAAAAAAGTTAAGGGTAGGATAAAGCCAAGTGAAGGACGTTAGCGATGAATAAGAGGAGTAATAAGTATGAAAAAGGAATGACAATTAG
AAAGATAGAAAAGAGATAGAAGATGAAAATATTGGAAAGTAGATGAAATTAGTGCATTCTATTACACTAAAACAAATCAACTAAATTATTGAAAAAACA
GTCGTTACAATTACTGGTGATGTGTCAGCTGGCAATGAATATGTCAAATAGGAATTGAATACGATATAGAAAAGTACAACTTTTGTTATTGTTGTTGAA
AGTAAAAATATAAAGTAGAGAATAAATAGTAACAGATAATGAAACAATGATGAAACAATGATGAAACAAATAGAGAAAAAGATTAAATTTCGTTAATTA
AAAATAAAATAAATAAATAAATATATTAAATGAAAATGAATCAAAAGAAAAGTAAATTAAAAGAAAAGTAAATTAAAATAAAATAAAACAATATAAATT
AAAAAAAATATTAAAACTTGATAAAATAAACTAAAATAAAATAAAACAACATAAAATTCACTAAATTATAACACGTTAAAATAAACTAATATGGTTACG
GTTATTAGGTGATGTATAAGTTGGAAGTGAATAAATAGAATCGAATAAAAAGTAGTAAGAACAAAAGGTTGGAAACTGGTTTGAGCATCTGTCAGAGGT
AAAGACAGTAATAGATATCACAACAC----------AAAATATTTATTTGCCAAACAACTGCAAGCGGTAGGCAATTGAAA
ATTTAAAATAAACTATAAATTTATTAATAAATTTAAATAAAATTATTGGAAAAGCTCGGTCAGTTGTATGTCGTCGGGGAAAAAGAGGAATGGTAAATG
TGGAGGAATGGTAAATGCGATAGAGTATAGTAAAGCTAGATAACGAAATGGATAAAGTCTCGTTATTGAAAAGAGGAAGGAAGTATGAAGGACAAGTGG
AGTAAGTAGTGAGTCAGTCAAGTTAATGAGTCATGAAGAAGATGGTAGAGGCAGGCGAGTAACAAAGTAACAAAGAGAGAGAATAAGAATGAGAATGAA
ATGGGGATAGAGTAGAGAGAGTAGGGAGAGTGAGTAGAGAGAGTAGAGTAAGTATGTTGAAATACGTAAATGAATAAATTAGACAAGTAACTATATAGA
AGAAAAGAAATAGAAGAAACAAAATAAACAGAATAATTAAATTAGATTCAATTACAACGAAAGAAAACAACAAAGTGATTGAAATAGTGATTGGGGCAA
AG----------------ATAAAGATAATAATAAATAACAATTGAAAATAACAGAAAATATTTATTTGCCAATGAACAGCAAACAGTAGAGAGTTGAAA
ATTGAAAAGAAAGCATGAGTTTATTATTAGATTTAGAAAAAATTATTGAGAAAGCTCGTTCAGTTGTATGTTGTCGGGGAGAATGAGAAATGTAGGGGA
AGGTTCATATGCATCGTGGAATGGTTAAGGTGGAAAGACGAAGTAAAGTGTGTTGGGATTCATTAAGTATAATAGAGATATATAGTGTGATATGATAAT
TGAGAAGGATGAAAAATTGAAGTTGACTCATTGATGTGTGAAGCGAGTAACGAGCCATAAGGCGAGTAATCGTTAAGCGTTGTTAAGGAGGGATAAGGA
AATAAAGGAATGAGAGTAGAAGAGGAAAGTAGTAAAGGAGAGTAGAAGAGGTAGAGTAGAAGAGGAAAGTAGAAGAGGTAGGAGTATAGAAGTAGTAAA
TAAGAGAATAAGAGAATAAGGGAATAAAAGAGCAATAAAACGAGGAAAAGAAAGTGAAAAGAATAAAAGAGATAAAGGGATAAAGTA
GGAGAATAAAGAAGTAAGTAAAGTAAGAAAAAAGAAGTTGATTAAGTTAA-AGTATGTTGGAGTGCAGTAAGTATAATAGGAACATGCGGTGTACTTTG
ATAAATGAAGTGTTTGAAAATTGAGCTTATGTTAGATATGTTAAATAAATAATATGATTGCGTAAAATAAATAAAAAAATAAAATGAAATGAAATAAAA
TAAAATAAAATATTTAAAAACGAATAAATATATAGAAATAAATAAAATTAATGAAATAAAATTAA-----CAAAATAAAATGAAATAAAAAATATAAAT
ACAAATATTGGAACACAAATAAAGACAATATTAAATAATGATAGAAAATAATGAAAAATATTTATTTGCCAATGAATGACAAACAGCAGGCAGTTGAAA
ATTGAAAATAAATTATAAAATGATTAATAAAAGAAAAAAAAATTATTGAAAAAGTTTGTTGAGTGATATGTCGTCGGGAAGAATGA
GTAGTGTAGTGTGGTAGTGAAGATGGACAAAACAC-----TGAAATGAGTG-GAAGT-GAGTGTGTTGGGATGCAGTAAGTATAATAAGGGGAT
GTAGTGTGCATGAGTGAATAAAACGGATGAAAAA----TTTGAAGTTGATTTGAATT-GAGTGTGCTGGAGTACGTTAAAGGTGATAGGGACAA
GTAGTGTACTATAATAATTAGGATGGTTAAAAAATA----TGAAGTTGACTCAGTTTTGATTCAGGTGGGT----------
AACGAGCAGTAAAGCGAGTAATCGTTAAACAT-GTTAACGAAATATAAAAGAGGTAAAACATAGGAAGTATAAATATTGG
ATAGGTAGAATAAGTAAAGC-AGAATAGGTAGA
----GTAGAATAGGTATAGGTAGAATAAGTAGA
GTAGAGTATGGAAGAAAGAAGAGGAAGAAAGTAATGGAATAAGAAATAATAATAAAAAATG
GGAGAACAAAGAAGTAAGTAAAGTAAGAAAAAAGAAGATACAATGAAGGGGACTATGTTGGAGTGCAGTAAGTATAATAGGAACATGCGGTGTACTTTG
ATAAATGAAGTGTTTGAAAATTGAGCTTATGTTAGTCATGTTAAACAAATAATATGATTGTGTAAAATAAATAAAAAAATAAAATGAAATGAAATGAAA
TAAAATAAAATAAAATATTTAAAAACGAATAAATATATAGAAATAAATAAAATTAATGAAATAAAATTAACAAAATAAAATGAAATAAAAAATATAAAT
ACAAATATTGGAACACAAATAAAGACAATATTAAATAATGATAGAAAATAATGAAAAATATTTATTTGCCAATGAATGACAAACAGCAGGCAGTTGAAA
ATTGAAAATAAATTATAAAATGATTAATAAAAGAAAAAAAAATTATTGAAAAAGTTTGTTGAGTGATATGTCGTCGGGAAGAATGA
GTAGTGTAGTGTGGTAGTGAAGATGGACAAAACAC-----TGAAATGAGTG-GAAGT-GAGTGTGTTGGGATGCAGTAAGTATAATAAGGGGAT
GTAGTGCAATATGGCGAATGAGGATAGATGAATGAAAACTTGAAGTTTATAGAATAA-GGCTGTGTTGTAATGCAGTAAGTTGAATAAAGATAC
GCAGTGTATTATGATAATCAAGATGGATGAAAAATA----TGAAGTTGACTCAGTTTTGATTCAGGTGGGT----------
AACGAGCAGTAAAGCGAGTAATCGTTAAACAT-GTTAACGAAATATAAAAGAGGTAAAACATAGGAAGTATAAATATTGG
ATAGGTAGAATAAGTAAAGC-AGAATAGGTAGA
ATAGGTAGAATAAGTAAAGC-AGAATAGGTAGA
GTAGAATATGGAAGAAAGAAGAGGAAGAAAGTAATGGAATAAGAAATAATAATAAAAAATG
GGAGAACAAAGAAGTAAGTAAAGTAAGAAAAAAGAAGATACAATGAAGGGGACTATGTTGGAGTGCAGTAAGTATAATAGGAACATGCGGTGTACTTTG
ATAAATGAAGTGTTTGAAAATTGAGCTTATGTTAGTCATGTTAAACAAATAATATGATTGTGTAAAATAAATAAAAAAATAAAATGAAATGAAATGAAA
TAAAATAAAATAAAATATTTAAAAACGAATAAATATACAGAAATAAATAAAATTAA--------------CAAAATAAAATGAAATAAAAAATATAAAT
ACAAATATTGGAACACAAATAAAGACAATATTAAATAATGATAGAAAATAATGAAAAATATTTATTTGCCAATGAACGACAAACAGCAGGCAGTTGAAA
ATTGAAAATAAATTATAAAATGATTGATAAAAGAAAAAAAAATTATTGAAAACGTTTGTTGAGTGATATGTCGTCGGGGAGGATGG

