The adaptive molecular landscape of reprogrammed telomeric sequences

Melania D'Angiolo[1,#,*], Benjamin P. Barré[1], Sakshi Khaiwal[1], Julia Muenzner[2], Johan Hallin[3], Matteo De Chiara[1], Nicolò Tellini[1], Jonas Warringer[3], Markus Ralser[2,4,5,6], Eric Gilson[1,7], Gianni Liti[1,*]

[1]Université Côte d'Azur, CNRS, INSERM, IRCAN, Nice, France; [2]Charité Universitätsmedizin Berlin, Department of Biochemistry, 10117 Berlin, Germany; [3]Department of Chemistry and Molecular Biology, University of Gothenburg, Gothenburg, Sweden; [4]The Francis Crick Institute, Molecular Biology of Metabolism Laboratory, London NW1 1AT, United Kingdom; [5]The Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7BN, United Kingdom; [6]Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany; [7]Department of Genetics, IHU RespirERA, CHU, Nice, France. [#]Current address: Institute of Healthy Ageing, Department of Genetics, Evolution & Environment, University College London, London, United Kingdom.

* to whom correspondence should be addressed: M.D. m.angiolo@ucl.ac.uk and G.L. gianni.liti@cnrs.fr

Abstract

Telomeric sequences vary across the tree of life and intimately co-evolve with telomere-binding protein complexes. However, the molecular mechanisms allowing organisms to adapt to new telomeric sequences are difficult to gauge from extant species. Here, we reprogrammed multiple yeast lines to human-like telomeric repeats to unveil their molecular and fitness response to novel telomeres. Initially, the exchange of telomere sequences resulted in genome instability, proteome remodelling and severe fitness decline. However, adaptive evolution experiments selected for repeated mutations that drove adaptation to the humanized telomeres. These consisted of the recurrent amplification of the telomere-binding protein *TBF1*, by complex aneuploidies, or in repeated mutations that attenuate the DNA damage response. Overall, our results outline a response that defines the adaptive molecular landscape to novel telomeric sequences.

2

Introduction

The correct maintenance of the genome is of fundamental importance in order to prevent instability and degeneration. In most eukaryotes, the genome is organized in linear segments whose extremities are exposed to degradation and need to be protected. This function is performed by ribonucleoproteins called “telomeres”, consisting of a repeated DNA sequence and its associated proteins (Greider & Blackburn, 1987). Telomeres are maintained via the specialised enzyme telomerase. This enzyme is a reverse transcriptase that synthesizes new repeats using a RNA moiety as template (Chan & Blackburn, 2002; De Lange, 2005; Greider & Blackburn, 1989; Lundblad & Szostak, 1989; Marcand et al., 1997).

Despite variations in their DNA sequence having occurred during evolution, the function of telomeres in genome maintenance is remarkably conserved across the tree of life. In fact, all telomeric DNA motifs share common characteristics, like being short, tandemly-repeated and G-rich, and the principles of telomere biology are widely conserved even among distant taxa (Kupiec, 2014). Nevertheless, variations in both telomeric sequence and length occur (Lyčka et al., 2024; Whittemore et al., 2019). The vertebrate-like TTAGGG (T_2AG_3) repeat is the most widespread, being present in the majority of Metazoa, but also in some classes of protists, fungi and plants, and is thought to be the ancestral unit from which all the others evolved (Fulnečková et al., 2013). Variations of this motif are observed in other taxa, like TTTAGGG in plants (Peska & Garcia, 2020), TTAGG in insects (Kuznetsova et al., 2020), TTAGGC in nematodes (Zetka & Müller, 1996) and degenerate motifs in *Saccharomyces* yeasts (Teixeira & Gilson, 2005). In other organisms, telomeric repeat units can be very long and complex, like in species of the genus *Candida* or in fission yeasts (Sepsiova et al., 2016; Peska et al., 2021). Therefore, telomeric DNA sequences have changed at multiple times during evolution, but how organisms subsequently adapt to accomodate and capitalize on these changes remains uncharacterised.

Telomeric DNA varies in length with a difference of several orders of magnitude across species and extensive within-species variation. Telomere length ranges from 20 bp in the ciliate *Oxytricha trifallax* (Cavalcanti et al., 2004), 150-600 bp in budding yeasts (D’Angiolo et al., 2023; Garrido et al., 2025; O’Donnell et al., 2023), to 5-15 kb in humans (Karimian et al., 2024) and over 100 kb in some mice species (Hemann & Greider, 2000). The budding yeast *Saccharomyces cerevisiae* has been a leading model in telomere biology. *S. cerevisiae* telomeres are constituted by degenerate TG_{1-3} repeats (Wellinger & Zakian, 2012). Moreover, yeast telomeres are flanked by repetitive subtelomeric DNA like the X

61 and Y′ elements, and interstitial telomeric sequences (ITS) that are often located between
62 them, adding further complexity (Louis, 1995; Wellinger & Zakian, 2012). *S. cerevisiae*
63 cells can survive for several cell divisions after the complete loss of telomerase, and often
64 generate stable survivors that can replicate indefinitely thanks to recombination-based
65 maintenance of telomeres (Lundblad & Blackburn, 1993).
66 It is possible to generate yeasts with human-like T_2AG_3 telomeres by engineering the
67 telomerase RNA template encoded by the gene *TLC1* (Henning et al., 1998; Forstemann
68 & Lingner, 2001). Such a humanization of yeast telomeres entails a switch in telomere-
69 binding proteins, in which Rap1p, originally bound to native TG_{1-3} repeats, is replaced by
70 Tbf1p, a regulatory factor known to bind TTAGGG repeats in the yeast genome (Brigati et
71 al., 1993; Koering et al., 2000; Berthiau et al., 2006). Humanized telomeres are
72 dysfunctional and show deregulation of their length and loss of subtelomeric transcriptional
73 silencing. Furthermore, they trigger the activation of the DNA damage response (DDR)
74 with consequent cell cycle arrest (Alexander & Zakian, 2003; Auriche et al., 2008; Bah et
75 al., 2004; Bah et al., 2011; Brevet et al., 2003; Di Domenico et al., 2009, 2013; Fukunaga
76 et al., 2012; Ribaud et al., 2012).
77 In this study, we investigated the evolutionary response of yeast cells to a sudden
78 reprogramming of all telomeres by propagating them through experimental evolution. We
79 sequentially combined two experimental evolution paradigms: first, we propagated cells
80 over many mitotic generations through mutation accumulation lines (MAL) to minimize
81 selection and study the long-term effects of telomere reprogramming. Next, we submitted
82 MAL to adaptive evolution (AEL) by repeated serial transfers of large populations in which
83 multiple clones are competing against each other, to map mutations that counteract the
84 fitness decline. The latter procedure ensures that if some clones develop a beneficial
85 mutation, they will outcompete the others and lead the mutation to fixation. During MAL,
86 yeasts with humanized telomers showed a progressive fitness decay accompanied by
87 genome instability and partial proteomic remodelling, but recovered near-wild-type fitness
88 after AEL by amplifying *TBF1* or modulating the DNA damage response.

89 Results

90 An evolutionary framework of novel telomeres

91 Previous work has shown that *S. cerevisiae* yeasts engineered to carry point mutations in
92 the *TLC1*-encoded 16 nt telomerase RNA template have growth and chromosome
93 segregation defects (Lin et al., 2004), suggesting that this motif is under strong
94 evolutionary constraints that limit its freedom to vary in nature. To quantify these

95 constraints, we examined the template and its flanking regions in 8 *Saccharomyces*
96 species spanning nearly 20 MY of evolution (Shen et al., 2018). Our multiple sequence
97 alignment revealed that while flanking regions show variation that reflects the species
98 evolutionary relationships the template region is extremely conserved, with only two
99 variants: in the first and last nt positions (**Fig. 1a**). We further analyzed intraspecific
100 variation in *TLC1* in a collection of 3380 natural *S. cerevisiae* isolates to explore recent
101 mutations across the ecological and geographic range of this model species. We found
102 two rare variants, again in the first or the last nt of the template, with minor allele
103 frequencies of 0.002 and 0.0003 respectively (**Fig. 1a**). Thus, purifying selection acting on
104 the *TLC1* sequence in natural yeast lineages is remarkably strong.
105 In order to reconfigure *S. cerevisiae* telomere repeats, we used a genetic system that
106 generates isogenic strains with either wild-type or humanized telomeres (Bah et al., 2004;
107 Bah et al., 2011) (**Table S1**). In brief, these two strains differ by a low-copy centromeric
108 plasmid with either a native (p*TLC1*-y) or humanized (p*TLC1*-h) copy of *TLC1*, with the
109 chromosomal *TLC1* having been deleted from its native locus. These strains contain an
110 additional plasmid with a copy of *TLC1*-y (pAZ1) that serves as a back-up and pAZ1 loss
111 can be controlled by plating cells in a selective medium (**Fig. 1b**).
112 We initiated the telomere reprogramming by streaking the ancestors on 5-FOA for two
113 single-cell bottlenecks (scb) to select for *URA3*-containing pAZ1 loss. We investigated the
114 effect of pAZ1 loss on telomere DNA length and composition, as well as on mitotic cell
115 growth and lifespan. First, we sequenced the ancestors' genomes before (0 scb) and after
116 pAZ1 loss (2 scb) to estimate telomere length and composition from short-read data
117 (D'Angiolo et al., 2023). The TG_{1-3} telomere length decreased upon pAZ1 loss, for both the
118 wild-type and humanized ancestor, consistent with telomere length being affected by the
119 copy number of *TLC1* (Zappulla & Cech, 2004). The wild-type did not contain T_2AG_3
120 repeats, while a small amount was detected (5% T_2AG_3 reads) in the humanized ancestor
121 at both timepoints (**Fig. 1c and Table S2**). The rare T_2AG_3 telomeric repeats and the loss
122 of pAZ1 did not immediately affect the cell's doubling time across several conditions or
123 chronological lifespan (**Fig. 1d-f, Tables S5-7**).
124 Next, we investigated the impact of pAZ1 loss on the global proteome, using liquid
125 chromatography tandem mass spectrometry. We only detected 4 differentially expressed
126 proteins in the humanized ancestor at 0 scb as compared to the wild-type (**Table S11**).
127 However, we identified 14 differentially expressed proteins in the humanized ancestor at 2
128 scb. In humanized yeasts, *ADE2* is inactivated by a loss-of-function mutation in its native
129 locus and a functional copy is inserted into the right subtelomere of chrV and acts as

130 subtelomeric transcriptional silencing reporter. The replacement of Rap1p by Tbf1p results
131 in transcriptional derepression of subtelomeric genes most likely due to the lack of the
132 deacetylases Sir2-4 binding to the decreased number of telomeric Rap1p (Alexander &
133 Zakian, 2003) and/or the insulator activity of Tbf1p (Fourel et al., 1999). Consistent with an
134 early loss of transcriptional silencing, we found Ade2p to be upregulated in yeast cells
135 carrying the humanized *TLC1* at 2 scb (**Fig. 1g, TableS S11, S12**). Overall, we found the
136 controlled loss of pAZ1 to initiate the telomeric replacement of yeast TG_{1-3} with human
137 T_2AG_3 repeats and that this initiation coincided with modest proteomic remodelling. Thus
138 we conclude that it can be used to study the long-term consequences of a telomeric repeat
139 reprogramming.

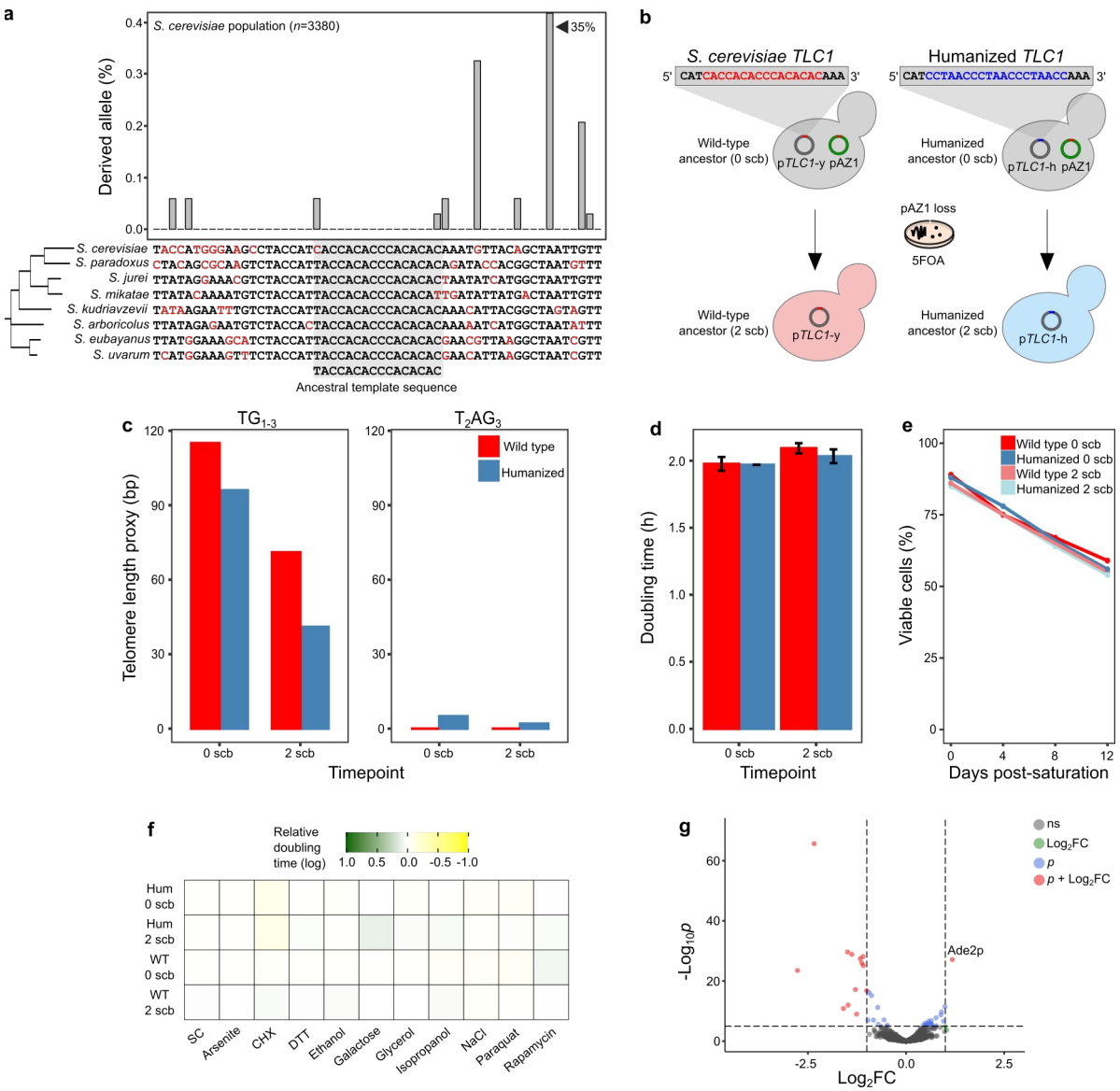

Fig. 1 – *TLC1* evolution and reprogramming. a, Natural sequence variation of the telomerase RNA template (*TLC1*). The barplot represents the percentage of *S. cerevisiae* strains carrying a derived allele at each position of the alignment (*n*=3380). The alignment is based on the 8 species of the *Saccharomyces* genus. Black letters denote ancestral alleles while red letters denote the inferred derived alleles. The upper panel represents an intra-specific comparison, while the bottom panel represents an inter-specific comparison. **b**, Schematics of the plasmids contained in the ancestor strains at 0 and 2 scb. **c-e**, Telomere length, population doubling time and survival rate of wild-type and humanized ancestors at 0 and 2 scb. Barplots: mean + standard deviation. Telomere length measurements were performed in one replicate. Population doubling time measurements were performed in triplicate. Survival rate measurements were performed in one replicate. **f,** Relative doubling time of wild-type and humanized ancestors at 0 and 2 scb. Colours indicate growth normalized using wild-type at 2 scb (MJD5). Relative doubling time measurements were performed in 24 replicates for SC and 12 replicates for all the other conditions. **g**, Volcano plot showing differentially expressed proteins in the humanized ancestor at 2 scb as compared to wild-type 2 scb. Proteins supported by fold-change (x), *p*-value (y) or both are indicated in color, but only the ones supported by both (red) were retained for analysis.