Supplementary Figure S2



E

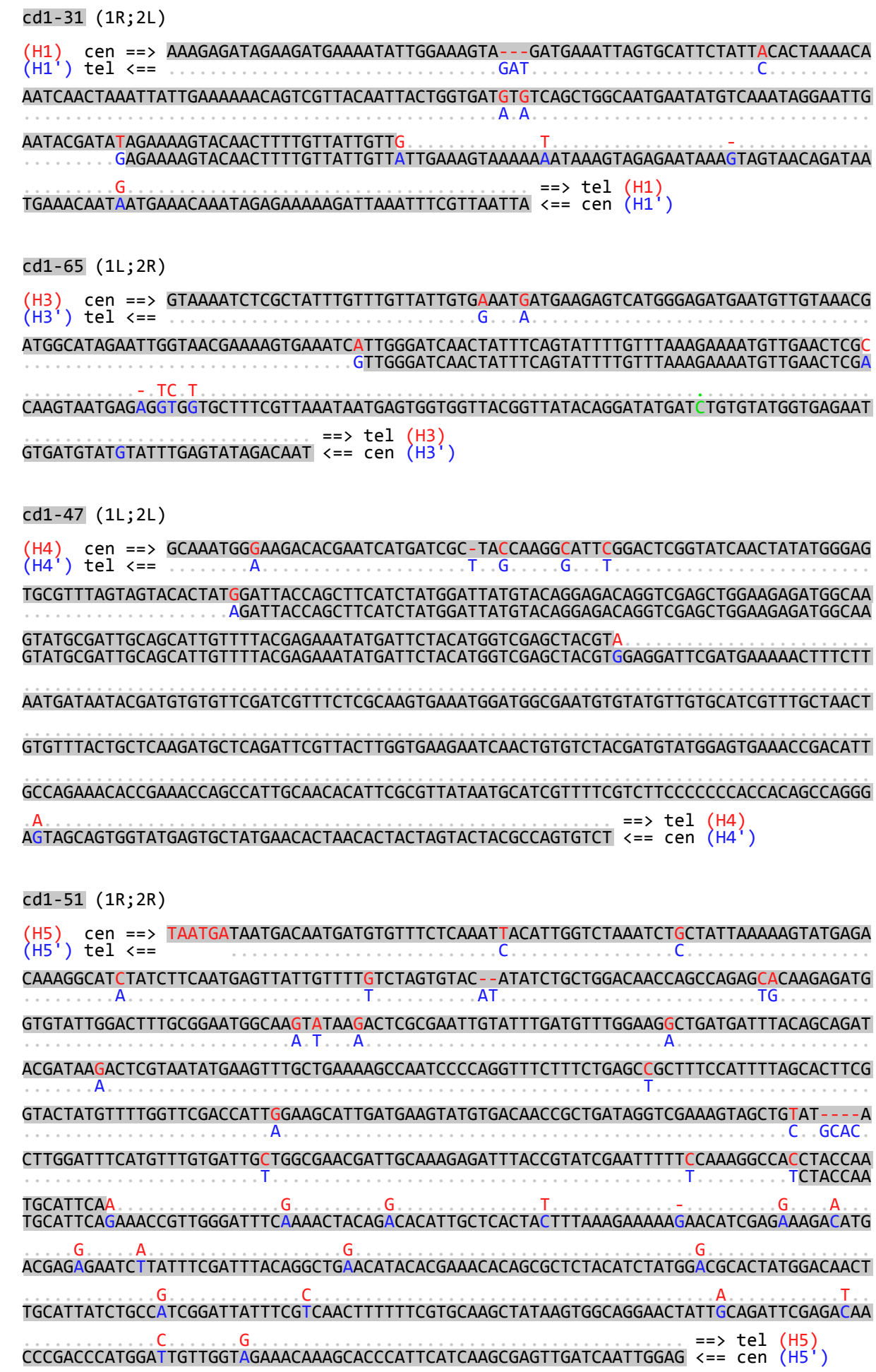


TAS2-4(LIC)

```
ATGCATCGTGGAATGGTTAAGGTGGAAAGACGAAGTAAAGTGTGTTGGGATTCATTAAGTATAATAGAGATATATAGTGTGATATGATAATTGAGAAGG
ATGAAAAATTGAAG-CTGACTCATTGATGTGTGAAGCGAGTAACGAGCCATAAGGCGAGTAATCGTTAAGCGTTGTTAAGGAGGGATAAGGAAATAAAG
GAATGAGAGTAGAAGAGGAAAGTAGTAAAGGAGAGTAGAAGAGGTAGAGTAGAAGAGGAAAGTAGAAGAGGTAGGAGTATAGAAGTAGTAAATAAGAGA
ATAAGGGAATAAAAGAGCAATAAAACGAGGAAAAGAAAGTGAAAAGAATAAAAGAGATAAAGGGATAAAGTA
GGAGAATAAAGAAGTAAGTAAAGTAAGAAAAAAGAAGTTGATTAAGTTAA-AGTATGTTGGAGTGCAGTAAGTATAATAGGAACATGCGGTGTACTTTG
ATAAATGAAGTGTTTGAAAATTGAGCTTATGTTAGATATGTTAAATAAATAATATGATTGCGTAAAATAAATAAAAAAATAAAATGAAATGAAATAAAA
TAAAATAAAATAAAATATTTAAAAACGAATAAATATATAGAAATAAATAAAATTAATGAAATAAAATTAACAAAATAAAATGAAATAAAAAATATAAAT
ACAAATATTGGAACACAAATAAAGACAATATTAAATAATGATAGAAAATAATGAAAAATATTTATTTGCCAATGAATGACAAACAGCAGGCAGTTGAAA
ATTGAAAATAAATTATAAAATGATTAATAAAAGAAAAAAAAATTATTGAAAAAGTTTGTTGAGTGATATGTCGTCGGGAAGAATGA
GTAGTGTAGTGTGGTAGTGAAGATGGACAAAACAC-----TGAAATGAGTG-GAAGT-GAGTGTGTTGGGATGCAGTAAGTATAATAAGGGGAT
GTAGTGTGCATGAGTGAATAAAACGGATGAAAAA----TTTGAAGTTGATTTGAATT-GAGTGTGCTGGAGTACGTTAAAGGTGATAGGGACAA
GTAGTGTACTATAATAATTAGGATGGTTAAAAAATA----TGAAGTTGACTCAGTTTTGATTCAGGTGGGT----------
AACGAGCAGTAAAGCGAGTAATCGTTAAACAT-GTTAACGAAATATAAAAGAGGTAAAACATAGGAAGTATAAATATTGG
ATAGGTAGAATAAGTAAAGC-AGAATAGGTAGA
----GTAGAATAGGTATAGGTAGAATAAGTAGA
GTAGAGTATGGAAGAAAGAAGAGGAAGAAAGTAATGGAATAAGAAATAATAATAAAAAATG
GGAGAACAAAGAAGTAAGTAAAGTAAGAAAAAAGAAGATACAATGAAGGGGACTATGTTGGAGTGCAGTAAGTATAATAGGAACATGCGGTGTACTTTG
ATAAATGAAGTGTTTGAAAATTGAGCTTATGTTAGTCATGTTAAACAAATAATATGATTGTGTAAAATAAATAAAAAAATAAAATGAAATGAAATGAAA
TAAAATAAAATAAAATATTTAAAAACGAATAAATATATAGAAATAAATAAAATTAATGAAATAAAATTAACAAAATAAAATGAAATAAAAAATATAAAT
ACAAATATTGGAACACAAATAAAGACAATATTAAATAATGATAGAAAATAATGAAAAATATTTATTTGCCAATGAATGACAAACAGCAGGCAGTTGAAA
ATTGAAAATAAATTATAAAATGATTAATAAAAGAAAAAAAAATTATTGAAAAAGTTTGTTGAGTGATATGTCGTCGGGAAGAATGA
GTAGTGTAGTGTGGTAGTGAAGATGGACAAAACAC-----TGAAATGAGTG-GAAGT-GAGTGTGTTGGGATGCAGTAAGTATAATAAGGGGAT
GTAGTGCAATATGGCGAATGAGGATAGATGAATGAAAACTTGAAGTTTATAGAATAA-GGCTGTGTTGTAATGCAGTAAGTTGAATAAAGATAC
GCAGTGTATTATGATAATCAAGATGGATGAAAAATA----TGAAGTTGACTCAGTTTTGATTCAGGTGGGT----------
AACGAGCAGTAAAGCGAGTAATCGTTAAACAT-GTTAACGAAATATAAAAGAGGTAAAACATAGGAAGTATAAATATTGG
ATAGGTAGAATAAGTAAAGC-AGAATAGGTAGA
ATAGGTAGAATAAGTAAAGC-AGAATAGGTAGA
GTAGAATATGGAAGAAAGAAGAGGAAGAAAGTAATGGAATAAGAAATAATAATAAAAAATG
GGAGAACAAAGAAGTAAGTAAAGTAAGAAAAAAGAAGATACAATGAAGGGGACTATGTTGGAGTGCAGTAAGTATAATAGGAACATGCGGTGTACTTTG
ATAAATGAAGTGTTTGAAAATTGAGCTTATGTTAGATATGTTAAATAAATAATATGATTGCGTAAAATAAATAAAAAAATAAAATGAAATGAAATAAAA
TAAAATAAAA-----TATTTAAAAACGAATAAATATATAGAAATAAATAAAATTAATGAAATAAAATTAACAAAATAAAATGAAATAAAAAATATAAAT
ACAAATATTGGAACACAAATAAAGACAATATTAAATAATGATAGAAAATAATGAAAAATATTTATTTGCCAATGAATGACAAACAGCAGGCAGTTGAAA
ATTGAAAATAAATTATAAAATGATTAATAAAAGAAAAAAAAATTATTGAAAAAGTTTGTTGAGTGATATGTCGTCGGGAAGAATGA
GTAGTGTAGTGTGGTAGTGAAGATGGACAAAACAC-----TGAAATGAGTG-GAAGT-GAGTGTGTTGGGATGCAGTAAGTATAATAAGGGGAT
GTAGTGCAATATGGCGAATGAGGATAGATGAATGAAAACTTGAAGTTTATAGAATAA-GGCTGTGTTGTAATGCAGTAAGTTGAATAAAGATAC
GCAGTGTATTATGATAATCAAGATGGATGAAAAATA----TGAAGTTGACTCAGTTTTGATTCAGGTGGGT----------
AACGAGCAGTAAAGCGAGTAATCGTTAAACAT-GTTAACGAAATATAAAAGAGGTAAAACATAGGAAGTATAAATATTGG
ATAGGTAGAATAAGTAAAGC-AGAATAGGTAGA
ATAGGTAGAATAAGTAAAGC-AGAATAGGTAGA
ATAGGTAGAATAAGTAAAGC-AGAATAGGTAGA
GTAGAATATGGAAGAAAGAAGAGGAAGAAAGTAATGGAATAAGAAATAATAATAAAAAATG
GGAGAACAAAGAAGTAAGTAAAGTAAGAAAAAAGAAGATACAATGAAGGGGACTATGTTGGAGTGCAGTAAGTATAATAGGAACATGCGGTGTACTTTG
ATAAATGAAGTGTTTGAAAATTGAGCTTATGTTAGTCATGTTAAACAAATAATATGATTGTGTAAAATAAATAAAAAAATAAAATGAAATGAAATAAAA
TAAAATAAAATAAAATATTTAAAAACGAATAAATATACAGAAATAAATAAAATTAA----------CAAAATAAAATGAAATAAAAAATATAAAT
ACAAATATTGGAACACAAATAAAGACAATATTAAATAATGATAGAAAATAATGAAAAATAATTATTTGCCAATGAACGACAAACAGCAGGCAGTTGAAA
ATTGAAAATAAATCATAAAATGATTGATAAAAGAAAAAAAAATTATTGAAAACGTTTGTTGAGTGATATGTCGTCGGGGAGGATGG
GAAATTTTGAGGATGGTCATTTTAATAAAGGAAAAAAATTCAAAGTTCACTTAGTCAGGTGTGAAGCGAGC
AACGAGACGTAAGGCGAGTAATCGTTAAATGTTGTTAACGAAA
GACAAGAATGTAACAGATAGAATTTAAGAGTAAAATATAGAGAGAAGAGAGTAGGATGAAGGAGCAGGATATATAAGTAGAATATATAGAAGAAATAGA
GTAAGGAGAG
TAAAATAAAATAAATAAAATAAAATAAAATAAATAAAATAAATAAAATAAATAA
GGTAGGAGAATAAAGAAGCAATTAAAGTAAGAGGAAAGAAGATAAAAGCAGGGGACTATATTGGAGTGCATTAAGTAGAATAGGGATGAGTAATGAACT
ATAGTAATTATGATAGACAAGTGGTTGAAATAATGTGTTGGGATGCAGGAATATAAATAGAGAATTGTAGTGGACTTTGGTAATTGAAGGGGATGAAAA
ATTTGAAGTTGACTCAGTCATGGTGAGCTGGATAACGAAGTAACAATATGGAATAAGGAAATGAAGGAGAGAATATAAGAGTAGGATATATAGAAGAGA
GTAGAAATCAAAAAAATAAAGTAAGGAGAG
TAAA-TTAAATATTTAAAAACGAATAAATAAAATAAAATAAAATAAAATGAATAAAATAAATAAA
GTAGGAGAAT----------GAAGAAGTAATCAAA---------GTAAGAGAGTATTAGA--------AAGTAAAGTAAA-TAAGGATGGATAC
ATAGTGTATATTGGAAAGTCAGATGCAT
```

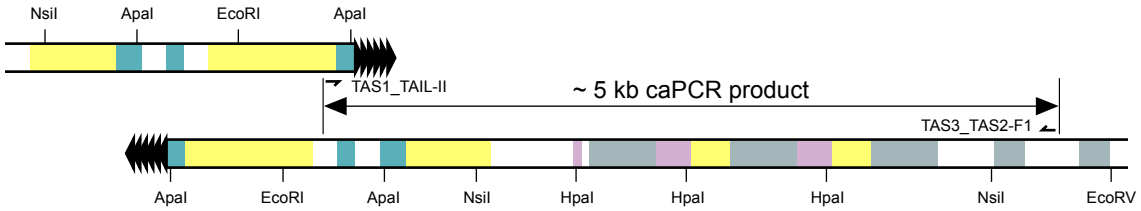
Supplementary Figure S2. The fission yeast subtelomere DNA structure. (A) Chromosomal view of the repeats located proximal to the telomeres. The nomenclature of the repeats is the same as that used in Figure 1A. The copy number of STE2 in each TAS2 is also indicated. (B) Graphical summary of the DNA structures found in the fission yeast telomeric clones in the Sanger Institute database (pNSU70, pNSU71, pNSU77, pNSU21, pNSU64, and pNSU65; available from the Sanger Institute ftp site: ftp://ftp.sanger.ac.uk/pub/yeast/sequences/pombe/telomeres/). The numbers associated with the restriction sites refer to the original numbering of the clones. The repetitive units STE1, STE1', STE2, and STE2' are indicated in yellow, cyan, grey, and lavender, respectively. STE1 is a tandem repeat of approximately 88 bp, whereas the other repeats are essentially composed of unique sequences, although STE2' harbours a direct repeat (see (D) for details). The tel repeat region is represented by a stack of filled triangles. The relative positions of TAS1 (ApaI/EcoRI), TAS2 (NsiI/NsiI), and TAS3 (EcoRV/HindIII) are also indicated for each clone. For pNSU77, the fused rDNA sequence and the break point (BP) beyond the tel repeats are indicated. (C) Sequence alignment of the STE1 repeats found in pNSU70. The numbers refer to the original numbering of the clone. (D) Sequence analyses of pNSU70, pNSU71, pNSU77, pNSU21, pNSU64, and pNSU65. The shaded sequences indicate the repetitive units of STE1, STE1', STE2, and STE2', with the same colour coding used in (B). The restriction sites shown in (B) are highlighted in magenta. The primers used for amplification of STE2 by qPCR are underlined in every STE2 repeat (to confirm their capacity to bind to each repeat with one nucleotide degeneration, as indicated by a dashed line). The reference sequence of telomeric DNA (43) takes into account the sequence of pNSU70. (E) The genomic sequence of the longest TAS2 clone. The wild-type genome was digested with NsiI, and the fragments encompassing TAS2 were cloned directly into a vector via ligation-independent cloning (LIC) technology. The structural properties found in the longest TAS2 clone (TAS2-4(LIC)) were colour-coded as described in (B).