7

155 **Chromosome end structures of humanized yeasts**

156 We induced the telomere reprogramming by evolving 16 clones from the wild-type and
157 humanized ancestors for 100 scb through a mutation accumulation lines (MAL) protocol,
158 which minimizes selective pressure (**Fig. 2a, Supplementary Fig. 1 and Table S1**).
159 During MAL, humanized yeasts' TG_{1-3} repeats are gradually replaced by T_2AG_3 repeats,
160 but they maintain a short centromere-proximal stretch of TG_{1-3} repeats (Auriche et al.,
161 2008). Five wild-type lines became "*petites*" (growth deficiency caused by mutation or loss
162 of mitochondrial DNA (Vowinckel et al., 2021)) and were not further characterized. We
163 estimated that the mutation accumulation lines underwent ~2140 and ~1640 generations
164 for wild-type and humanized lines, respectively (**Table S1**). Several wild-type and
165 humanized lines subsequently were used to initiate adaptive evolution lines through 31
166 serial transfers (st, ~400 generations) of large population sizes to explore adaptation to
167 new telomeric sequences (**Fig. 2a and Supplementary Fig. 1**).

168 To characterize the telomere DNA upon reprogramming, we applied short-read and long-
169 read technology to both MAL and AEL. Short-read sequencing confirmed the absence of
170 T_2AG_3 repeats in wild-types, while their amount increased in the humanized lines over
171 mitotic generations (**Fig. 2b and Table S2**). ITS content and Y' copy number (CN)
172 increased in humanized yeasts, a hallmark of type I survivors when cells overcome
173 telomerase inactivation via recombination-mediated telomere lengthening. Humanized
174 yeasts also carried ITS with T_2AG_3 repeats suggesting telomere-subtelomere
175 recombinations (**Fig. 2c-d and Table S2**).

176 Long-read sequencing can resolve telomeres at single chromosome-end resolution
177 (Karimian et al., 2024; Sanchez et al., 2024; Schmidt et al., 2024; Sholes et al., 2022;
178 Tham et al., 2023). We thus applied Nanopore sequencing to the humanized ancestor at 2
179 scb, 4 MAL and 2 single clones from the same AEL (**Fig. 2e**). At 2 scb, three telomeres
180 (chrVI_R, chrXI_L, chrXVI_L) contained small amounts of terminal T_2AG_3 repeats,
181 confirming that telomeric reprogramming starts at the first single-cell bottlenecks. At 100
182 scb, almost all telomeres are constituted by an internal TG_{1-3} (median 45.5±23.3 bp) and a
183 terminal T_2AG_3 stretch (median 224±357.5 bp). While the TG_{1-3} length is similar across
184 telomeres, the T_2AG_3 lengths show remarkable variation with the Hum15 line with ultra-
185 long telomeres (>3 kb in chrIV_R) (**Fig. 2e, Supplementary Fig. 2a and Table S3**).

186 ITS and Y' elements consistently increased in MAL and AEL, with chrVIII_L and chrXVI_R
187 being preferential sites for their amplification (**Fig. 2f**). We confirmed the presence of
188 T_2AG_3 repeats at ITS, with 5 conformations, implying recombination between telomeres
189 and ITS (**Fig. 2g, Supplementary Fig. 2b-c and Tables S3-4**). In contrast, X elements

were largely stable with 1 copy per chromosome-end, except for chrXVI_R that carried 2 copies and occasionally amplified on chrVIII_L (**Fig. 2f and Table S4**). Abundant chromosome size variations that are compatible with subtelomeric amplifications were also observed (**Supplementary Discussion 1 and Supplementary Fig. 2d**). These results raise the possibility that the amplification of subtelomeric ITS and Y' elements compensates for the inadequate chromosome end protection in humanized yeasts. To further probe this hypothesis, we analyzed co-variation between telomere length, ITS and Y' CN. Y' and ITS CN were strongly correlated both in MAL and in AEL (r=0.96 and r=0.99, $p<2.2e^{-16}$) (**Supplementary Fig. 2e**).

We investigated the impact of telomere humanization on proteins that directly bind telomeric repeats and the telomere silencing reporter through mass spectrometry. The abundance of Ade2p was significantly higher in humanized lines than in wild-types in both MAL and AEL, consistent with loss of transcriptional silencing in subtelomeres (p=0.0003 in MAL at 100 scb). This shift was mirrored by lower Rap1p levels (p=0.09 in AEL from 100 scb) and higher Tbf1p levels (p=0.02 in MAL at 100 scb), consistent with humanized yeasts adjusting the level of their telomere-binding-proteins according to telomere DNA sequence composition or length (**Fig. 2h**). In contrast, only 15 of the 147 known telomere-length-maintenance proteins that were present in our proteomics dataset showed significant difference between humanized and wild-type yeasts (**Supplementary Discussion 2**).

Overall, we disentangled the telomere and subtelomere structure of yeasts with humanized telomeres at unprecedented resolution, revealing high telomere/subtelomere dynamics and heterogeneity.

9

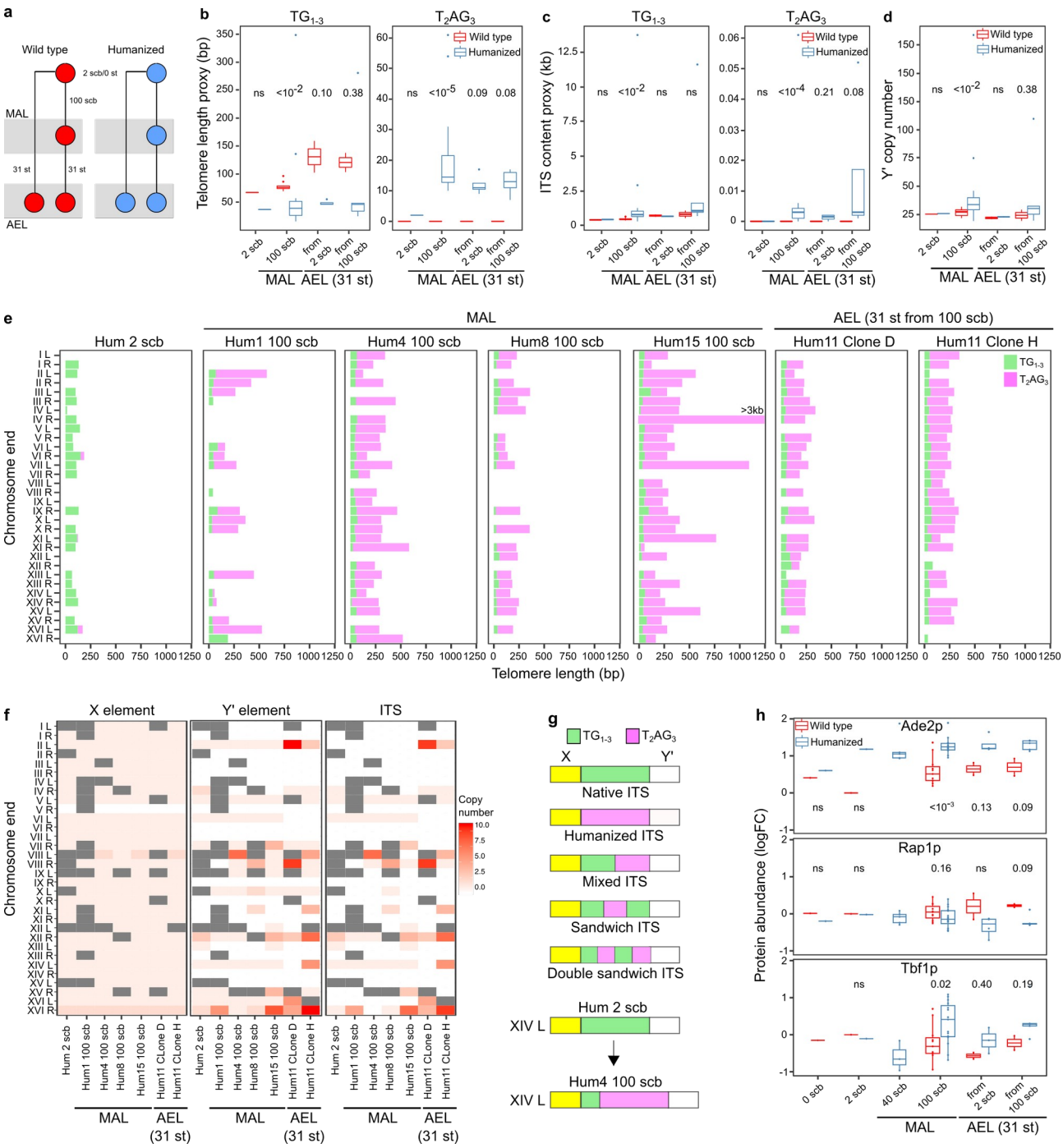

Fig. 2 – Chromosome-end dynamics during MAL and AEL. a, Experimental protocols used to evolve wild-type (red) and humanized (blue) ancestor strains in the absence (MAL-top panel) or in the presence (AEL-bottom panel) of selection. AEL was performed in two ways: starting from MAL at 100 scb or from the ancestors at 2 scb. **b-d** Telomere DNA length, ITS content and Y' CN estimations in MAL at 2 (*n*=1) and 100 scb (*n*=11 and *n*=16 for wt and hum), and after AEL started from 2 (*n*=2 and *n*=4, respectively) and 100 scb (*n*=2 and *n*=5, respectively). Telomere length estimations derive from the number of telomeric reads present in short-read sequencing data. As the calculation parameters differ between TG_{1-3} and T_2AG_3 motifs, the respective values are not comparable between each other. Boxes: horizontal line: median, upper/lower hinge: interquartile range (IQR), whiskers: largest/smallest value within upper/lower hinge +/- 1.5x IQR. Numbers in the plots represent *p*-values resulting from comparison of wild-type and humanized lines at the underlying timepoints (two-tailed Wilcoxon test). **e**, Telomere lengths at single chromosome-ends derived from long-read genome assemblies. The order of TG_{1-3} (green) and T_2AG_3 (pink) repeats reflects telomere structure. Empty lines denote truncated chromosome-ends with unavailable data. In TG_{1-3}-only telomeres the mapping of long reads onto the chromosome-ends showed a mixed composition comprising reads with only TG_{1-3} repeats and reads with TG_{1-3} followed by T_2AG_3 repeats. This underlies high heterogeneity within a line. **f**, Number of X, Y' elements and ITS at single chromosome-ends. Grey boxes denote missing data due to truncated chromosome-ends. **g**, Upper panel: schematics of ITS structures with i) TG_{1-3}-only, ii) T_2AG_3-only,

iii) mixed ITS with a distal TG_{1-3} stretch followed by T_2AG_3 repeats, iv) sandwich ITS with T_2AG_3 repeats embedded in two TG_{1-3} stretches, v) double sandwich ITS with a repeated sequence of two mixed ITS. Double sandwich ITS are observed only in AEL. Lower panel: example of ITS that was native at 2 scb but became mixed at 100 scb. **h**, Protein abundance (log_2FC) of Ade2p, Rap1p and Tbf1p during MAL at 0 (*n*=1), 2 (*n*=1), 40 (*n*=5 for humanized only) and 100 scb (*n*=11 and *n*=16 for wild type and humanized, respectively), and after AEL started from 2 (*n*=2 and *n*=4, respectively) and 100 scb (*n*=2 and *n*=5, respectively). Protein abundance measurements were performed in triplicate for all strains except the wild-type MAL at 100 scb (*n*=1).

11

238 **The fitness and molecular landscape of humanized yeasts**

239 We investigated the impact of telomere humanization on yeast fitness by measuring
240 mitotic cell growth and chronological lifespan. During MAL, the cell doubling time remained
241 largely stable for the wild-types, while humanized lines showed early growth defects
242 followed by partial recovery (median doubling time 2.0/2.8 h for WT/Hum at 100 scb,
243 p=1.5e^{-7}). This indicates that some adaptation occurred despite the MAL being bottle-
244 necked to random single cells (or at least extremely small populations) (**Fig. 3a,**
245 **Supplementary Fig. 3a-b and Table S5**).

246 We next tested whether telomere humanization impacts cell growth across environments,
247 including different carbon sources and stresses, and observed that rapamycin and glycerol
248 magnified the fitness defects of humanized yeasts (**Fig. 3b and Table S6**). We further
249 measured the survival rate across a 12-days incubation in a rich medium without
250 refreshing. Humanized lines progressively decreased their survival rate without recovering
251 it (median viability 51/21 % for WT/Hum at day 12 and 100 scb, p=1.7e^{-5}) (**Fig. 3c,**
252 **Supplementary Fig. 3c and Table S7**).

253 Next, we characterized the mitotic growth of AEL started from MAL after 100 scb. During
254 this adaptive evolution, we observed complete growth recovery (median doubling time
255 2.0/2.0 h for WT/Hum at 31 st, p=0.85). This recovery was evident also in non-adaptation
256 environments, including rapamycin and glycerol (p=0.05 combining all conditions).
257 However, it is accompanied by a decrease in survival rate (median viability 32/18 % for
258 WT/Hum at day 12 and 31 st, p=0.09) (**Fig.s 3a-c, Supplementary Fig. 4a and Tables**
259 **S5-7**). When started from 2 scb, both wild-type and humanized AEL had the same growth
260 and survival and even improved their growth in certain environments when compared to
261 the starting clones. Thus, if telomere reprogramming and selection progress in concert,
262 cells keep their fitness (**Fig. 3a-c, Supplementary Fig. 4b and Tables S5-7**).