A

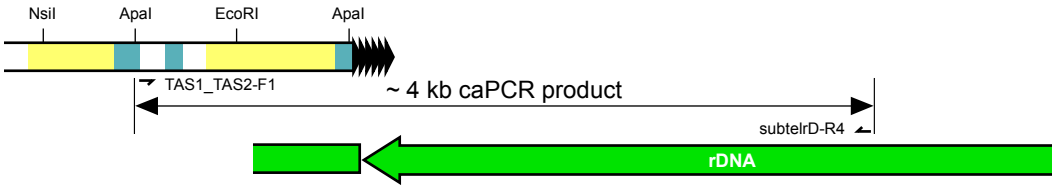

cd1-31 (1R;2L)
(H1) cen ==> AAAGAGATAGAAGATGAAAATATTGGAAAGTA---GATGAAATTAGTGCATTCTATTACACTAAAACA
(H1') tel <==
AATCAACTAAATTATTGAAAAAACAGTCGTTACAATTACTGGTGATGTGTCAGCTGGCAATGAATATGTCAAATAGGAATTG
AATACGATATAGAAAAGTACAACTTTTGTTATTGTTG
GAGAAAAGTACAACTTTTGTTATTGTTATTGAAAGTAAAAAAATAAAGTAGAGAATAAAGTAGTAACAGATAA
TGAAACAATAATGAAACAAATAGAGAAAAAGATTAAATTTCGTTAATTA <== cen (H1')
==> tel (H1)
cd1-65 (1L;2R)
(H3) cen ==> GTAAAATCTCGCTATTTGTTTGTTATTGTGAAATGATGAAGAGTCATGGGAGATGAATGTTGTAAACG
(H3') tel <==
ATGGCATAGAATTGGTAACGAAAAGTGAAATCATTGGGATCAACTATTTCAGTATTTTGTTTAAAGAAAATGTTGAACTCGC
GTTGGGATCAACTATTTCAGTATTTTGTTTAAAGAAAATGTTGAACTCGA
CAAGTAATGAGAGGTGGTGCTTTCGTTAAATAATGAGTGGTGGTTACGGTTATACAGGATATGATCTGTGTATGGTGAGAAT
GTGATGTATGTATTTGAGTATAGACAAT <== cen (H3')
==> tel (H3)
cd1-47 (1L;2L)
(H4) cen ==> GCAAATGGGAAGACACGAATCATGATCGC-TACCAAGGCATTCGGACTCGGTATCAACTATATGGGAG
(H4') tel <==
TGCGTTTAGTAGTACACTATGGATTACCAGCTTCATCTATGGATTATGTACAGGAGACAGGTCGAGCTGGAAGAGATGGCAA
AGATTACCAGCTTCATCTATGGATTATGTACAGGAGACAGGTCGAGCTGGAAGAGATGGCAA
GTATGCGATTGCAGCATTGTTTTACGAGAAATATGATTCTACATGGTCGAGCTACGTA
GTATGCGATTGCAGCATTGTTTTACGAGAAATATGATTCTACATGGTCGAGCTACGTGGAGGATTCGATGAAAAACTTTCTT
AATGATAATACGATGTGTGTTCGATCGTTTCTCGCAAGTGAAATGGATGGCGAATGTGTATGTTGTGCATCGTTTGCTAACT
GTGTTTACTGCTCAAGATGCTCAGATTCGTTACTTGGTGAAGAATCAACTGTGTCTACGATGTATGGAGTGAAACCGACATT
GCCAGAAACACCGAAACCAGCCATTGCAACACATTCGCGTTATAATGCATCGTTTTCGTCTTCCCCCCCACCACAGCCAGGG
AGTAGCAGTGGTATGAGTGCTATGAACACTAACACTACTAGTACTACGCCAGTGTCT <== cen (H4')
==> tel (H4)
cd1-51 (1R;2R)
(H5) cen ==> TAATGATAATGACAATGATGTGTTTCTCAAATTACATTGGTCTAAATCTGCTATTAAAAAGTATGAGA
(H5') tel <==
CAAAGGCATCTATCTTCAATGAGTTATTGTTTTGTCTAGTGTAC--ATATCTGCTGGACAACCAGCCAGAGCACAAGAGATG
GTGTATTGGACTTTGCGGAATGGCAAGTATAAGACTCGCGAATTGTATTTGATGTTTGGAAGGCTGATGATTTACAGCAGAT
ACGATAAGACTCGTAATATGAAGTTTGCTGAAAAGCCAATCCCCAGGTTTCTTTCTGAGCCGCTTTCCATTTTAGCACTTCG
GTACTATGTTTTGGTTCGACCATTGGAAGCATTGATGAAGTATGTGACAACCGCTGATAGGTCGAAAGTAGCTGTAT----A
CTTGGATTTCATGTTTGTGATTGCTGGCGAACGATTGCAAAGAGATTTACCGTATCGAATTTTTCCAAAGGCCACCTACCAA
TGCATTCAA G G T - G A
TGCATTCAGAAACCGTTGGGATTTCAAAACTACAGACACATTGCTCACTACTTTAAAGAAAAAGAACATCGAGAAAGACATG
ACGAGAGAATCTTATTTCGATTTACAGGCTGAACATACACGAAACACAGCGCTCTACATCTATGGACGCACTATGGACAACT
TGCATTATCTGCCATCGGATTATTTCGTCAACTTTTTTCGTGCAAGCTATAAGTGGCAGGAACTATTGCAGATTCGAGACAA
CCCGACCCATGGATTGTTGGTAGAAACAAAGCACCCATTCATCAAGCGAGTTGATCAATTGGAG <== cen (H5')
==> tel (H5)

B

cd1-98 (1L:2L)

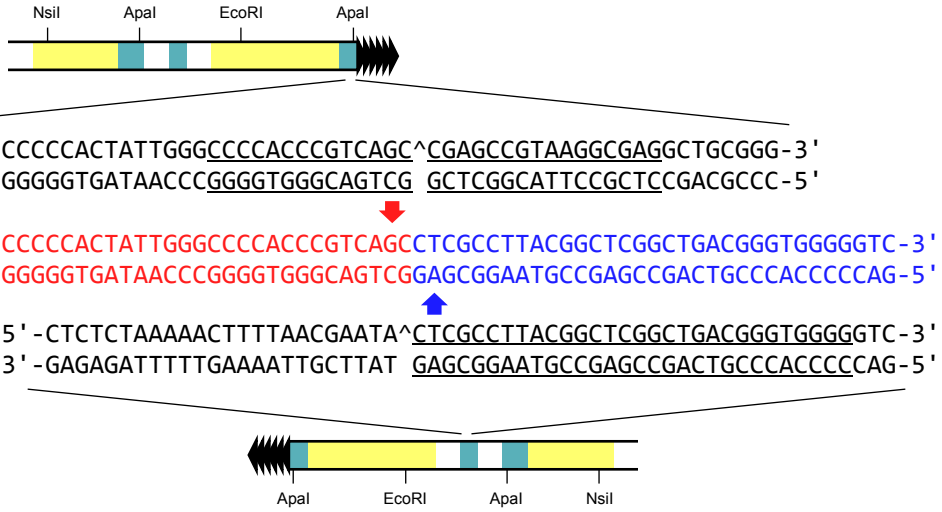


cd1-53, cd1-63 (1R:3L)

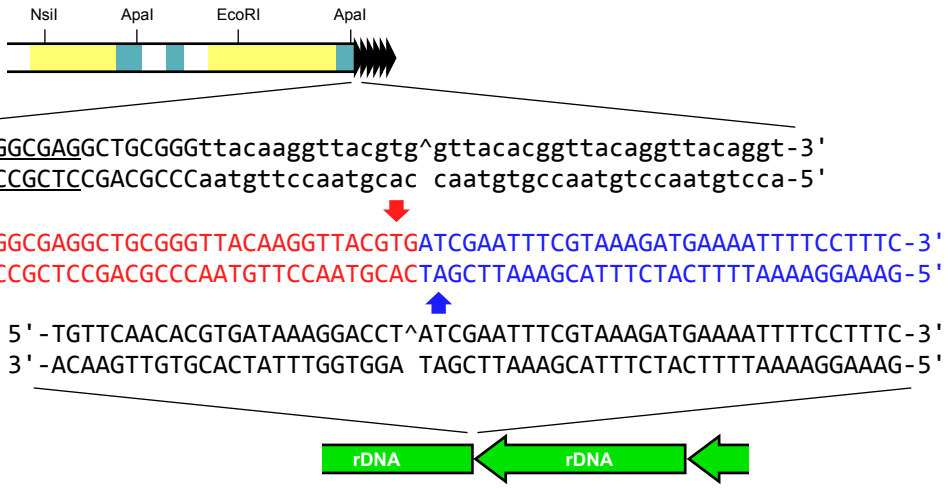


C

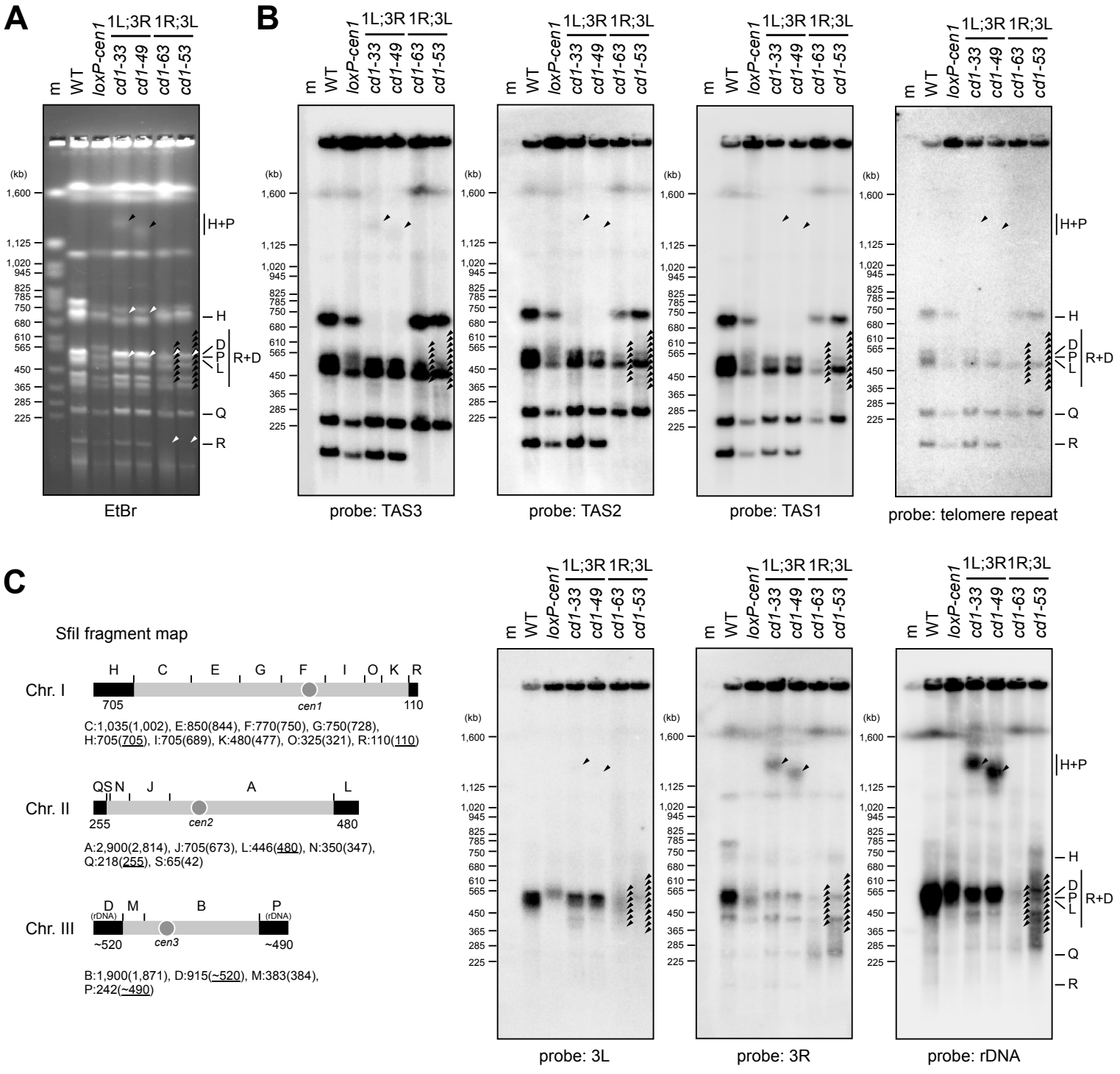
cd1-98 (1L:2L)



cd1-53, cd1-63 (1R:3L)

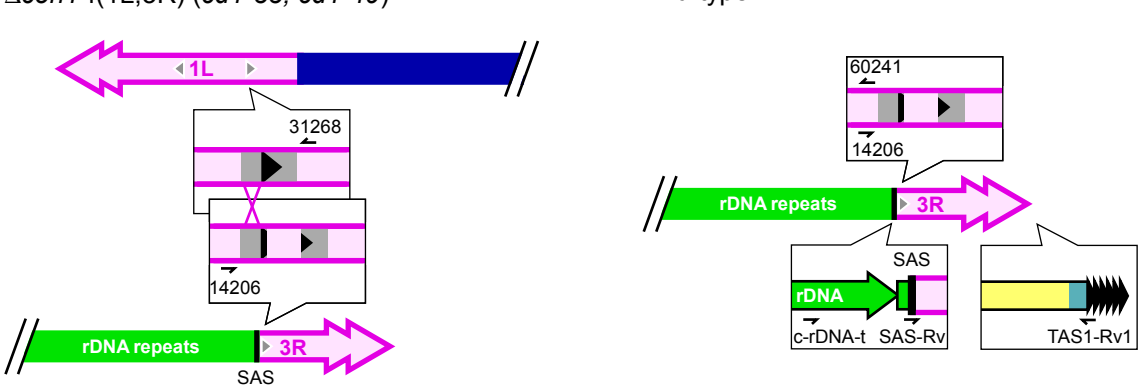


Supplementary Figure S3. Sequence analysis of the chromosome fusion junction. (A) Alignment of the inversely homologous subtelomere sequences and the actual fusion sequences of *cd1-31*, *cd1-65*, *cd1-47*, and *cd1-51* (grey-shaded text). Nucleotide degenerations in the original subtelomere sequences are highlighted in red and blue. The actual fusion sequences were determined by sequencing the gPCR products. The site of crossover in the fusogenic sequence can be confined to a limited segment according to the choice of degenerated nucleotides, which is indicated by the overlapping grey-shaded text. The nomenclature of the homology segments (H1–H5 and H1'–H5') follows that described by Wang and Baumann (43). (B) Schematic diagram of the NHEJ-type fusion (*cd1-98*, *cd1-53*, and *cd1-63*). Colour coding of the DNA structures is the same as that used in Supplementary Figure S2B. The primers sets used for the gPCR analyses of the *Δcen1*-f rearrangements are also shown. (C) Experimentally determined junction sequences. Subcloning and subsequent sequencing of the gPCR products revealed that all of the fusion events were attributable to the canonical NHEJ pathway.


A
B
C
m
WT
loxP-cen1
1L;3R
1R;3L
cd1-33
cd1-49
cd1-63
cd1-53
(kb)
1,600
1,125
1,020
945
825
785
750
680
610
565
450
365
285
225
H+P
H
D
P
L
R+D
Q
R
EtBr
probe: TAS3
probe: TAS2
probe: TAS1
probe: telomere repeat
SfiI fragment map
Chr. I
H C E G F I O K R
705
cen1
110
C:1,035(1,002), E:850(844), F:770(750), G:750(728),
H:705(705), I:705(689), K:480(477), O:325(321), R:110(110)
Chr. II
Q S N J A L
255
cen2
480
A:2,900(2,814), J:705(673), L:446(480), N:350(347),
Q:218(255), S:65(42)
Chr. III
D (rDNA) M B P (rDNA)
~520
cen3
~490
B:1,900(1,871), D:915(~520), M:383(384),
P:242(~490)
probe: 3L
probe: 3R
probe: rDNA

Supplementary Figure S4. SfiI digestion of chromosome III-related fusion survivors. (A, B) PFGE analyses of SfiI-digested chromosomes in the *Δcen1*-f(1;3) survivors. The gel was subjected to EtBr staining (A) and Southern blotting with the indicated probes (B). The white arrowheads indicate the SfiI bands that disappeared in the survivors due to fusion, and the black arrowheads indicate the newly-generated fusion bands. The identities of the bands with altered expression levels are shown at the right-hand side of the gels. The nomenclature of the SfiI fragments follows that described by Fan et al. (64) (see (C)). Because the rDNA repeat length was variable, particularly in the survivors, the band intensity of the SfiI fragment containing rDNA was low on some occasions. m, molecular size marker. (C) Schematic diagram of SfiI fragments localised along the fission yeast chromosomes, with emphasis on the terminal fragments. The SfiI-digested fragment sizes reported by Fan et al. (64) are indicated below each chromosome. The values in parentheses correspond to the fragment sizes calculated from fission yeast whole-genome sequences. The underlined values in parentheses correspond to terminal fragments that were not precisely defined in the database; thus the experimentally determined values from our wild-type laboratory yeast strain and the values described by Fan et al. (64) are shown.