263 We examined the proteomic profiles of humanized and wild-type yeasts and derived a list
264 of 110 proteins whose abundance changed significantly in at least one evolution condition
265 (MAL or AEL). We used hierarchical clustering to group proteins based on their $\log_2$FC co-
266 variation, uncovering 8 major clusters (**Fig. 3e**). Clusters 1,2,3,5 contain proteins down-
267 regulated in humanized yeasts, with an enrichment for the pentose phosphate
268 pathway/NADPH regeneration, transport/localization, aminoacid metabolism, and stress
269 response, respectively. Proteins in clusters 1,5 remained downregulated after AEL from
270 100 scb, while those in clusters 2,3 returned to normal abundance. Of note, proteins in
271 cluster 3 exhibit opposite trends between MAL and AEL from 2 scb. Cluster 4 contains
272 proteins up-regulated in humanized yeasts and is enriched for deoxyribonucleotides

273 synthesis. Clusters 6 and 7 comprise proteins down- or upregulated in AEL from 2 scb,
274 respectively. Cluster 8 contains proteins upregulated in all strains, with no functional
275 enrichment (**Fig. 3d and Table S12**). Proteomic profiles of humanized and wild-type
276 yeasts were similar in AEL initiated from 2 scb., indicating that their differential protein
277 abundance is mainly driven by growth advantage and faster metabolism, rather than
278 telomere structure. This effect might underlie a common adaptation signature to the
279 dropout medium through a modification of aminoacid metabolism. Furthermore, we
280 performed Principal Component Analysis to group strains based on their proteomic profile
281 and separated humanized and wild-type MAL. Humanized AEL from 100 scb clustered
282 together with the wild type MAL and the ancestors, confirming that they recovered a
283 proteomic profile similar to the wild types as they recovered their growth capacity. In
284 contrast, both humanized and wild-type AEL from 2 scb clustered together (**Fig. 3e**). We
285 further defined a "Telomere Humanization Proteomic Response" (THPR) as the proteins
286 that are differentially abundant in at least 2 humanized MAL at 40 or 100 scb (n=283). We
287 compared the THPR with genes whose RNA is differentially expressed upon telomere
288 dysfunction induced by a *cdc13* temperature-sensitive mutation and that are present in our
289 dataset (n=232), and found a significant overlap between them (n=79, two-tailed X^2 test,
290 p=6.5e^{-20}, **Table S12**) (Greenall et al., 2008). In contrast, we did not find a significant
291 overlap between the THPR and Rap1p targets (n=345) (Brindle et al., 1990; Duveau et al.,
292 2021; Hu et al., 2007; Lickwar et al., 2012; MacIsaac et al., 2006; McNeil et al., 1990;
293 Song et al., 2020, Stanway et al., 1989), indicating that the THPR is not driven by Rap1p-
294 dependent transcription (**Supplementary Discussion 2**).
295 Taken together, our experiments revealed that reconfiguring telomeric DNA sequences
296 causes a severe fitness decline accompanied by a proteomic deregulation and partial
297 reprogramming, but yeasts can recover wild-type growth.

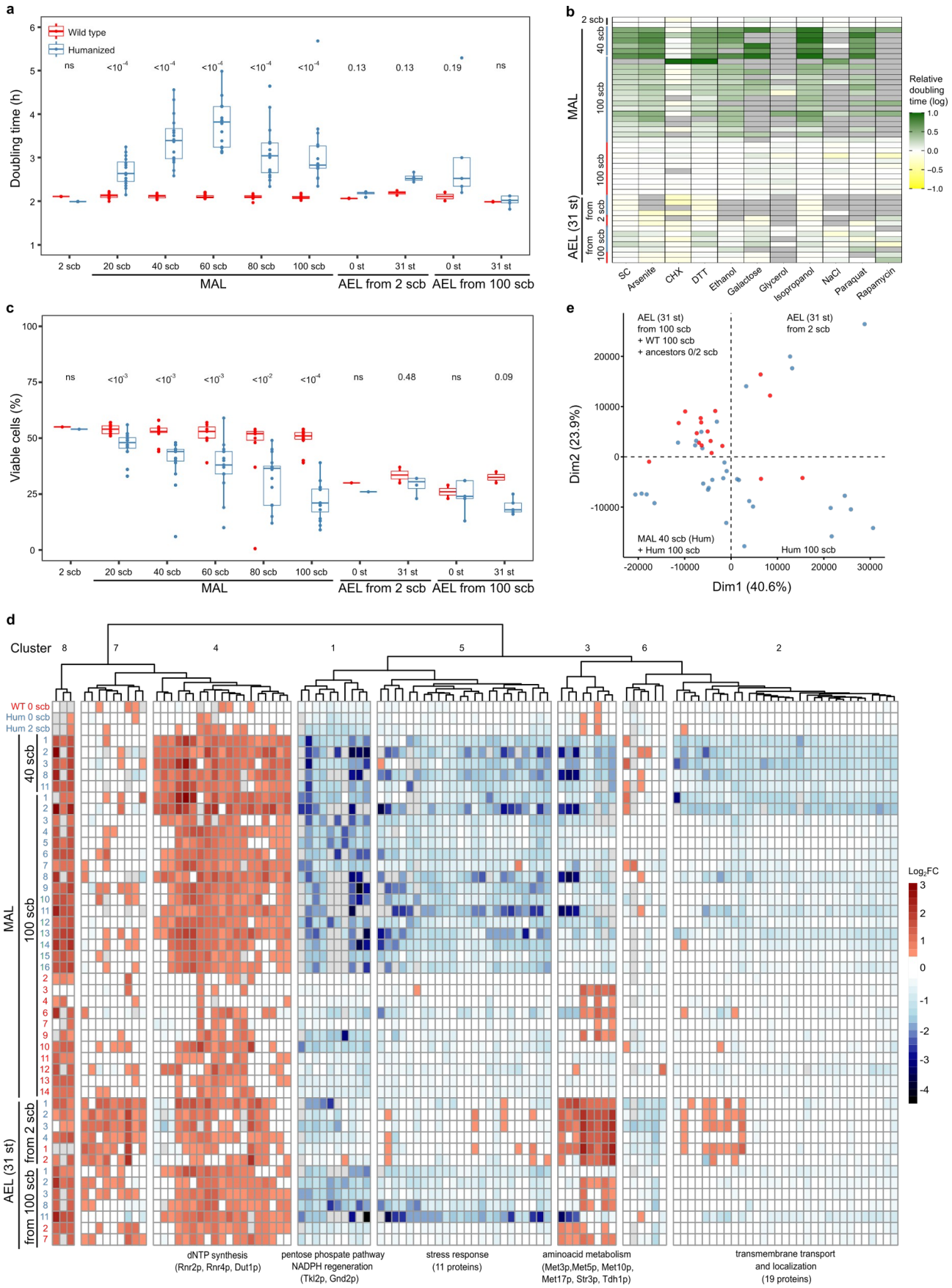

Fig. 3 – Fitness decay and recovery. a-c, Population doubling time, survival rate and relative doubling time of MAL every 20 scb, and of AEL. Boxplots as in Fig. 2. In Fig. **b**: green, white and yellow tones denote poor, equal or better growth respect to the control strain used for normalization (WT1 at 2 scb). Grey boxes denote missing data due to extremely poor growth. Red and blue lines denote WT and Hum yeasts, respectively. $p<10^{-16}$ (WT vs Hum 100 scb), $p<10^{-6}$ (WT vs Hum 100 scb/0 st), p=0.05 (WT vs Hum 100 scb/31 st), p=ns

14

(WT vs Hum 2 scb/0 or 31 st). Fig. **c** shows values at 12 days post-saturation.Population doubling time measurements were performed in triplicate. Survival rate measurements were performed in one replicate. Relative doubling time measurements were performed in 24 replicates for SC and 12 replicates for all the other conditions. Numbers or asterisks in the plots represent *p*-values resulting from the comparison of wild-type and humanized lines at the underlying timepoints (two-tailed Wilcoxon test). **p*<0.05, ***p*<0.01, ****p*<0.001, *****p*<0.0001. **d**, Hierarchical clustering of 110 proteins whose abundance changed ($\log_2$FC >1 or >-1, *p*<0.05) relative to a reference wild-type strain at 2 scb. Rows represent MA or AEL, while columns represent proteins. Upregulated and downregulated proteins are shown in red/blue, respectively (starting from log2FC >0.25 or <-0.25). Gray squares indicate missing data. Significant GO terms enrichments for the clusters and their underlying genes are indicated. Protein abundance measurements were performed in triplicate for all strains except the wild-type MAL at 100 scb (*n*=1). **e**, Principal Component Analysis based on proteomic signatures across humanized/wild-type strains at different stages of experimental evolution. Strains belonging to the 4 quadrants and separated across the two dimensions are indicated as text annotations.

15

317 **Mutations in the DNA damage response rescue humanized yeasts fitness**
318 We analyzed whole-genome sequencing data from MAL and AEL to characterize the
319 mutational spectrum and genetic alterations underlying fitness recovery. We identified 250
320 new single-nucleotide variants (SNVs) in the 16 humanized MAL (49/51 % within
321 subtelomeres/core), and 87 in the 11 wild-type MAL (17/83 % within subtelomeres/core).
322 Given that subtelomeres cover 11% of the genome (Yue et al., 2017) and assuming that
323 selection had a negligible influence on variant distributions across MAL, we thus find that
324 subtelomeres constitute a mutational hotspot in humanized yeasts (binomial test,
325 $p<0.0001$). Humanized yeasts have significantly higher mutation rates than wild-types,
326 which instead are in line with previous estimates (Sharp et al., 2018; Tattini et al., 2019)
327 (**Fig. 4a, Supplementary Fig. 5a-c and Table S8**). A GO-term enrichment analysis
328 revealed that wild-type MAL mutations were enriched in genes encoding ion binding
329 proteins ($p=0.03$), while mutations in humanized MAL were enriched in genes encoding
330 ceramid biosynthesis ($p=0.10$), cell periphery ($p=0.06$) and DNA damage response
331 proteins (MAL Hum8 - *TEL1* and MAL Hum4 - *MEC1*; $p=0.09$), suggesting that these
332 pathways could underlie their fitness recovery (**Fig. 4b and Table S9**). Both the *MEC1*
333 (R1824) and *TEL1* (L2308) mutations are aminoacid substitutions located in the FAT
334 domain of the respective proteins, with the *TEL1* mutation located in a highly conserved
335 region and therefore predicted to be impactful. The humanized MAL carrying the *TEL1*
336 L2308R mutation underwent massive ITS/Y' amplifications, chromosomal rearrangements
337 and whole-genome duplication. The *MEC1* R1824H mutation occurred between 40-60 scb
338 whereas the *TEL1* L2308R mutation occurred between 60-80 scb. The occurrence of
339 *TEL1* L2308R precedes fitness recovery, supporting that DDR inactivation rescues fitness
340 in humanized yeasts (**Fig. 4c, Supplementary Fig. 3a-b and Table S5**).
341 Bulk sequencing of AEL at 31 st unveiled only few new mutations ($n=13/38$ in AEL starting
342 from 2/100 scb), with the tryptophan tRNA genes recurrently affected ($n=6$). Mutations
343 affecting components of the MRX complex, which binds DNA breaks or uncapped
344 telomeres and then activates the DDR through the recruitment of Tel1p and Mec1p,
345 occurred recurrently and independently in three AEL: Hum1 - *MRE11 K477I*; Hum2 -
346 *MRE11 Q447P*; Hum3 - *RAD50 L171P*). These lines grew poorly at 0 st but completely
347 recovered their fitness after AEL (31 st), and all three mutations are predicted impactful
348 through SIFT score analysis (**Fig. 4c**). The two *MRE11* mutations in the DNA binding
349 domain and *RAD50* mutation in the ATPase domain were all fixed, in line with that they are
350 strongly beneficial (**Fig. 4b-c and Tables S8-9**). To confirm the adaptive role of the MRX
351 mutations, we engineered *MRE11* Q447P into the slow-growing humanized line 3 at 100

352 scb. This mutation produced a full growth recovery across conditions, but decreased the
353 survival rate (**Fig. 4d and Tables S5-7**). We tested whether the beneficial effect would
354 also manifest in fit humanized yeast that had not yet suffered a loss of reproductive
355 capacity by introducing *MRE11* Q447P in the humanized line 1 at 2 scb. However, we
356 observed no improvement in growth or lifespan, consistent with that its benefits are strictly
357 limited to compensating for the fitness loss arising from telomere reprogramming
358 (**Supplementary Fig. 5d and Tables S5-7**). The full deletion of *MRE11* improved both
359 growth and survival rate but did not match the growth of the wild type MAL
360 (**Supplementary Fig. 5e and Tables S5-7**). Taken together, these results suggest that
361 *MRE11* mutations induce partial loss-of-function.
362 To further disentangle how DDR modulation rescues fitness in humanized yeasts, we
363 inspected protein abundance of known DDR targets in the proteomic dataset. We
364 observed upregulation of the ribonucleotide reductase (RNR) subunits Rnr2p/Rnr4p and
365 the deoxyuridine triphosphate diphosphatase Dut1p, and the downregulation of the RNR
366 inhibitor Sml1p. The upregulation of Rnr2p/Rnr4p/Dut1p is highest in MAL at 40 scb while
367 it decreases after AEL from 100 scb ($p<10^{-2}$ for all proteins at 100 scb) (**Fig. 4e**). We
368 further checked the abundance of Rnr2p/Rnr4p/Dut1p in a humanized AEL that
369 spontaneously developed a mutation in *MRE11* (Hum2 31 st *MRE11* Q447P), as well as a
370 line where we introduced the same mutation by CRISPR/Cas9 genetic engineering (Hum3
371 100 scb *MRE11* Q447P). The three proteins decrease when *MRE11* Q447P is present,
372 suggesting attenuation of the DDR (**Fig. 4f**).
373 Overall, these findings reveal that humanization of yeast telomeres increases mutation
374 rate and the partial inactivation of DDR rescues their fitness through its downstream
375 targets.

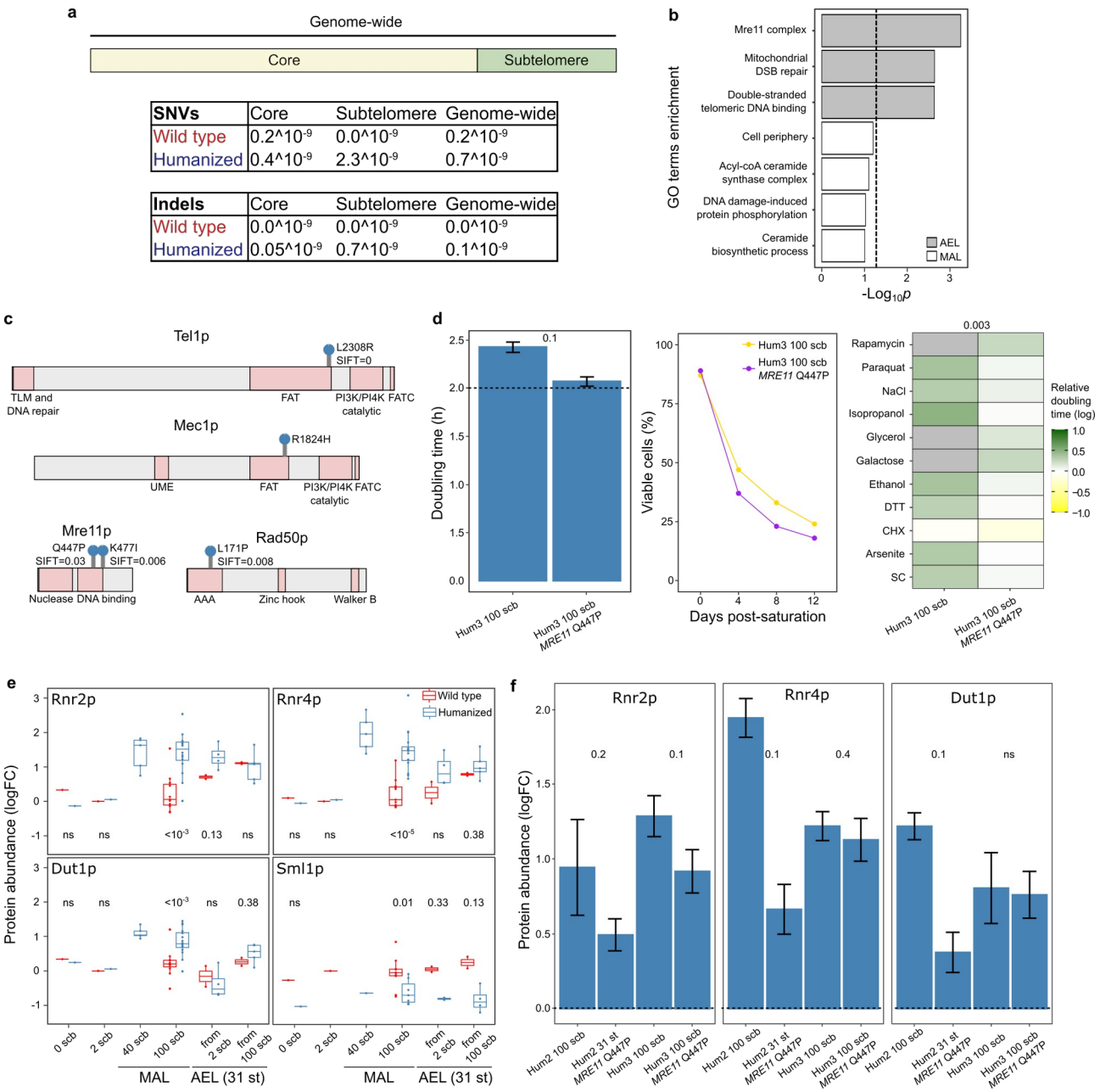

Fig. 4 – Effect of telomere humanization on mutation rate and DDR targets. a, Base substitution and insertion/deletion rate per bp per generation in wild-type and humanized MAL, in core (yellow), subtelomeres (green) and genome-wide. **b**, GO terms enriched for mutations occurred in humanized MAL (white bars) and AEL (grey bars). The dashed line indicates the significance threshold at $p<0.05$. **c**, Location of mutations in the DDR proteins. Protein domains are indicated by pink boxes. **d**, Population doubling time, survival rate and relative doubling time of one humanized line at 100 scb (Hum3) before and after the introduction of a missense mutation in *MRE11* (Q447P). The dashed line indicates the population doubling time of wild-type. Colour codes and number of replicates are as described in Fig. 3. $p=0.003$ for doubling time across combined conditions. Barplot: mean + standard deviation. **e-f**, Protein abundance ($\log_2$FC *vs* the reference wild-type at 2 scb) of Rnr2p, Rnr4p, Dut1p (**e**,**f**) and Sml1p (only in **e**) in wild-type and humanized yeasts during MAL at 0, 2 ($n=1$ for both), 40 ($n=5$ for hum only) and 100 scb ($n=11$ and $n=16$, respectively), and after AEL started from 2 ($n=2$ and $n=4$, respectively) and 100 scb ($n=2$ and $n=5$, respectively). In **f,** Sml1p has been removed due to unavailable data, and the dashed line indicates the $\log_2$FC of the reference. Protein abundance measurements were performed in triplicate for all strains except the wild-type MAL at 100 scb. Boxplots as in Fig. 2. Barplots: mean + standard deviation. Numbers in the plots represent *p*-values resulting from the comparison of wild-type and humanized lines from the underlying experimental evolution timepoints or genetic backgrounds (two-tailed Wilcoxon test).

393 **Chromosome XVI aneuploidy is adaptive in humanized yeasts**
394 After revealing the beneficial effect of DDR point mutations, we investigated whether
395 adaptation also occurred via large-scale copy number and structural variants (SVs). All the
396 MAL remained haploid except for a whole-genome duplication in the humanized MAL 8
397 with a *TEL1* loss-of-function mutation. This line likely underwent endoreduplication since
398 the *MAT* locus remained homozygous (*MATa/MATa*) (**Fig. 5a and Table S10**) (Harari et
399 al., 2018). We next analyzed long-read-sequencing-derived genome assemblies from 4
400 humanized MAL at 100 scb (Hum1,4,8,15) and an ancestor at 2 scb. We did not detect
401 any major SVs except for Hum8 at 100 scb, which underwent a translocation between
402 chromosomes V and XV mediated by the two *ADE2* copies (**Supplementary Fig. 6, Table**
403 **S10 and Supplementary Discussion 1)**.
404 We next analysed read coverage along chromosomes to determine whether large CNV
405 and aneuploidies occurred during MAL. We discovered one segmental duplication of
406 chromosome IV (~100 kb, WT2) and 17 whole chromosome aneuploidies in our lines.
407 These were all chromosome gains and exclusively occured in humanized lines (except
408 one in WT6). Chromosome XVI was aneuploid in 12 lines suggesting a possible adaptive
409 role (**Fig. 5a and Table S10**). We further explored whether an additional chromosome XVI
410 is beneficial by comparing humanized lines carrying 1 or 2 copies of this chromosome.
411 Lines with 2 copies had shorter telomeres, and this difference was driven by fewer T_2AG_3
412 repeats (p=0.01) (**Fig. 5b, Supplementary Fig. 7a and Table S2**). Furthermore, they had
413 a shorter doubling time in multiple environments but lower survival rate (p=8.8e^{-8} for
414 doubling time across conditions, p=0.02 for survival rate at day 12) (**Fig. 5b,**
415 **Supplementary Fig. 7b and Tables S5-7)**. To better pinpoint the emergence of the
416 aneuploidy during MAL evolution, we probed its existence in the stored records through
417 qPCR and found that it occurred between 60-100 scb (only one case at 40-60 scb), with its
418 emergence correlating well with the decrease of both doubling time and survival rate,
419 confirming a causative role (p=0.22/0.02 for doubling time/survival at day 12) (**Fig. 5a,**
420 **Supplementary Fig. 7c-e and Tables S5-7**).
421 Aneuploidies are usually detrimental in nature, especially when involving long
422 chromosomes, but can also fuel adaptation (Pompei & Cosentino Lagomarsino, 2023;
423 Rojas et al., 2024; Sheltzer & Amon, 2011). We inspected whether any telomere-binding-
424 proteins were encoded on chromosome XVI, detecting Tbf1p, which directly binds T_2AG_3
425 repeats (Brigati et al., 1993). We tested the effect of *TBF1* duplication by engineering its
426 copy number, first by deleting one *TBF1* copy in a humanized MAL carrying two
427 chromosomes XVI (Hum9) and then by inserting a second *TBF1* copy in a humanized

428 MAL carrying a single chromosome XVI (Hum12). We found that the addition of a second
429 *TBF1* copy to a humanized MAL carrying one improved growth, while the deletion of one
430 out of two *TBF1* copies increased the doubling time by 17% (**Fig. 5c**). The beneficial effect
431 of an extra *TBF1* in humanized yeasts likely derives from better telomere capping and/or
432 from Tbf1p's inhibitory action towards the binding of the MRX complex to telomeres and its
433 subsequent activation of DDR, as supported in our proteomic data by the inverse
434 relationship between protein levels of *TBF1* and DDR targets (Rnr2p, Rnr4p, Dut1p)
435 (Berthiau et al., 2006; Fukunaga et al., 2012) (**Supplementary Fig. 8a-b**). These results
436 suggest that the fitness advantage conferred by *TBF1* duplication is due to a combination
437 of telomere protection and DDR-inhibitory functions.
438 Next, we investigated the evolutionary dynamics of large-scale structural variants during
439 AEL. All the AEL derived from ancestors at 2 scb, i.e. from before substantial telomere
440 reprogramming, maintained their original ploidy and showed no major structural variants.
441 In contrast, AEL derived from MAL at 100 scb and that had gained extra chromosomes
442 during the MAL evolution, frequently lost the extra chromosomes during the adaptive
443 evolution and thereby restored a balanced genome content. Humanized lines 2,8,3 lost
444 their additional copies of chromosomes V, XI and XVI, respectively, while Hum11 revealed
445 a mixed population of clones carrying different segments of the extra copy of chromosome
446 XVI (**Fig. 5a and Table S10**). We isolated 8 single clones from this population and
447 uncovered 4 distinct chromosome XVI conformations by qPCR and whole-genome
448 sequencing. All the segments' breakpoints were flanked by repetitive sequences such as
449 Ty and long terminal repeats that likely mediated the recombination events. Notably, the
450 region encompassing the *TBF1* gene remained duplicated in all conformations and
451 included the centromere (CEN16) (**Fig. 5d**). We applied long-read sequencing to two of
452 these clones to disentangle their chromosome XVI structure. Clone D carried almost the
453 full version of duplicated chromosome XVI, except for a terminal segment, and the
454 duplicated segment was stabilized by the duplication and translocation of chrVIII-R
455 telomeres. Clone H carried only a central duplicated segment of chromosome XVI and this
456 segment was circularized with the circularization likely mediated by Ty elements and
457 stabilized by CEN16 (**Fig. 5e**). This circular structure is likely present in all conformations
458 lacking terminal duplicated blocks (i.e. conformations 2,3,4) and has the advantage of not
459 needing telomeres. We tested whether these segmental duplications resulted in a double
460 amount of proteins. Proteomic profiles mirrored read coverage ones for the 4 chrXVI
461 conformations (**Supplementary Fig. 8c**). However, the average abundance for proteins
462 encoded by genes in these segments was less than 2-fold higher than that of other

463 proteins, indicating a buffering effect. An extension of this analysis to the whole genome
464 revealed a partial gene dosage compensation at the proteome level, which was probably
465 selected during the long-term evolution (**Fig. 5f and Supplementary Discussion 3**)
466 (Muenzner et al., 2024). We reasoned that this conformations mixture might be the result
467 of selection favouring conformations carrying the shortest duplicated block containing
468 *TBF1*, in order to minimize the detrimental effects associated with the duplication of all
469 chrXVI genes. To test this hypothesis we measured growth in the 8 clones and confirmed
470 an inverse relationship between growth and chrXVI duplicated content (p=2.9e^{-5} between
471 conformations 1-2 and 3-4) (**Fig. 5g and Tables S5-7**).
472 Taken together, these results show that chrXVI aneuploidy is a primary evolutionary route
473 to fitness recovery driven by the duplication of *TBF1*.

21

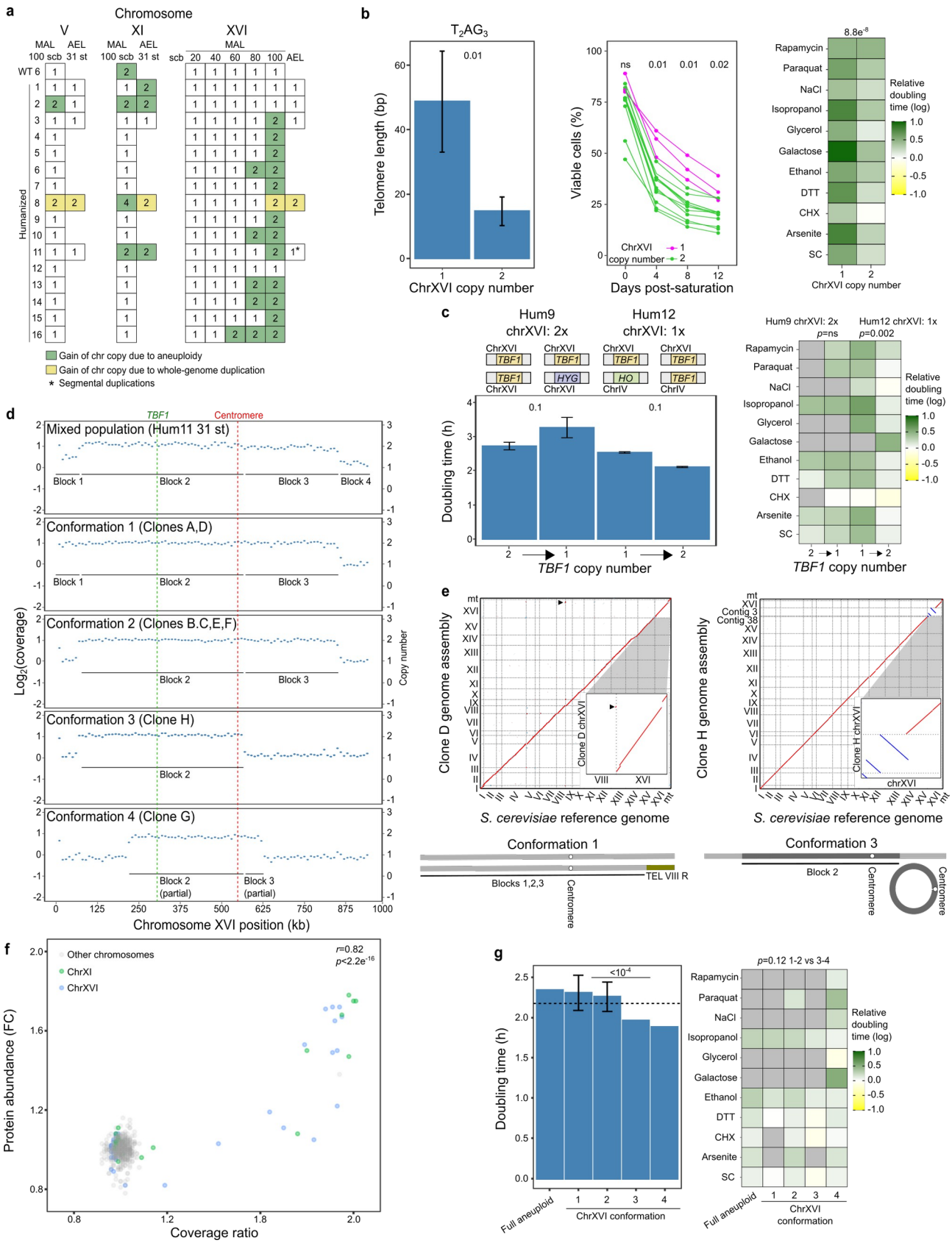

Fig. 5 – Structural variation during MAL and AEL. a, Chromosome-scale copy number of chromosomes that underwent copy number variation in the MAL and AEL. White boxes represent euploid chromosomes that maintained their original copy number (1), while green/yellow boxes indicate gain of copy by aneuploidy or whole-genome duplication, respectively. **b**, T_2AG_3 telomere length, survival rate and relative doubling time in humanized MAL that are either euploid (one copy) or aneuploid (two copies) for chromosome XVI. Barplot: mean + standard deviation. Color codes and number of replicates are as described in Fig. 3. Hum8 was excluded due to its diploidy. Numbers in the plots represent *p*-values resulting from the comparison of