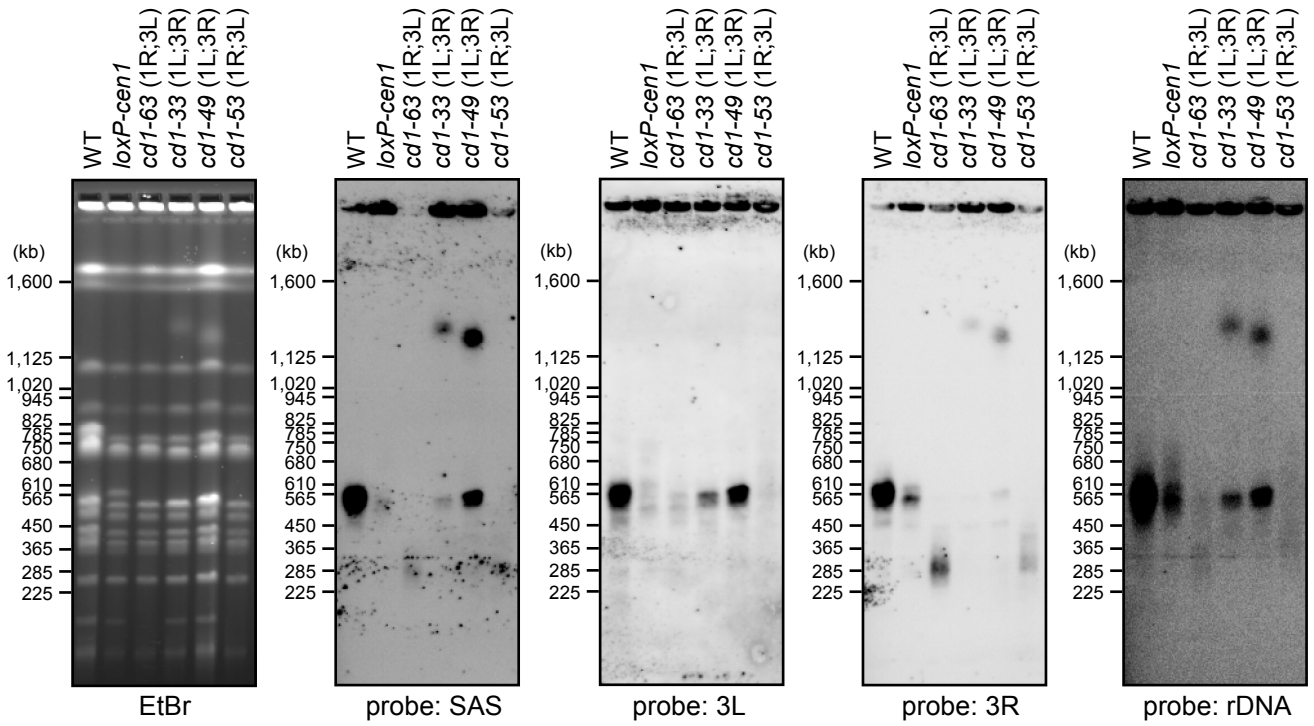
A



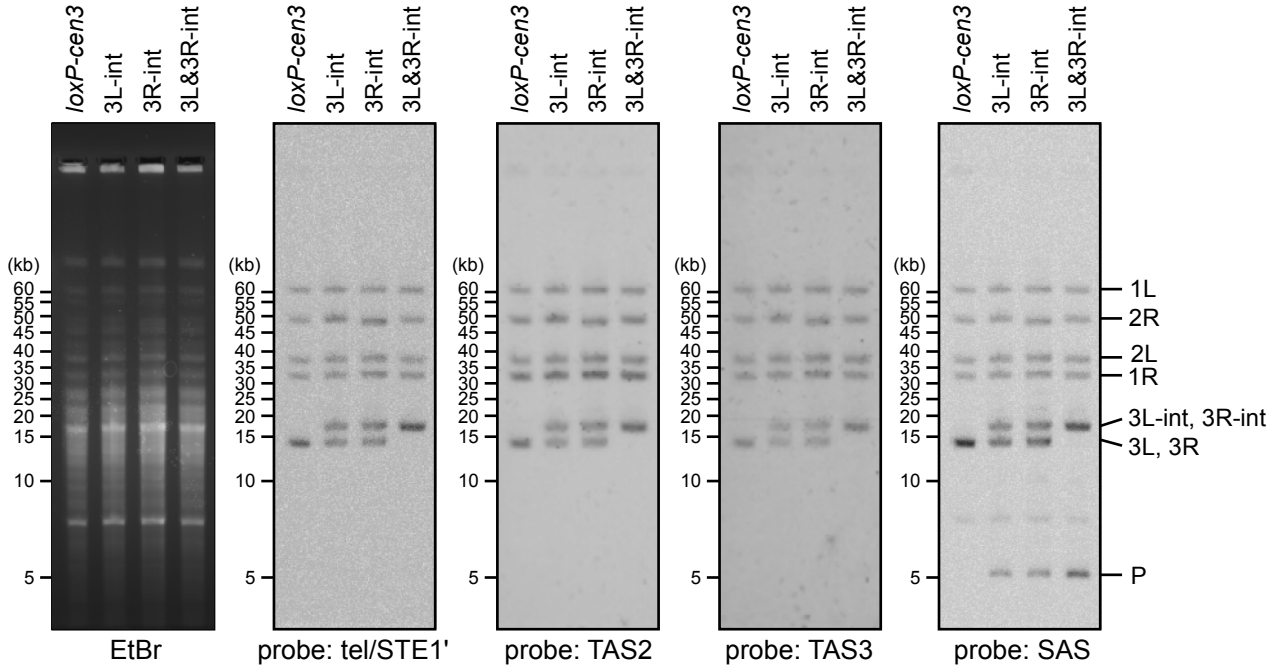
B

```
(rDNA) ==> TAGAGGTGGAGATGGGTAGGATTGGATGGTTTGTGAGATGGGGAAATGACGTAAGAATTGGAATTTAACGAGAGAGAACGAGTCTATCTCGGTTGTTTTGG
CATGTCCTATCTCAGTTATGAATAGGGAGAGTTAGGGATAGAGAACGTATTTCGATGAGAATCAACGATGAGTAACGAGAGGGAATGAAATGGAACGAGAACAGACCAGAGT
CATCGTTTCAATGCAATGCCATGCCATGTCAATGCCATCTCACTTACGAATAACGTAGAGTAAGGAAAAAAGGACATCTCAGTTATGAAATGAGTCGTATCTCACTTATGAA
TAAGGAGGAGTAGGGAAAAAGGAAAAAATAAAAATAAAAAAAAATAAATAAATTAAAATGAATAAAAAATAAAAAATAAAAAATAAAAATAAAAATAAAAATAAAAATAAAT
TAAAATGAATAAAAAAATTGTTTATTCAATGAAATTGTTTAAATAAATATAAAATTACAAATTTCGAAAAAAAACAATTGTATTACTTTTCTTTTTTTTTTTCATTGCAATA
AAATGGAATTGTTTGGTTACACAACTTATTATTTCATGAACCTTTAAAAAAATTTGTTGAATAAAAAGAAGAAAAGAAAAGAAAAGAAATCATTTTAATTTTATCAAATTCG
CTATACATAAAGATTTTAGAGAGAAATGGATTATCTTTTGGTATTCAAAGTAATAGGAAGGGTGAGGCAAAACCCTTTGGTGGGAAGTGGATTTTTCTGTAGTACTGCGTAT
TGCGTTCGACAAATGTCAAACCCTACCCTAGAAGCACCGATATGGATGAAATTCACAAAGTAGCACCACTTTCTACAAAATCGAGGAAAACTAGGTCAAAGGTAGTATGTAA
AGGGGTAGAGGGAGGAGAATAAAAAAAGTTTCAAGTCGAAACTTGTCAAGTTGAAAAAAAAGTTTCAATCTTTTTATACTTGGATGCTCTTTGACATGGCTGTCATTTCTGT
CATTTCGTATGGCTGGTCTATTTCACTTTTTAGGCTGGTTCAAACAGGGTTTGATAAGCGATGGCTGACTGGGCTGGCTACTGTTTTACATTCGTTATATAAATGCAAACTT
TTTTTATTTGTTGATACATTATTTATCCCTTCATTTTGTTGTTTTCACATTTCTACTGCC ==> (subtelomere)
```

C

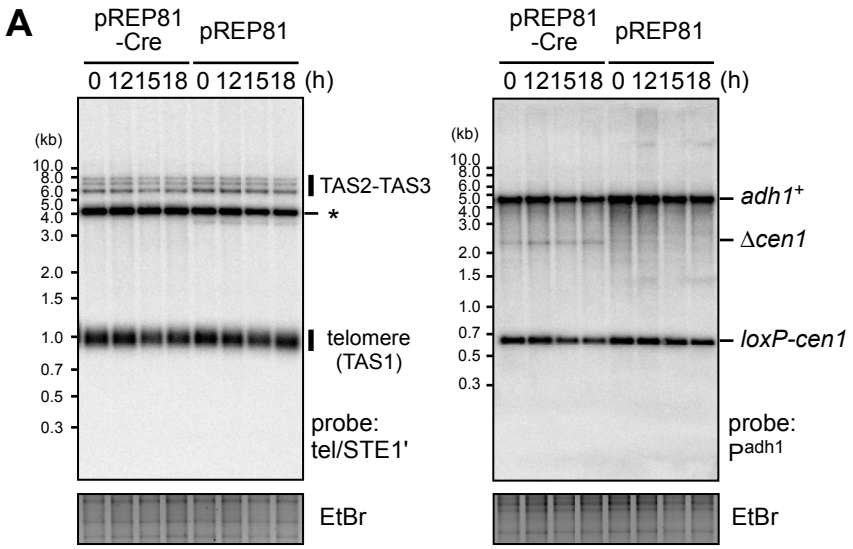


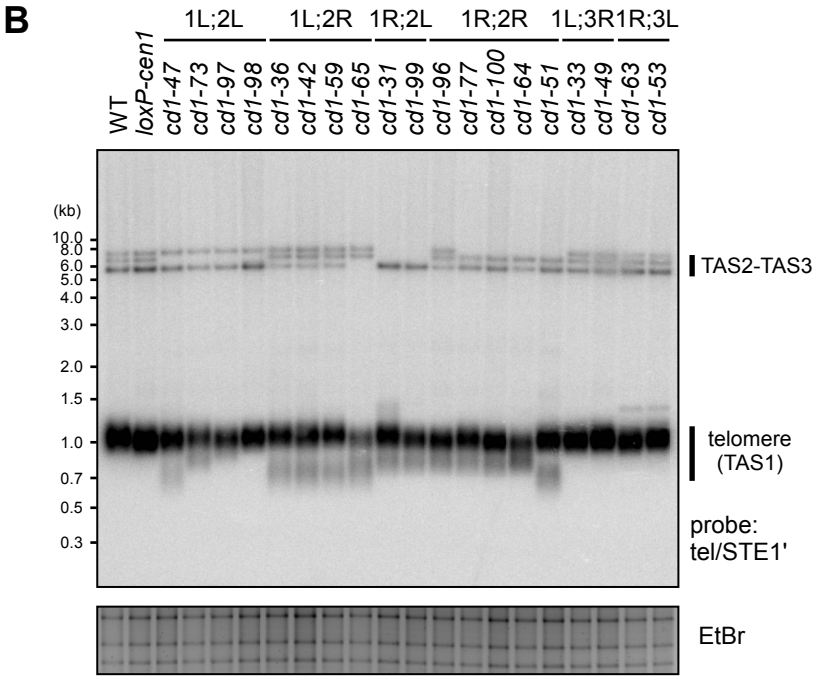
D



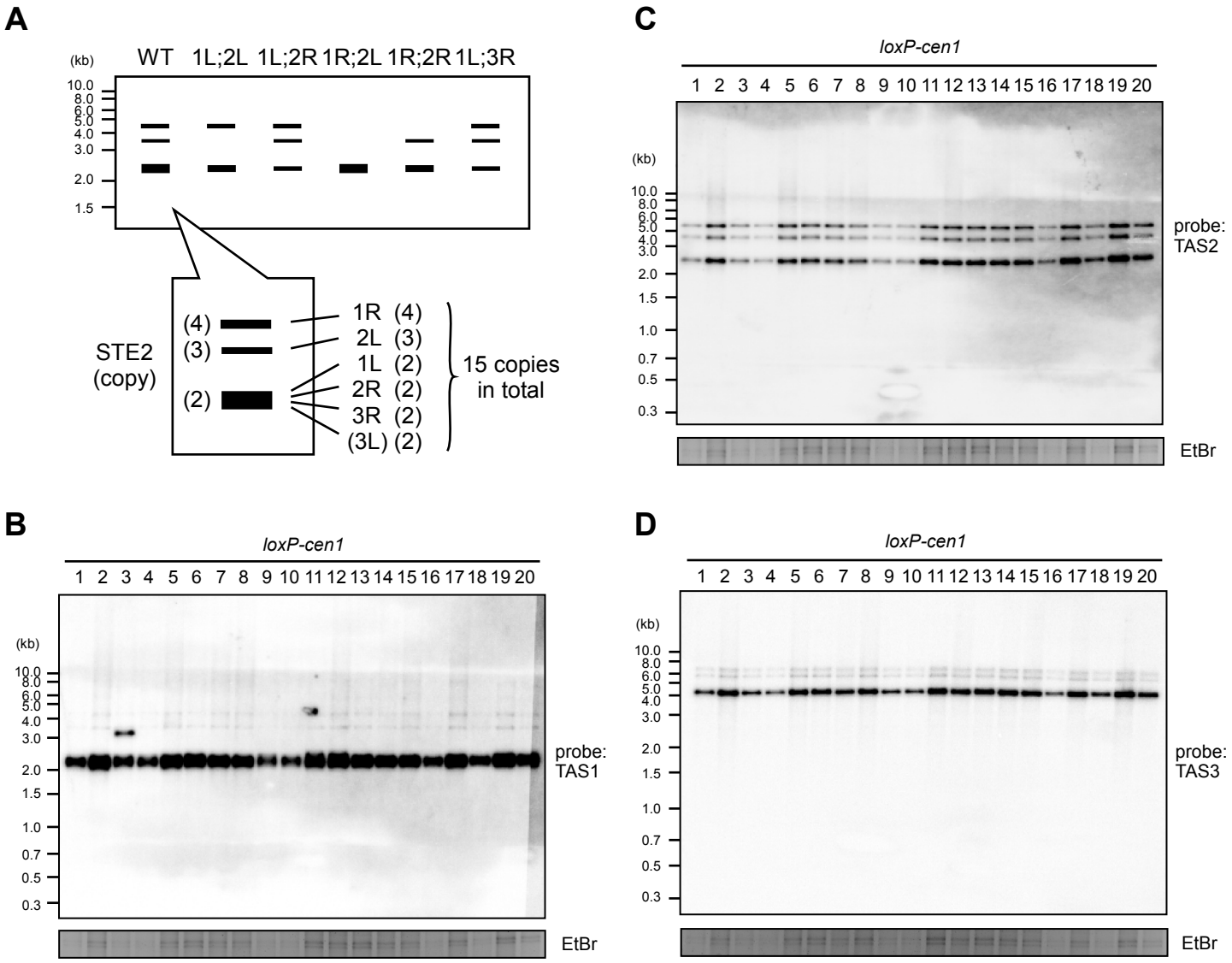
Supplementary Figure S5. Identification of a novel sequence (SAS) located between the rDNA repeats and the subtelomere region of chromosome III. (A) Overview of the gPCR analyses. The left panel shows a gPCR analysis using primers 31268 and 14206 to identify the fusion point in the *Δcen1*-f(1;3) survivors (*cd1-33* and *cd1-49*). The success of the gPCR analysis suggested the existence of a 31268- or 14206-corresponding sequence at the terminal region of 3R. The right panel shows a gPCR analysis of the wild-type genome using primer 60241, which was complementary to 14206, together with a series of rDNA primers that hybridised in the outward orientation. One of these primers, c-rDNA-t, yielded a gPCR product containing a novel 1169 bp sequence (SAS) located between the rDNA repeats and the subtelomere sequence. The end-adjacency of SAS was confirmed further by a gPCR analysis with an SAS-specific primer (SAS-Rv) and a terminal primer (TAS1-Rv1). (B) Sequence analysis of SAS. The SAS sequence showed 89% homology with a newly identified sequence at a chromosome BP of the *S. pombe* isolate strain CBS2777 (54). It also exhibited weak homology to STE2 (81% identity over 54 bp), possibly causing cross-hybridisation with TAS2 in some Southern blot analyses (see (D)). (C) The existence of the SAS at both ends of chromosome III, as confirmed by Southern blotting of SfiI-digested chromosomes. A pair of chromosome III terminal fragments in our wild-type laboratory yeast strain migrated at almost the same position in PFGE experiments, making it difficult to assess the existence of the SAS. However, fusogenic rearrangement in the *Δcen1*-f(1;3) survivors (*cd1-33* and *cd1-49*) altered the migration position of one of the fragments. The observation of two SAS hybridisation signals in these strains effectively localised the SAS at both ends of chromosome III. (D) The presence of subtelomere repeats at the SAS-distal ends of 3R and 3L. A plasmid containing a BamHI site was integrated into one or both SASs at 3R and 3L in the *loxP-cen3* strain (40). To confirm the presence of TAS3, TAS2, TAS1 (STE1'), and tel repeats in the SAS-containing fragments, BamHI-digested chromosomes from the resulting integrants were subjected to PFGE followed by Southern blotting. The identities of the other subtelomeric bands were determined by PFGE and Southern blot analyses of the integrants in which TAS3 of any of the chromosomal ends could be distinguished by a newly created BamHI site (data not shown). P, plasmid-sized genomic fragment generated due to the sequential integration of multiple plasmids into the target locus of the host genome.

Supplementary Figure S6





Supplementary Figure S6. The telomere length remains constant during the induction of centromere deletion. (A) Southern blot analyses of genomic DNAs recovered from *loxP-cen1* cells harbouring a Cre-inducible plasmid (pREP81-Cre) or empty vector (pREP81) (42). The DNAs were digested with EcoRI (left), or BsiWI plus BssHII (right), at the indicated time (h) after induction. Southern blotting was performed using probes encompassing both the tel repeats and the STE1' repeats (tel/STE1', left), or the *adh1*$^{+}$ promoter region (P^{adh1}, right). EtBr-stained images of the gels are shown as loading controls. The TAS2-TAS3 bands in the left-hand gel indicate signals that were most likely derived from the STE1' repeats located between TAS1 and TAS2 (see Supplementary Figure S1). The asterisk indicates a non-specific hybridisation signal that was not reproducible. The *loxP-cen1* signals in the right-hand gel were derived from the intact *cen1* genome, whereas the *Δcen1* signals correspond to *cen1*-excised DNA. The *adh1*$^{+}$ signals were derived from the endogenous *adh1*$^{+}$ gene, which remained at constant levels. By contrast, decreased *loxP-cen1* and increased *Δcen1* signals were observed upon Cre induction. (B) Southern blot analyses of genomic DNAs recovered from the indicated telomere-fusion survivors. The experiment was performed as described in (A). Variations in the TAS2-TAS3 signal intensity are representative of subtelomere instability in the survivors. Alterations in the telomere signal were stable and specific to a given chromosome; hence, they most likely reflect instability of the TAS1 repeat rather than the tel repeat.