22

humanized lines from the underlying genetic backgrounds (two-tailed Wilcoxon test). **c**, Population and relative doubling time of humanized lines that underwent the knock-out (left) or the knock-in of one copy of *TBF1*, before and after the genetic engineering. The arrow indicates the direction of the genetic engineering. Numbers in the plots represent *p*-values resulting from the comparison of humanized lines from the underlying genetic backgrounds (two-tailed Wilcoxon test). Barplot: mean + standard deviation. **d**, Coverage profiles on chromosome XVI for an AEL started from 100 scb (top panel) and of 8 single clones derived from it (bottom panels). Dashed lines indicate the position of *TBF1* (green) and the centromere (red). Plain lines indicate the length of duplicated blocks as defined by the points of coverage breakage in the mixed population (top panel). **e**, Comparison of the *S. cerevisiae* reference genome (x axis) and our Nanopore-derived genome assemblies for two clones (D,H) derived from the mixed population. Sequence homology signals are indicated in red (forward match) or blue (reverse match). Insets show a zoom-in of chromosome XVI. Karyotypes on the right show the structure of the two chromosome XVI copies. **f**, Correlation between chromosome copy number estimated from short-read sequencing coverage (x axis) or from normalized FC (y axis), across samples for whom we possessed both sequencing and proteomics data (n=44). Aneuploid chromosomes (XI and XVI) are highlighted in colour. **g**, Population and relative doubling time of clones grouped by their chrXVI conformation (n=2,4,1,1 for conformation 1,2,3,4 respectively). The fully aneuploid (for chrXVI) Hum11 at 100 scb is shown on the left (n=1). The dashed line indicates the doubling time of the mixed population (Hum11 at 31 st). Numbers in the plots represent *p*-values resulting from the comparison of humanized lines from the underlying conformations (two-tailed Wilcoxon test). Barplot: mean + standard deviation.

23

501 **Discussion**

502 In this work, we studied the evolution of budding yeast cells carrying humanized T_2AG_3
503 telomeric repeats through the combination of two laboratory evolution protocols: mutation
504 accumulation lines (MAL) and adaptive evolution lines (AEL). Humanizing the telomeric
505 repeats opens the possibility to investigate the long-term impact of telomeric
506 reprogramming enabling adaptation via the "pre-adapted" factor Tbf1p, in contrast to other
507 random unnatural telomeric repeats whose usage will require long-term evolution of other
508 proteins that will be outcompeted by the emergence of telomerase-negative mechanisms.
509 Our results show that humanized yeasts accumulate T_2AG_3 repeats, and that this telomere
510 sequence reprogramming impairs fitness and increases genome instability. This is in
511 agreement with humanized telomeres inducing replication fork stalling and cell cycle
512 delays, and being subject to a high repeat turnover (Bah et al., 2004; Bah et al., 2011; Di
513 Domenico et al., 2009), thus explaining the need for massive subtelomeric amplifications
514 of ITS and Y' elements to prevent chromosome-end uncapping. Increased genome
515 instability as a consequence of telomere dysfunction was already observed in humanized
516 yeasts (Di Domenico et al., 2013) and in telomerase-deficient *est1Δ S. cerevisiae* strains
517 (Hackett et al., 2001). However, this effect was restricted to chromosome-end regions,
518 while our results show that telomere dysfunction increases mutation rates across the
519 whole genome.

520 After an initial decline, humanized yeasts recovered their fitness through two evolutionary
521 routes: by acquiring mutations in components of the DNA damage response or by
522 increasing the copy number of telomere-binding proteins. The mutations in DDR genes are
523 likely partial or complete loss-of-functions, as their appearance coincides with a decreased
524 expression of DDR targets like *RNR* genes, suggesting a release of cell-cycle blockage
525 induced by telomere uncapping. Previous studies showed that DDR is mildly but
526 chronically activated in humanized yeasts and causes a G_2/M delay (Di Domenico et al.,
527 2009), while the deletion of the DDR mediator *RAD9* abolished DDR and the cell cycle
528 delay (Bah et al., 2011), in line with our findings. While DDR inactivation relieves the cell
529 cycle blockage and rescues growth, it also increases genome instability. Our humanized
530 line 8, carrying a loss-of-function mutation in *TEL1*, exemplifies this tradeoff, as the
531 appearance of the mutation is followed by complete growth recovery coupled with multiple
532 chromosomal rearrangements, extensive subtelomeric amplifications and even whole-
533 genome duplication. Given its important role in DNA repair, DDR inactivation likely
534 constitutes a temporary solution while its long-term inactivation would probably lead to
535 genomic melt-down. A more likely scenario of telomere sequence evolution might involve

24

536 the transient inactivation of the DDR followed by the occurrence of compensatory
537 mutations in telomere-binding proteins to adapt to their novel telomeric sequences.
538 The second route that humanized yeasts use to rescue their fitness passes through the
539 duplication of chromosome XVI, ultimately to increase the abundance of Tbf1p, a T_2AG_3
540 binding protein (Brigati et al., 1993). While aneuploidy is generally considered detrimental,
541 it can be beneficial under different environments (Gilchrist & Stelkens, 2019; Pavelka et
542 al., 2010; Rancati et al., 2008). Our results are consistent with aneuploidy as a faster
543 strategy to recover fitness, due to the high chromosome duplication rate (Yona et al.,
544 2012). Even though MAL are generally considered a close-to-neutral evolutionary
545 scenario, adaptive aneuploidies have been previously observed in these settings, e.g.
546 chromosome XII is frequently duplicated in MAL propagated under rapamycin treatment to
547 maintain a balanced rDNA content after rapamycin-induced rDNA repeat contraction (Li et
548 al., 2023). However, aneuploidy also entails a fitness cost due to the indiscriminate
549 increase in expression of many genes and the stoichiometric imbalance of proteins
550 involved in protein complexes (Papp et al., 2003). Thus, the ultimate evolutionary outcome
551 is dictated by a tradeoff between the fitness costs and gains of the aneuploidy. Removing
552 the parts of the duplicated chromosome that are not driving the fitness improvement
553 constitutes a successful strategy, as exemplified in one of our AEL where the clones with
554 the shortest duplicated chromosome XVI segments were the fittest. The formation of
555 extrachromosomal circles carrying duplicated genes, as well as other chromosomal
556 rearrangements, has already been observed both in laboratory evolution and telomerase
557 negative mutants (Liti & Louis, 2003; Hull et al., 2017; Libuda & Winston, 2006). Our
558 results provide a temporal dynamic to the recent observation of complex aneuploidies
559 associated with structural variations, showing that aneuploidies precede structural
560 variations (O'Donnell et al., 2023). These results are also in line with a theoretical
561 framework in which the fitness cost of an aneuploidy is directly proportional to the number
562 of genes located on that chromosome and it is counterbalanced by the fitness benefit
563 resulting from the dosage increase of specific genes (e.g. *TBF1*) (Pompei & Cosentino
564 Lagomarsino, 2023). DDR mutations and *TBF1* duplication might ultimately impinge on the
565 same outcome: the attenuation of the DDR and release of cell cycle blockage. In fact, the
566 MRX complex binds critically short telomeres, recruits telomerase and activates the DDR
567 through the downstream action of Tel1p. Artificially placing Tbf1p-bound T_2AG_3 repeats in
568 subtelomeric regions shortens the respective telomeres, as Tbf1p inhibits MRX binding of
569 telomeres, thus preventing the initiation of the DDR, cell cycle blockage and telomerase
570 recruitment (Berthiau et al., 2006; Fukunaga et al., 2012; Ribaud et al., 2012). The

571 inhibition of telomerase recruitment by Tbf1p also explains the shorter T_2AG_3 tracts
572 detected in humanized yeasts carrying two copies of chromosome XVI.
573 Overall, our results point to an evolutionary scenario in which aneuploidy of chromosome
574 XVI and amplification of Tbf1p is the first adaptation to the modification of telomeric
575 sequences, due to its high frequency. However, given its fitness cost, it is removed as
576 soon as less burdening point mutations in the DDR and MRX complex appear. In contrast,
577 if the adaptive point mutations do not occur, another strategy is to remove the parts of the
578 duplicated chromosome XVI that are not immediately useful, only retaining duplicated
579 segments encompassing the *TBF1* gene. This scenario is in line with previous research
580 showing that aneuploidy and copy number variations are quick responses to
581 environmental stress whose role is to give time to the cells to develop more specific and
582 losg-lasting genetic changes in the form of point mutations (Yona et al., 2015).
583 T_2AG_3 repeats are thought to be the ancestral ones from which all the others evolved
584 (Fulnečková et al., 2013). Going back to an ancestral state might be easier than adapting
585 to completely new challenges. This view is consistent with the fact that Tbf1p is
586 functionally related to the telomere protective proteins TRFs in animals, and particularly in
587 its DNA binding domain interacting specifically with T_2AG_3 repeats (Bilaud et al., 1996).
588 Therefore, yeasts might behave differently if evolved with other telomeric sequences and
589 in different environmental conditions. Mutations in the telomerase RNA template cause
590 growth defects and telomere length dysregulation (Lin et al., 2004; McEachern &
591 Blackburn, 1995), but it is not known how these strains would survive and adapt when
592 propagated under selection for a long time. Recently, a variant telomeric repeat has been
593 identified in a human family and shown to be inherited across one generation, indicating
594 that telomere sequence modification is possible even in humans (Hinchie et al., 2024).
595 Moreover, telomeric sequences can be introgressed from other species following
596 horizontal gene transfer events, as exemplified by the discovery of mixed telomeric
597 sequences in *S. cerevisiae* natural isolates, constituted by centromere-proximal
598 *Torulaspora* telomeric repeats followed by native TG_{1-3} telomeric repeats (O' Donnell et al.,
599 2023). Future studies investigating the long-term impact of other telomeric repeats, as well
600 as their interactions with the external environment, will further elucidate how cells can
601 evolve dysfunction in key cellular processes.

26

Acknowledgements

We thank Simone Mozzachiodi for help with CRISPR/Cas9 genetic engineering, Jing Li and Jia-Xing Yue for help in generating and analyzing long-read sequencing data, and Daniela Ludwig, Kathrin Textoris-Taube and Michael Mülleder at the Core Facility - High Throughput Mass Spectrometry, Charité University Medicine, Berlin, for help with sample preparation and mass spectrometry analysis. We also thank Miguel Ferreira and Zhou Xu for critical reading of the manuscript. This study was supported by ANR (ANR-18-CE12-0004, ANR-20-CE12-0020, ANR-23-CE12-0031, ANR-24-CE12-7740), Fondation pour la Recherche Médicale (EQU202003010413), Fondation ARC (No. PJA32020070002320), Impulscience - Fondation Bettencourt Schueller to G.L, VR-2022-03024 to JW and the European Research Council (ERC) under grant agreement ERC-SyG-2020 951475 to M.R. M.D. was supported by PhD fellowships from the LABEX SIGNALIFE ANR-11-LABX-0028-01 and the Fondation pour la Recherche Médicale (FDT201904008453).

Contributions

M.D., B.B., E.G. and G.L. designed the experiments; M.D. performed the experiments and the genomic analyses; M.D.C. estimated copy number variants from short-read sequencing data; N.T. analyzed *TLC1* variation data; J.H. and M.D. generated and analysed the growth phenotype data; J.M., S.K. and M.D. generated and analysed proteomics data; J.W., M.R., E.G. and G.L. contributed with resources and reagents; E.G. and G.L. conceived and supervised the project; G.L. coordinated the project; M.D. and G.L. wrote the paper.

Data availability

The genome sequencing generated in this study are available at Sequence Read Archive (SRA), NCBI under accession codes BioProject ID PRJNA985049, Biosample ID SAMN35786841-SAMN35786899. The phenotyping and proteomics data are available at: https://github.com/mdangiolo89/Humanized_yeasts_project and within the supplementary information files. All the strains generated in this work are available upon request. Raw and processed mass spectrometry data have been deposited at the ProteomeXchange consortium via the PRIDE partner repository with the dataset identifier PXD064562 and will be publicly released upon publication of the study.

Code availability

YeaISTY (Yeast ITS, Telomeres and Y′ elements estimator) and other custom scripts are

27

available at: https://github.com/mdangiolo89/Humanized_yeasts_project.

Competing interest

The authors declare no competing interests.

28

637 ## Materials and methods

638 ### Yeast strains and plasmids

639 The two ancestor strains (one for the humanized and one for the wild-type lines) derive
640 from (Bah et al., 2004) and their genetic background is: RWY12 (BY4741 *Mat*a, *ura3-52*,
641 *lys2–801*, *ade2-101*, *trp1*-Δ*1*, *his3*-Δ*200*, *leu2*-Δ*1*, *tlc1*Δ*::LEU2,* VR-*ADE2*-T) (Dionne &
642 Wellinger, 1996). These strains are haploid and auxotrophic for uracil, lysin, tryptophan
643 and histidine. The native *ADE2* gene on chromosome XV is not functional and another
644 copy has been inserted into the right subtelomere of chromosome V. *ADE2* inactivation
645 results in red colonies through the accumulation of purine precursors in the vacuole and
646 can be used as a marker of transcriptional silencing at subtelomeric sites (Zonneveld &
647 Van Der Zanden, 1995). The replacement of the *TLC1* gene with an auxotrophic marker
648 ensures that the strains cannot use their native RNA template to elongate telomeres.
649 In addition, the two ancestors contain two centromeric plasmids: pAZ1 and p*TLC1*. The
650 pAZ1 plasmid derives from (Dionne & Wellinger, 1996) and contains the wild-type version
651 of *TLC1* and the selection marker *URA3*. The p*TLC1* plasmid backbone originally derives
652 from pRS314 in (Sikorski & Hieter, 1989) and contains the *TLC1* gene and the selection
653 marker *TRP1*. The *TLC1* gene in the p*TLC1* plasmid is different in the two ancestors: the
654 wild-type ancestor carries the native p*TLC1*-y gene while the humanized ancestor carries
655 a modified p*TLC1*-h gene in which the 16 template nucleotides (5'-
656 CACCACACCCACACAC-3') have been replaced by a humanized counterpart (5'-
657 CUAACCCU-3').

658 ### Mutation accumulation lines

659 The humanized and wild-type ancestors were revived from frozen stocks and plated on
660 Leucin/Tryptophan/Uracil dropout medium to select for cells containing both the pAZ1 and
661 p*TLC1* plasmids. Plates were incubated for 3 days at 30 °C. After that, we streaked for
662 single colonies on 5-FOA plates to retrieve cells that had lost the *URA3*-containing pAZ1
663 plasmid, and therefore rely only on the *TLC1* copy present in the p*TLC1*-y/p*TLC1*-h
664 plasmids. We selected 16 colonies derived from the wild-type ancestor and 16 derived
665 from the humanized ancestor, which contained only the p*TLC1*-y and p*TLC1*-h plasmids,
666 respectively. Since two streaking-for-single were performed during the pAZ1 exclusion
667 procedure, we dated this timepoint as 2 single cell bottlenecks (scb).
668 We propagated the 32 lines on Leucin/Tryptophan dropout solid medium in order to
669 maintain cells with *tlc1::LEU2* genotype and to retain the p*TLC1*-y/p*TLC1* at the final
670 timepoint (100 scb) and at four intermediate timepoints (20, 40, 60, 80 scb) were stored at

29

671 -80 °C. Moreover, we stored 4 humanized and 2 wild-type lines at 2 scb, immediately after
672 the loss of pAZ1.
673 To estimate the number of generations occurring every 72 h, three colonies derived from
674 three independent humanized lines at every timepoint and one colony from a wild-type line
675 at 2 and 100 scb were resuspended in 100 µl of sterile water and serially diluted. Twenty µl
676 of each dilution were plated on solid YPD medium and incubated for 72 h at 30 °C. The
677 number of colonies was manually counted in the plate with a suitable dilution and the
678 number of generations (G) was estimated according to $G=\log_2(n\times d)$, where n is the
679 number of cells on the plate and d is the corresponding dilution factor. The total number of
680 generations occurred during the MAL protocol for the wild-type and humanized lines (T_{wt}
681 and T_{hum}) were calculated as: $T_{wt}=G\times 100$ where G was averaged across the only
682 representative line at 2 and 100 scb; $T_{hum}=(G_{20}\times 20)+(G_{40}\times 20)+(G_{60}\times 20)+(G_{80}\times 20)+$
683 $(G_{100}\times 20)$ where G_x was averaged across the three representative lines at each timepoint.
684 The humanized lines at 2 scb were not considered for the calculation.

685 **Adaptive evolution**

686 The 4 humanized and 2 wild-type lines at 2 scb, as well as 5 humanized lines and 2 wild-
687 type lines derived from the MAL protocol (Hum1,2,3,8,11,WT2,7) were further submitted to
688 adaptive evolution (AEL) by multiple serial transfers (STs) of large population sizes.
689 Strains were inoculated in tubes containing 5 ml of Leucin/Tryptophan dropout liquid
690 medium and incubated at 30 °C in a shaking incubator. Serial transfers were performed
691 every 48h by diluting 50 µl of the culture into 5 ml of fresh Leucin/Tryptophan dropout
692 medium (1:100 dilution). We performed a total of 31 serial transfers.

693 **Estimation of the growth performance in liquid medium**

694 Strains were incubated two nights in 200 uL of liquid Leucin/Tryptophan dropout medium
695 in a non-shaking incubator at 30 °C. The day after, strains were mixed and 2 ul of the
696 mixed culture was resuspended in 200 uL of fresh Leucin/Tryptophan dropout liquid
697 medium in a 96-well plate. The plate was incubated in a TECAN plate reader machine
698 (Tecan, Infinite F200 Pro) at 30 °C for 72 hours and optical density (600 nm) was
699 measured every 15 minutes during the experiment. Resulting data were analyzed using
700 the software PRECOG (Fernandez-Ricaud et al., 2016). We phenotyped 3 biological
701 replicates for each strain. Given the high number of strains in this study, phenotyping was
702 performed using several 96-well plates. To correct for the batch effect and compare data
703 across plates, we included the humanized ancestor at 0 scb (MJD548) in all the plates and

704 used it as internal control. We normalized the data in each plate by dividing by the average
705 of the 3 control replicates in that plate. We then recalibrated all the measurements by
706 multiplying for constant values: 1.97 for population doubling time, 5.82 for lag duration and
707 3.20 for yield. These values represent the average of the 3 control replicates in the first
708 plate. Values reported in Table S5 and in the Fig.s are already recalibrated.

709 **Estimation of the growth performance across multiple conditions**

710 Strains were grown in 5 ml Synthetic Complete (SC) shaking liquid culture overnight. 100
711 uL from each of these cultures were resuspended into four 96-well plates, each with a
712 different randomized layout. These four 96-well plates were then transferred onto a 1536
713 SC PLOS agar plate using the Rotor HDA Robot (Singer ltd). Each randomized layout was
714 replicated 3 times onto the 1536 plate, resulting in a total of 12 replicates (24 for SC). Each
715 plate also contained 384 positions with spatial controls (Zackrisson et al., 2016). The
716 spatial control used was a wild-type ancestor at 2 scb (MJD5). Due to the size differences
717 of the colonies when transferring from liquid to solid, the 1536 solid plate was grown for 3
718 days and then replica-plated onto a new SC agar plate. This new plate is the preculture
719 plate. After 3 days of growth on the preculture plate, the strains were transferred onto the
720 plates with the different stressors and phenotyped. Lag duration, population doubling time
721 and yield were extracted from the resulting growth curves using Scanomatic (version 2.2).
722 Lag was calculated as the intercept between the tangent of the initial colony size (average
723 of the first three timepoints) and the tangent from the timepoints used to calculate the
724 growth rate. Population doubling time was calculated at the steepest incline of the growth
725 curve, using five timepoints for a local regression. Yield was calculated as the difference
726 between the initial colony size, as calculated for lag, and the final colony size (average of
727 the last three timepoints). All measurements were calculated from the smoothened growth
728 curves.

729 **Estimation of survival rate**

730 Strains were grown two nights in liquid Leucin/Tryptophan dropout medium, diluted 100x in
731 200 µL of fresh Leucin/Tryptophan dropout in a 96-well plate and incubated at 30 °C for 15
732 days. We measured viability at the beginning of the stationary phase when cultures
733 reached saturation (3 days), and then every 4 days until day 15. Data are shown by days
734 post-saturation, meaning that day 0 post-saturation corresponds to day 3 after the
735 inoculation of the culture, while day 12 post-saturation corresponds to day 15 post-
736 inoculation. At each time point, 5 µL of cells were transferred in 100 µL of staining solution

737 (Phosphate-buffered saline-PBS + 3 µM propidium iodide + 200 nM YO-PRO-1) in a 96-
738 well plate and incubated 10 minutes in the dark at 30 °C. Cell viability was measured by
739 high-throughput flow cytometry on a Cytoflex machine, using the plate reader module.
740 Cells were excited with the 561nm Yellow/Green laser. Fluorescence is read with the
741 Y610-mCHERRY filter for propidium iodide and with the B610-ECDA filter for YO-PRO-1.
742 Propidium iodide and YO-PRO-1 are membrane-impermeable nucleic acid binding
743 molecules that can only enter into dead cells whose cellular membrane is damaged. Non-
744 fluorescent cells were counted as viable whereas fluorescent cells were counted as dead.
745 We measured one biological replicate with 3000 events per sample. Flow cytometry data
746 were subsequently analysed using the CytExpert software.

747 **Ploidy estimation and pulsed-field gel electrophoresis**

748 Ploidy was estimated for MAL at 2 and 100 scb and after AEL (at 31 st) using a propidium
749 iodide (PI) staining assay. Ploidy was estimated at all timepoints during the MAL protocol
750 for humanized line 8, which showed a diploid genome content at 100 scb. Cells were
751 incubated in 100 uL of Leucin/Tryptophan dropout liquid medium in a 96-well plate and left
752 two nights at 30 °C. After that, 3 µl were incubated in 100 µl of cold 70% ethanol for 3 h at
753 4 °C, washed twice with PBS, and resuspended in 100 µl of staining solution (15 µM PI,
754 100 µg/ml RNase A, 0.1% v/v Triton-X, in PBS). After that, samples were incubated for 3 h
755 at 37 °C in the dark. Samples were analyzed on a FACS-Calibur flow cytometer using the
756 HTS module for processing 96-well plates. Cells were excited at 488 nM and fluorescence
757 was collected with a FL2-A filter. The distributions of both FL2-A and FSC-H values were
758 processed to find the two main density peaks, which correspond to the two cell populations
759 in G1 and G2 phase of the cell cycle, as described in (Peter et al., 2018).
760 Whole chromosome size was estimated for MAL at 2 and 100 scb through Pulsed-Field
761 Gel Electrophoresis, as described in (Louis, 1998).

762 **Generation of *MRE11* and *TBF1* mutants**

763 The deletion of *MRE11* was performed by standard homologous recombination. The
764 *URA3*-containing cassette was amplified from the laboratory S288C strain carrying a
765 functional *URA3* gene. 80-bp primers used for the amplification were designed to contain
766 the first and last 20 bp of *URA3* followed by 60 bp of homology with flanking regions
767 upstream and downstream of *MRE11*. Samples were transformed using the standard
768 lithium-acetate protocol, plated on selective medium lacking uracil and incubated at 30 °C
769 for 3–7 days. Candidate transformed clones were validated by PCR using one primer

770 designed inside *URA3* and another primer designed on the outside region of *MRE11*.
771 The deletion of one copy of *TBF1* in humanized strains carrying two chromosome XVI
772 copies was performed by standard homologous recombination. The *HYGMX*-containing
773 cassette was amplified from a plasmid carrying this antibiotic marker. 80-bp primers used
774 for the amplification were designed to contain 20 bp of the MX region followed by 60 bp of
775 homology with flanking regions upstream and downstream of *TBF1*. Samples were
776 transformed using the standard lithium-acetate protocol, plated on selective medium
777 containing hygromycin (200 ug/mL) and incubated at 30 °C for 3–7 days. Candidate
778 transformed clones were validated by PCR using a combination of two PCR reactions: one
779 reaction carried one primer designed inside *HYGMX* and another primer designed on the
780 outside region of *TBF1*; the other reaction carried both primers designed inside *TBF1*.
781 Only clones positive for both reactions were selected, as we wanted to delete only one
782 copy of *TBF1*.
783 The introduction of *MRE11* Q447P, as well as the introduction of *TBF1* in the *HO* locus,
784 were performed by using CRISPR/Cas9 genome editing. The plasmid harboring Cas9 was
785 obtained from Addgene pUDP004 and linearized with BsaI. The resistance to acetamide
786 was replaced with the resistance cassette to hygromycin which was amplified from a
787 plasmid harboring the *HYGMX* gene. The gRNAs with the necessary nucleotides for self-
788 cleavage were designed using the software UGENE and ordered as synthetic oligos from
789 Eurofins Genomics (TM). Plasmid constructions were performed as described in
790 (Mozzachiodi et al., 2022). The cassette used for the engineering of *MRE11* Q447P was
791 designed to overlap the mutation site. The cutting site for Cas9 was modified to make it
792 unrecognizable, by replacing one nucleotide with another one that did not change the
793 aminoacid of the corresponding codon. The single-stranded parts of this cassette were
794 ordered as a unique synthetic oligo at Eurofins Genomics. They were subsequently mixed
795 at equimolar ratio, heated at 95 °C for 15 minutes and cooled down at room temperature to
796 be used for the transformation. The cassette used for the knock-in of *TBF1* into the *HO*
797 locus was amplified from the humanized ancestor at 0 scb (MJD548). 80-bp primers used
798 for the amplification were designed to contain the first and last 20 bp of *TBF1* followed by
799 60 bp of homology with flanking regions upstream and downstream of *HO*. Yeast samples
800 were transformed using the standard lithium-acetate protocol. Cells were plated on
801 selective medium containing hygromycin (200 µg/mL) and incubated at 30 °C for 3–7
802 days. *MRE11* Q447P candidate transformed clones were validated by Sanger sequencing
803 of selected colonies. *ho::TBF1* candidate transformed clones were validated by PCR using
804 one primer designed inside *TBF1* and another primer designed on the outside region of

33

805 *HO*. Positive clones were streaked on Leucin/Tryptophan dropout medium and grown for 2
806 days at 30 °C to allow plasmid loss and stored at -80 °C in 25% glycerol stocks.

807 **Quantitative PCR to track chrXVI copy number**

808 To track the occurrence of chrXVI aneuploidy during MAL, we extracted DNA from yeast
809 cells at various timepoints during the MAL protocol using a MasterPure Yeast DNA
810 purification kit (Epicentre), following the manufacturer's instructions. Primers were
811 designed for the *TBF1* locus and for one control locus (*PWP2*) on chromosome III, since
812 this chromosome never underwent aneuploidy in any line. To determine the copy number
813 of chrXVI blocks in 8 single clones derived from Hum11 at 31 st (MJD460), we extracted
814 DNA from the clones and designed one pair of primers for each block. Primers were
815 designed on the following loci: *FUM1*-Block1; *TBF1*-Block2; *RPC40*-Block 3; *NUT2*-
816 Block4. We also used the same control locus on chrIII (*PWP2*). DNA was diluted to 8
817 ng/uL for all the samples and the qPCR was performed using iTaq Universal SYBR Green
818 Supermix (total volume: 10 µL) and run on an Applied Biosystems StepOne machine using
819 the comparative ΔCt method. The PCR protocol was: initial denaturation (95 C; 15 min)
820 followed by 45 cycles of: denaturation (95 C; 15 s), annealing (60 C; 30 s), extension (72
821 C; 30 s). We performed two replicates per sample per pair of primers. We calculated the
822 copy number of the chrXVI regions covered by the primers by first subtracting the Ct of the
823 chrIII control to all the primer pairs (ΔCt). We further normalized each primer pair's ΔCt by
824 subtracting the corresponding ΔCt of an internal control strain that has one copy of each
825 region (humanized ancestor at 0 scb-MJD548). We calculated the copy number (CN) of
826 each region as $CN=2^{-\Delta\Delta Ct}$. In addition to the investigated samples, we also included a
827 humanized line aneuploid for the full length of chromosome XVI (Hum3-MJD420) and the
828 humanized ancestor at 0 scb (MJD548) that has one copy of each chromosome, as
829 positive and negative controls, respectively.