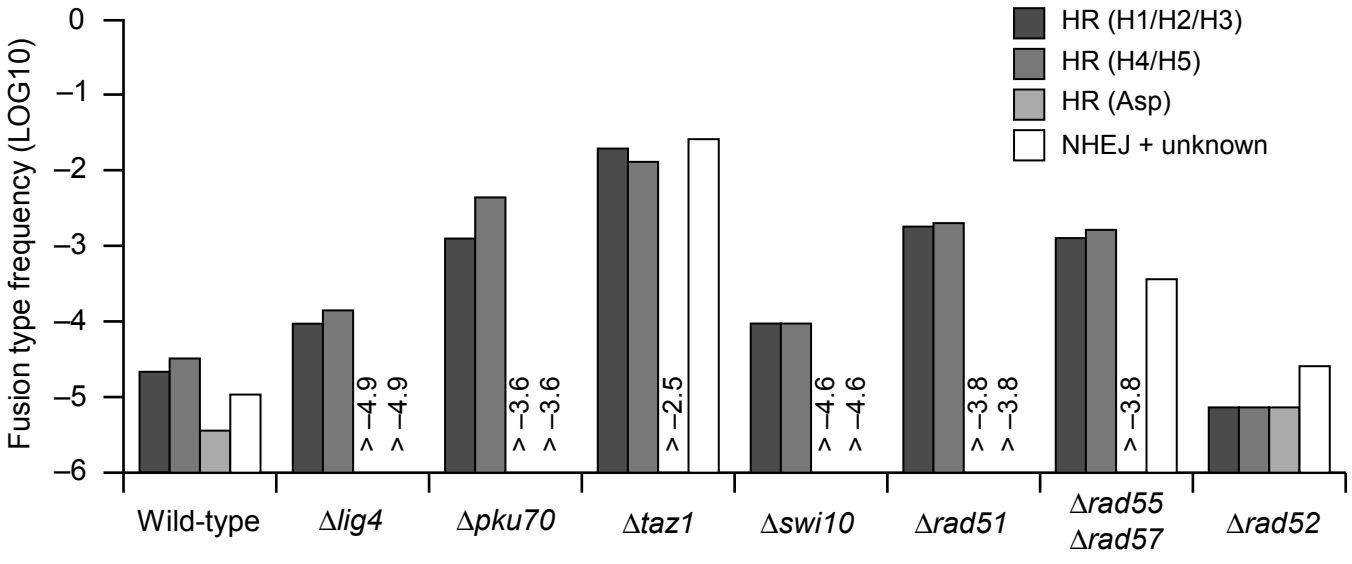

A
WT 1L;2L 1L;2R 1R;2L 1R;2R 1L;3R
(kb)
10.0
8.0
6.0
5.0
4.0
3.0
2.0
1.5
STE2
(copy)
(4)
(3)
(2)
1R (4)
2L (3)
1L (2)
2R (2)
3R (2)
(3L) (2)
15 copies
in total
B
loxP-cen1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
(kb)
10.0
8.0
6.0
5.0
4.0
3.0
2.0
1.5
1.0
0.7
0.5
0.3
probe:
TAS1
EtBr
C
loxP-cen1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
(kb)
10.0
8.0
6.0
5.0
4.0
3.0
2.0
1.5
1.0
0.7
0.5
0.3
probe:
TAS2
EtBr
D
loxP-cen1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
(kb)
10.0
8.0
6.0
5.0
4.0
3.0
2.0
1.5
1.0
0.7
0.5
0.3
probe:
TAS3
EtBr

Supplementary Figure S7. The wild-type subtelomere DNA structure. (A) Schematic diagram of representative TAS2 hybridisation patterns in wild-type cells and the indicated fusion strains. The original data refer to (C) and Figure 3B. The identity of each band was deduced based on the disappearance of the common band(s) in assorted fusion survivors, as shown in the balloon. The *STE2* copy numbers are also indicated in parentheses. For rearrangements involving 3L, only NHEJ-type events were observed; therefore, the STE2 copy number could not be deduced accurately from the TAS2 Southern blot. The identity of the 3L band was determined based on the elimination of the other identified bands. (B–D) Southern blot analyses of genomic DNAs recovered from 20 independent *loxP-cen1* strain isolates of various ages. The DNAs were digested with NsiI and then subjected to Southern blotting with the subtelomeric TAS1 (B), TAS2 (C), and TAS3 (D) probes. EtBr-stained images of the gels are shown as loading controls. Almost no structural variation was observed between the *loxP-cen1* isolates, with the exception of an additional TAS1 band in isolate #3.

A

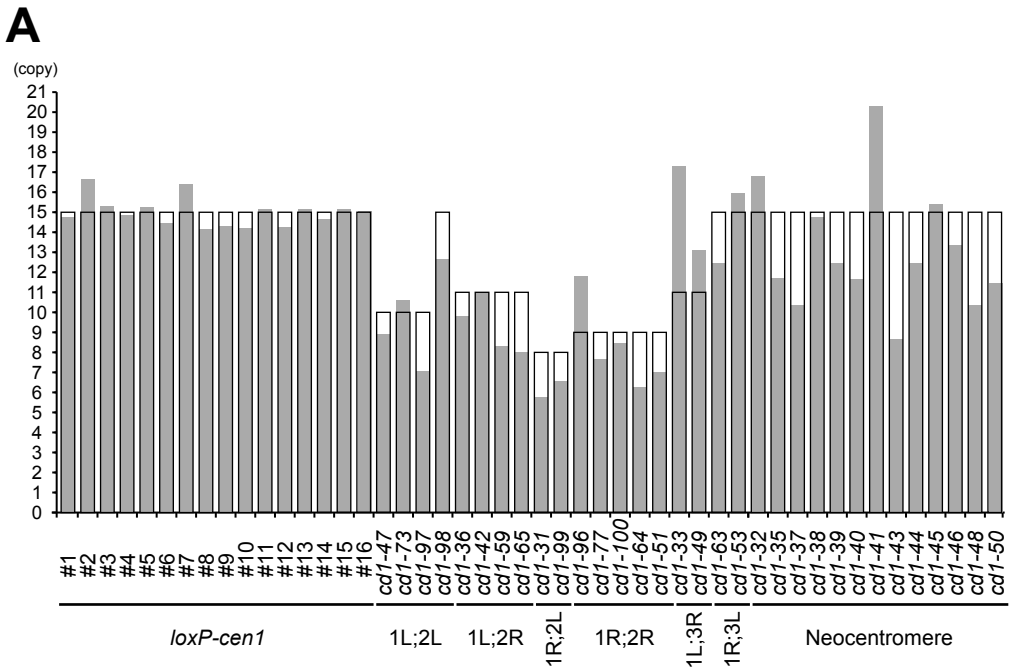
	class	I	II	III	IV	V
	PFGE-Southern blot	tel(–) TAS1(–) TAS2(–) TAS3(+)	tel(–) TAS1(–) TAS2(–) TAS3(+)	tel(–) TAS1(–) TAS2(–) TAS3(–)	tel(–) TAS1(–) TAS2(–) TAS3(–)	tel(+/–) TAS1(+) TAS2(+) TAS3(+)
	gPCR	H1/H2/H3	H4/H5	Asp	–	–
Wild-type	19	6 (31.6)	9 (47.3)	1 (5.3)	0 (>5.0)	3 (15.8)
Δ*lig4*	5	2 (40.0)	3 (60.0)	0 (>16.6)	0 (>16.6)	0 (>16.6)
Δ*pku70*	18	4 (22.2)	14 (77.8)	0 (>5.2)	0 (>5.2)	0 (>5.2)
Δ*taz1*	18	6 (33.3)	4 (22.2)	0 (>5.2)	0 (>5.2)	8 (44.4)
Δ*swi10*	6	3 (50.0)	3 (50.0)	0 (>14.2)	0 (>14.2)	0 (>14.2)
Δ*rad51*	19	9 (47.4)	10 (52.6)	0 (>5.0)	0 (>5.0)	0 (>5.0)
Δ*rad55*Δ*rad57*	18	7 (38.9)	9 (50.0)	0 (>5.2)	2 (11.1)	0 (>5.2)
Δ*rad52*	13	2 (15.4)	2 (15.4)	2 (15.4)	7 (53.8)	0 (>7.1)

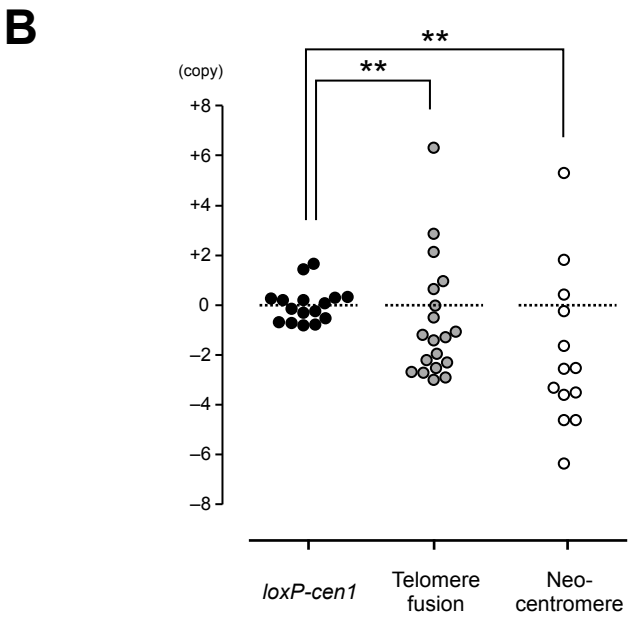
B



Supplementary Figure S8. Fusion spectrum analysis. (A) Classification of the *Δcen1*-f survivors in the indicated strain backgrounds. The five different types of classification (I–V) were based on the results of PFGE-Southern blot and gPCR analyses. The numbers in parentheses represent the percentage of each class of survivors in respective strains. Classes I–III are the HR-type and class V is the NHEJ-type; with the exception of those in the wild-type, *Δrad51*, and *Δrad52* backgrounds, the precise fusion points have not been determined. The identity of class IV remains unclear. (B) The calculated frequency of each fusion type in the indicated strain backgrounds. The calculations were based on the *Δcen1*-f survivor frequencies and fusion spectrum. The frequencies of classes IV and V were combined and are indicated as “NHEJ + unknown”.

Supplementary Figure S9





Supplementary Figure S9. TAS2 instability occurs also in neocentromere survivors. (A) The predicted and experimentally determined STE2 copy numbers in the *loxP-cen1* clones, the telomere-fusion survivors, and the neocentromere survivors (shown as described in Figure 3D). Like the telomere-fusion survivors, TAS2 was also destabilised in the neocentromere survivors. (B) Scatter plot showing the differences between the predicted and experimentally determined STE2 copy numbers in the strains shown in (A). $^{**}P < 0.01$ by a Welch's two-tailed t-test.