830 **DNA extraction and short-read sequencing**

831 Cells were inoculated in 5 ml of liquid Leucin/Tryptophan dropout medium and grown
832 overnight at 30 °C in a shaking incubator. DNA was extracted using "Yeast Masterpure" kit
833 (Epicentre, USA) following the manufacturer's instructions. Illumina paired-end libraries
834 (2×150 bp) were prepared according to manufacturer's standard protocols and sequenced
835 with an HiSeq 2500 instrument, at the NGS platform of Institut Curie. For the MAL
836 experiment, two initial strains at 2 scb (one with the humanized background and one with
837 the wild-type background) and all the strains at 100 scb were sequenced. For the AEL

838 experiment, all the 5 humanized and 2 wild-type lines at 31 ST were sequenced. We
839 additionally sequenced the humanized and wild-type ancestors at 0 scb, plus 8 single
840 clones derived from Hum11 at 31 st. DNA from these 10 samples was extracted as
841 described above but sequencing was performed by Novogene (Cambridge, UK).

842 **Long-read sequencing and structural variant analysis**

843 Cells from a humanized line at 2 scb (Hum1), 4 humanized lines at 100 scb (Hum1,4,8,15)
844 and two single clones derived from a AEL started from 100 scb (Hum11 at 31 st) were
845 grown for two nights in liquid Leucin/Tryptophan dropout medium at 30 °C. Genomic DNA
846 was extracted using "Yeast Masterpure" kit (Epicentre, USA) according to the
847 manufacturer's instructions. No vortexing was performed during the DNA extraction to
848 avoid excessive DNA fragmentation. The correct size distribution of DNA fragments after
849 the DNA extraction was verified by gel electrophoresis using the lambda DNA as
850 molecular weight size marker. The MINION (Oxford Nanopore) sequencing library was
851 prepared using the SQK-LSK109 sequencing kit according to the manufacturer's
852 instructions. The library was loaded onto a FLO-MIN106 flow cell and sequencing was run
853 for 72 hours. Long read basecalling, demultiplexing and scaffolding were performed using
854 the pipelines LRSDAY and LRSDAY_Patch (https://github.com/nicolo-tellini/LRSDAY-
855 Patch) and the assembler was set as "canu" (Yue & Liti, 2018). Dotplots were generated
856 using *nucmer* and *mummerplot* (Marçais et al., 2018). Structural variants were further
857 confirmed by manually inspecting reads encompassing the breakpoints using IGV
858 (v2.3.68) (Thorvaldsdottir et al., 2013).

859 **Short-read data preprocessing and CNVs estimation**

860 We mapped Illumina reads resulting from MAL and AEL on the *S. cerevisiae* SGD
861 reference genome assembly using *bwa mem* (v0.7.12) (Li & Durbin, 2009). Read mapping
862 runs were stored in the Sequence Alignment/Map (SAM) format and its binary form (BAM)
863 and post-processing steps were performed by picard tools (v2.8.0). Sorting, indexing and
864 per-position coverage calculation were performed using SAMtools (v1.2) (Li et al., 2009).
865 Regions with higher coverage values as compared to the genome-wide median were
866 considered as copy number variants and their effective copy number was calculated as the
867 ratio between their median coverage and the median genome-wide coverage for the strain
868 in which they were observed. The ORFs copy number was estimated as in (Peter et al.,
869 2018). Reads from each strain were mapped against the reference S288C genome from

870 the SGD database using *bwa mem* (Li & Durbin, 2009), to obtain the baseline coverage.
871 Then, the same reads were mapped against the pangenome file from (Peter et al., 2018)
872 using *bwa mem* and the option *-U 0*, and the median coverage was extracted for each
873 pangenomic ORF. Coverages from the bam files were extracted using a custom script
874 using the R library *Rsamtools*.
875 We used the software freebayes (v0.9.5) to call markers from the BAM files. Results were
876 stored in Variant Call Format (VCF) files. We filtered the VCF files to keep only
877 SNVs/indels with quality higher than 20 and coverage higher than 10.

878 **Identification of *de novo* SNVs in MAL and AEL**

879 For MAL, we subtracted the SNVs/indels in the humanized strain at 2 scb (MJD1) from the
880 humanized lines at 100 scb, and the SNVs/indels in the wild-type strain at 2 scb from the
881 wild-type lines at 100 scb, using the vcf-isec program included in the suite VCFtools
882 (v0.1.16). For AEL, we subtracted the SNVs/indels in the corresponding mutation
883 accumulation lines at 100 scb from their derived AEL. Next, we searched for shared SNVs
884 present in at least two MAL at 100 scb (or in two AEL for the AEL experiment) and
885 removed them from the VCF files, based on the assumption that identical *de novo* SNVs
886 are highly unlikely to arise independently and derive from incorrect mapping or poor
887 coverage. All the remaining SNVs/indels were manually validated using IGV (v2.3.68)
888 (Thorvaldsdottir et al., 2013). SNVs were annotated using the *S. cerevisiae* SGD genome
889 assembly implemented in the online software Variant Effect Predictor (VEP) (McLaren et
890 al., 2016). Further annotation of the functional impact of SNVs was carried out using
891 mutfunc (Wagih et al., 2018). We tested whether selection during MAL might have affected
892 base substitution rates. The proportion of observed vs expected coding SNVs (74 *vs* 75 %)
893 and non-synonymous SNVs (68 *vs.* 75 %) in wild-types supports a close-to-neutral
894 mutation accumulation scenario (binomial test: p=ns), while the corresponding proportions
895 in humanized yeasts (59 and 76 % for coding and non-synonymous mutations,
896 respectively) indicate an effect of selection in this group, with mutations less likely to occur
897 in coding regions (binomial test: $p<0.0001$). *De novo* variants derived from MAL were used
898 to estimate base substitution and insertion/deletion rates (μ) as $\mu=(n/bp)/g$, where n is the
899 number of *de novo* SNVs present in each line, bp is the (haploid) genome size of the *S.*
900 *cerevisiae* SGD genome assembly and g is the number of generations performed in the
901 mutation accumulation experiment (2140 for wild-type lines and 1640 for humanized lines).

902 **Experimental validation of *de novo* SNVs in MAL and AEL**

903 Two SNVs from humanized MAL were validated by Sanger sequencing. One variant was
904 located inside the gene *MEC1* (Hum4: *MEC1* R1824H) and the other was inside the gene
905 *TEL1* (Hum8: *TEL1* L2308R). A pair of primers (upstream and downstream) was designed
906 for each SNV using Unipro UGENE (Okonechnikov et al., 2012), and used to perform
907 Polymerase Chain Reaction (PCR). We checked the presence of the mutations at all
908 timepoints during the MAL protocol. PCR products were sequenced by Eurofins
909 Genomics. The presence and the genotype of the variants were checked by visual
910 inspection of the electropherograms. We confirmed the presence of both variants and
911 discovered that the mutations in *TEL1* occurred between 60 and 80 scb, while the
912 mutation in *MEC1* occurred between 40 and 60 scb. We used the same procedure to
913 confirm the presence of *MRE11* and *RAD50* mutations in AEL started from 100 scb
914 (Hum1: *MRE11* K477I; Hum2: *MRE11* Q447P; Hum3: *RAD50* L171P).

915 **Analysis of *TLC1* variation**

916 To analyze the extent of genetic variation in the *TLC1* template and its flanking regions
917 across the *Saccharomyces sensu stricto* species and across the *S. cerevisiae* population,
918 we first downloaded published *TLC1* sequences from (Liti et al., 2006), spanning several
919 *S. cerevisiae* and *S. paradoxus* isolates. We then downloaded the *TLC1* sequence from
920 the *Saccharomyces* Genome Database (SGD) and used it to search for orthologs in other
921 *Saccharomyces sensu stricto* species in the NCBI database, using the *blastn* algorithm.
922 We then performed a multiple sequence alignment using *clustal omega*
923 (https://www.ebi.ac.uk/jdispatcher/msa/clustalo). Next, we gathered publicly available
924 Illumina sequencing datasets relative to 3,139 *S. cerevisiae* strains, and further sequenced
925 241 new *S. cerevisiae* isolates (unpublished data). The latter were sequenced at The
926 Earlham genomics facilities (www.earlham.ac.uk) according to the LITE pipeline (Perez-
927 Sepulveda et al., 2021). We run whole-genome short-reads alignments with *bwa mem*
928 (v.0.7.17-r1198-dirty) (Li, 2013) and processed the bam files according to the latest
929 samtools guidelines (samtools fixmate, sort, markdup and view) (v1.19). For each isolate,
930 we produced a gVCF file using bcftools (mpileup, call, view and annotate) (v1.19) and
931 extracted the genotypes for the 16 nt template (S288C v.R64-3-1_20210421,
932 chrII:308054-308069) and 20 nt upstream and downstream this sequence (chrII:308034-
933 308053, chrII:308070-308089). We converted the gVCF genotypes in fasta sequences
934 concatenating the alleles across consecutive positions with vcf2phylip.py (v.2.8) and
935 created a multifasta file. The heterozygous positions were resolved according to the
936 IUPAC nomenclature. At the end, we extracted the unique sequences, their occurrences

937 and the list of isolates sharing the unique sequences with seqkit rmdup (v.2.4.0) (W. Shen
938 et al., 2016). The output was parsed with custom R scripts.

939 **Telomere length, ITS content and Y' copy number estimation from short-read**
940 **sequencing**

941 Telomere length, ITS content and copy number of Y' elements were determined using
942 Y^{ea}ISTY (D'Angiolo et al., 2023). Telomere length and ITS content were measured
943 separately for the TG_{1-3} and T_2AG_3 motifs. For TG_{1-3} we used the published version of
944 Y^{ea}ISTY, while we developed a new version compatible with the T_2AG_3 motif for this study.
945 We used a threshold of 20 bp of telomeric repeats stretch length for TG_{1-3}, and a threshold
946 of 18 bp for T_2AG_3. By comparing our results with the estimates derived from long-read
947 genome assemblies, we determined that our lengths were underestimated and the
948 underestimation was particularly pronounced for the T_2AG_3 repeat. This is likely due to the
949 fact that the T_2AG_3 motif is longer and more complex than TG_{1-3} and sequencing errors are
950 more prone to disrupt the continuity of T_2AG_3 stretches. As a consequence, we avoid direct
951 comparison of total telomere lengths (TG_{1-3}+T_2AG_3) between humanized and wild-type
952 yeasts.

953 **Chromosome-ends annotations from long-read genome assemblies**

954 X and Y' elements were annotated using LRSDAY (Yue & Liti, 2018). Elements annotated
955 as "partial" were further inspected by extracting the corresponding and surrounding
956 genomic regions (~10 kb) with *bedtools getfasta* and by performing a *blastn* local
957 alignment with the SGD reference genome. Most of the partial elements were instead full
958 versions and were re-annotated using the coordinates derived from the local alignment.
959 We annotated TG_{1-3} and T_2AG_3 repeats in the genome assemblies by using custom Perl
960 scripts. We merged the repeats which were interspaced by only 1 bp into a unique stretch
961 by using the software mergeBed included in the suite bedtools (version 2.17.0) with the
962 option "-d 1" (Quinlan & Hall, 2010). Then, we filtered the list to keep only the stretches
963 longer than 10 bp. The interstitial stretches were annotated as ITS if they closely preceded
964 a Y' element, while the stretches at the termini were annotated as telomeres. We further
965 mapped long reads derived from each strain onto its corresponding genome assembly
966 using *bwa mem* (v0.7.12) and manually visualized long reads encompassing telomeres
967 using IGV (v2.3.68) (Thorvaldsdottir et al., 2013).

968 **Mass spectrometry and proteomics data analysis**

969 *Cultivation of library:* strains were randomized manually in triplicates onto 96-well plates,
970 with the exception of strains MJD436-447, MJD450-455, MJD481-500, and MJD586-597,
971 each of which was cultivated once. Strains were pinned onto solid Leucine/Tryptophan
972 dropout agar using a Singer Rotor and incubated at 30 °C for 48 h. Subsequently, each
973 colony was pinned into 200 µL of Leucin/Tryptophan liquid dropout medium and the
974 samples were incubated at 30 °C without shaking. After 32 h, these pre-cultures were
975 mixed thoroughly by pipetting and 80 µL of each sample were added to 1.55 mL of pre-
976 warmed Leucin/Tryptophan liquid dropout medium in 96-well deep-well plates. One
977 sterilized borosilicate glass bead was added per well in order to aid mixing. Plates were
978 sealed with a Breathe-Easy membrane and incubated for 14.5 h at 800 rpm, 30 °C. Then,
979 1.4 mL of culture were transferred to a new deep-well plate, and cells were harvested by
980 centrifugation (2880 xg, 5 min, 4 °C). The optical density at harvest was measured using
981 an Infinite M Nano plate reader (Tecan). The supernatant after centrifugation was removed
982 by inverting the plate, and cell pellets were immediately cooled on dry ice. Cell pellets
983 were stored at -80 °C until lysis.
984 *Proteomic sample preparation*: samples for proteomics were prepared in a 96-well high-
985 throughput format as previously described (Muenzner et al., 2024). Briefly, cells were
986 lysed in 200 µL 100 mM ammonium bicarbonate, 7 M urea by beat beating using a
987 Geno/Grinder (Spex). Subsequently, samples were reduced and alkylated using DTT (20
988 µL, 55 mM) and iodoacetamide (20 µL, 120 mM), respectively. Samples were then diluted
989 by adding 460 µL of mL 100 mM ammonium bicarbonate. 500 µL per sample were
990 digested using 2 µg Trypsin/LysC (Promega, Cat#V5072). After 17 h of incubation at 37
991 °C, 25 µL 20% formic acid were added to the samples, and peptides were purified using
992 solid-phase extraction as described before (Messner et al., 2023). Eluted samples were
993 vacuum-dried and subsequently reconstituted in 90 µL 0.1% formic acid. An equivoluminal
994 pool of all samples (5 µL per sample) was generated to be used as technical controls
995 during MS measurements. Peptide concentrations of samples and the pool were
996 determined using a fluorimetric peptide assay kit (Thermo Scientific, Cat#23290). Samples
997 with a too low optical density at harvest (OD600 < 0.25) and/or too low peptide
998 concentrations (no peptide signal observed) were excluded from mass spectrometric
999 measurements.
1000 *LC–MS/MS measurements*: samples were analyzed in two batches (96 samples per batch
1001 plus analytical quality control (pool) injected every ten samples) on a SCIEX ZenoTOF
1002 7600 System mass spectrometer coupled to a Waters ACQUITY UPLC M-Class System.
1003 Prior to MS analysis, 200 ng peptides were chromatographically separated with a 60 min

39

1004 gradient on a Waters HSS T3 column (300um x 150mm, 1.8um) heated to 35 °C, using a
1005 flow rate of 5 µL/min where mobile phase A and B were 0.1% formic acid in water and
1006 0.1% formic acid in acetonitrile, respectively. For data independent acquisition, a
1007 zenoSWATH MS/MS acquisition scheme with 85 variable size windows and 11 ms
1008 accumulation time was used. Ion source parameters were set to: ion source gas 1 & 2 –
1009 12 and 60 psi respectively; curtain gas 25, CAD gas 7 and source temperature at 150°C;
1010 spray voltage – 4500V.
1011 *Data pre-processing*: raw data were processed with DIA-NN (version 1.8, available from
1012 https://github.com/vdemichev/DiaNN) (Demichev et al., 2020) using the following settings:
1013 fragment ion m/z range set from 300-1800, mass accuracy set to 20 ppm at the MS2 and
1014 MS1 level, respectively, scan window set to 6, MBR (Match-between-runs) enabled, and
1015 Quantification strategy set as "Robust LC (high Precision)". A spectral library-free
1016 approach and yeast reference proteome (UniProt UP000002311, accessed on
1017 23.04.2021) were used for annotation. The resulting report.tsv file was further pre-
1018 processed at the precursor level in R. First, all non-proteotypic peptides were excluded
1019 and precursors with q.values, PG.q.values, Global.q.values, Global.PG.q.values, or
1020 GG.q.values > 0.01 were excluded. Next, samples with very low precursor numbers (Z-
1021 score < -2.5) were removed. Batch correction was performed by equalizing the median
1022 precursor quantity of each batch. Protein quantities were determined for all proteins with
1023 two or more precursors mapped using the diann_maxlfq() function of the DIA-NN R
1024 package (https://github.com/vdemichev/diann-rpackage), resulting in a dataset covering
1025 3503 proteins across 181 samples. No imputation was applied to the proteomic dataset
1026 and undetected proteins were treated as missing values.
1027 *Data analysis*: after these pre-processing steps, we calculated the median of the three
1028 technical replicates and normalized them using the wild-type strain 1 at 2 scb (MJD5) as a
1029 reference, to quantify the relative changes. The quality of the technical replicates was
1030 confirmed by the high Pearson's correlation coefficients (>0.98) between each pair. To
1031 identify genes that are differentially expressed as compared to the reference, we
1032 calculated the fold-changes and the p-values corresponding to each gene using the edgeR
1033 library in R. The normalization factors for differences in measurement depth across
1034 different samples were calculated using the function 'calcNormFactors'. Dispersion factors
1035 were estimated using the functions 'estimateCommonDisp', 'estimateTagwiseDisp' and
1036 'estimateTrendedDisp' and we quantified the dispersion of the gene expression based on
1037 the mean values. The function 'exactTest' was used to identify significant differential
1038 expressions, which were then visualized as volcano plots highlighting significant fold-

1039 changes in expression using the 'EnhancedVolcano' function from the 'EnhancedVolcano'
1040 library. The Principal Component Analysis was performed using the function 'fviz_pca_ind'
1041 from the library 'factoextra' in R.
1042 Next, we derived lists of differentially abundant proteins for each strain by selecting those
1043 showing a $\log_2$FC>=1 or $\log_2$FC<=-1 and p<=0.05. Each list was further filtered on the
1044 basis of the aneuploidies carried by each strain, in order to retain only proteins that are
1045 differentially abundant due to upregulated/downregulated expression and not by copy
1046 number modifications of their encoding genes located on duplicated chromosomes. Once
1047 we produced a list for each strain, we derived lists of differentially abundant proteins for
1048 humanized and wild-type yeasts by group (MAL at 40 and 100 scb, AEL from 2 and 100
1049 scb) by selecting those present in >40% of the strains in each group and combined them,
1050 deriving 110 proteins whose abundance changed in at least one group as compared to the
1051 reference (**Fig. 3d**). In contrast, the list of proteins defining the THPR was derived by
1052 combining the proteins that are differentially abundant in at least 2 humanized MAL at 40
1053 or 100 scb. The THPR was compared with a set of genes that are differentially expressed
1054 upon *cdc13* temperature-sensitive mutation and were downloaded from (Greenall et al.,
1055 2008), as well as with a list of Rap1p gene targets. The list of Rap1p targets was
1056 downloaded from SGD and derived from multiple studies (Brindle et al., 1990; Duveau et
1057 al., 2021; Hu et al., 2007; Lickwar et al., 2012; MacIsaac et al., 2006; McNeil et al., 1990;
1058 Platt et al., 2013, Song et al., 2020, Stanway et al., 1989).

1059 **GO term analysis**

1060 Standard GO term analysis was performed separately on the variants extracted from wild-
1061 type and humanized MAL. Due to the small number of *de novo* variants occurred during
1062 AEL, we combined the variants in AEL from 2 and 100 scb and performed GO term
1063 analysis separately on the variants extracted from wild-type and humanized AEL. GO term
1064 analysis for *de novo* mutations was performed using the GO Term Finder tool available at
1065 *Saccharomyces* Genome Database (SGD). Significant GO terms were extracted by the
1066 algorithm implemented in the tool, with a False Discovery Rate (FDR) corrected α
1067 threshold of 0.1. GO term analysis for differentially abundant proteins was performed using
1068 the *gprofiler2* package in Rstudio (Kolberg et al., 2020).

41

References

Alexander, M. K., Zakian, V. (2003). Rap1p telomere association is not required for mitotic stability of a C3TA2 telomere in yeast. *The EMBO Journal*, *22*(7), 1688–1696. https://doi.org/10.1093/emboj/cdg154

Auriche, C., Di Domenico, E. G., Ascenzioni, F. (2008). Budding yeast with human telomeres: A puzzling structure. *Biochimie*, *90*(1), 108–115. https://doi.org/10.1016/j.biochi.2007.09.009

Bah, A., Bachand, F., Clair, E., Autexier, C., Wellinger, R. J. (2004). Humanized telomeres and an attempt to express a functional human telomerase in yeast. *Nucleic Acids Research*, *32*(6), 1917–1927. https://doi.org/10.1093/nar/gkh511

Bah, A., Gilson, E., Wellinger, R. J. (2011). Telomerase is required to protect chromosomes with vertebrate-type T2AG3 3′ ends in Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *286*(31), 27132–27138. https://doi.org/10.1074/jbc.M111.220186

Berthiau, A. S., Yankulov, K., Bah, A., Revardel, E., Luciano, P., Wellinger, R. J., Géli, V., Gilson, E. (2006). Subtelomeric proteins negatively regulate telomere elongation in budding yeast. *The EMBO Journal*, *25*(4), 846–856. https://doi.org/10.1038/sj.emboj.7600975

Brevet, V., Berthiau, A. S., Civitelli, L., Donini, P., Schramke, V., Géli, V., Ascenzioni, F., Gilson, E. (2003). The number of vertebrate repeats can be regulated at yeast telomeres by Rap1-independent mechanisms. *The EMBO Journal*, *22*(7), 1697–1706. https://doi.org/10.1093/emboj/cdg155

Brigati, C., Kurtz, S., Balderes, D., Vidali, G., Shore, D. (1993). An essential yeast gene encoding a TTAGGG repeat-binding protein. *Mol Cell Biol*, *13*(2):1306-14. https://doi.org/10.1128/mcb.13.2.1306-1314.1993

Brindle, P. K., Holland, J. P., Willett, C. E., Innis, M. A., Holland, M. J. (1990). Multiple factors bind the upstream activation sites of the yeast enolase genes ENOJ and EN02: ABFI protein, like Repressor Activator Protein RAP1, binds cis-acting sequences which modulate repression or activation of transcription. *Mol Cell Biol,* 10(9):4872-85. https://doi.org/10.1128/mcb.10.9.4872-4885.1990

Cavalcanti, R. O., Dunn, D. M., Weiss, R., Herrick, G., Landweber, L. F., Doak, T. G. (2004). Sequence features of Oxytricha trifallax (class Spirotrichea) macronuclear telomeric and subtelomeric sequences. *Protist 155*(3):311-22. https://doi.org/10.1078/1434461041844196

Chan, S. W. L., Blackburn, E. H. (2002). New ways not to make ends meet: telomerase, DNA damage proteins and heterochromatin. *Oncogene*, *21*(4), 553–563. https://doi.org/10.1038/sj.onc.1205082

D'Angiolo, M., Yue, J. X., De Chiara, M., Barré, B., Giraud Panis M. J., Gilson, E., Liti, G. (2023). Telomeres are shorter in wild Saccharomyces cerevisiae isolates than in domesticated ones. *Genetics,* 2;223(3):iyac186. https://doi.org/10.1093/genetics/iyac186

De Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes & Development*, *19*(18), 2100–2110. https://doi.org/10.1101/gad.1346005

Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S., Ralser, M. (2020). DIA-NN: Neural networks and interference correction enable deep proteome coverage in high throughput. *Nature Methods*, *17*(1), 41–44. https://doi.org/10.1038/s41592-019-0638-x

Di Domenico, E. G., Auriche, C., Viscardi, V., Longhese, M. P., Gilson, E., Ascenzioni, F. (2009). The Mec1p and Tel1p checkpoint kinases allow humanized yeast to tolerate chronic telomere dysfunctions by suppressing telomere fusions. *DNA Repair*, *8*(2), 209–218. https://doi.org/10.1016/j.dnarep.2008.10.005

Di Domenico, E. G., Mattarocci, S., Cimino-Reale, G., Parisi, P., Cifani, N., D'Ambrosio, E., Zakian, V. A., Ascenzioni, F. (2013). Tel1 and Rad51 are involved in the maintenance of telomeres with capping deficiency. *Nucleic Acids Research*, *41*(13), 6490–6500. https://doi.org/10.1093/nar/gkt365

Dionne, I., Wellinger, R. J. (1996). Cell cycle-regulated generation of single-stranded G-rich DNA in the absence of telomerase. *Proceedings of the National Academy of Sciences*, *93*(24), 13902–13907. https://doi.org/10.1073/pnas.93.24.13902

Duveau, F., Vande Zande, P., Metzger, B. P., Diaz, C. J., Walker, E. A., Tryban, S., Siddiq, M. A., Yang, B., Wittkopp, P. J. (2021). Mutational sources of trans-regulatory variation affecting gene expression in Saccharomyces cerevisiae. *eLife*, *10*, e67806. https://doi.org/10.7554/eLife.67806

Fernandez-Ricaud, L., Kourtchenko, O., Zackrisson, M., Warringer, J., Blomberg, A. (2016). PRECOG: A tool for automated extraction and visualization of fitness components in microbial growth phenomics. *BMC Bioinformatics*, *17*(1), 249. https://doi.org/10.1186/s12859-016-1134-2

Forstemann, K., Lingner, J. (2001). Molecular basis for telomere repeat divergence in budding yeast. *Mol Cell Biol*, *21*(21):7277-86. https://doi.org/10.1128/MCB.21.21.7277-7286.2001

Fourel, G., Revardel, E., Koering, C. E., Gilson, E. Cohabitation of insulators and silencing elements in yeast subtelomeric regions. *EMBO J*, *18*(9):2522-37. https://doi.org/10.1093/emboj/18.9.2522

Fukunaga, K., Hirano, Y., Sugimoto, K. (2012). Subtelomere-binding protein Tbf1 and telomere-binding protein Rap1 collaborate to inhibit localization of the Mre11 complex to DNA ends in budding yeast. *Molecular Biology of the Cell*, *23*(2), 347–359. https://doi.org/10.1091/mbc.e11-06-0568

Fulnečková, J., Ševčíková, T., Fajkus, J., Lukešová, A., Lukeš, M., Vlček, Č., Lang, B. F., Kim, E., Eliáš, M.,

42

Sýkorová, E. (2013). A broad phylogenetic survey unveils the diversity and evolution of telomeres in eukaryotes. *Genome Biology and Evolution*, *5*(3), 468–483. https://doi.org/10.1093/gbe/evt019

Garrido, C., Kornobis, E., Agier, N., Ilioaia, O., Fischer, G., Xu, Z. (2025). Natural diversity of telomere length distributions across 100 yeast strains. *bioRxiv*. https://doi.org/10.1101/2025.05.13.653712

Gilchrist, C., Stelkens, R. (2019). Aneuploidy in yeast: segregation error or adaptation mechanism? *Yeast*, *36*(9), 525–539. https://doi.org/10.1002/yea.3427

Greenall, A., Lei, G., Swan, D. C., James, K., Wang, L., Peters, H., Wipat, A., Wilkinson, D. J., Lydall, D. (2008). A genome wide analysis of the response to uncapped telomeres in budding yeast reveals a novel role for the NAD+ biosynthetic gene BNA2 in chromosome end protection. *Genome Biology*, *9*(10), R146. https://doi.org/10.1186/gb-2008-9-10-r146

Greider, C. W., Blackburn, E. H. (1987). The telomere terminal transferase of tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell*, *51*(6), 887–898. https://doi.org/10.1016/0092-8674(87)90576-9

Greider, C. W., Blackburn, E. H. (1989). A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. *Nature*, *337*(6205), 331–337. https://doi.org/10.1038/337331a0

Hackett, J. A., Feldser, D. M., Greider, C. W. (2001). Telomere dysfunction increases mutation rate and genomic instability. *Cell*, *106*(3), 275–286. https://doi.org/10.1016/S0092-8674(01)00457-3

Harari, Y., Ram, Y., Rappoport, N., Hadany, L., Kupiec, M. (2018). Spontaneous changes in ploidy are common in yeast. *Current Biology*, *28*(6), 825-835.e4. https://doi.org/10.1016/j.cub.2018.01.062

Hemann, M. T., Greider, C. W. (2000). Wild-derived inbred mouse strains have short telomeres. *Nucleic Acids Res, 28*(22):4474-8. https://doi.org/10.1093/nar/28.22.4474

Henning, K. A., Moskowitz, N., Ashlock, M. A., Liu, P. P. (1998). Humanizing the yeast telomerase template. *Proceedings of the National Academy of Sciences*, *95*(10), 5667–5671. https://doi.org/10.1073/pnas.95.10.5667

Hinchie, A. M., Sanford, S. L., Loughridge, K. E., Sutton, R. M., Parikh, A. H., Gil Silva, A. A., Sullivan, D. I., Chun-On, P., Morrell, M. R., McDyer, J. F., Opresko, P. L., Alder, J. K. (2024). A persistent variant telomere sequence in a human pedigree. *Nature Communications*, *15*(1), 4681. https://doi.org/10.1038/s41467-024-49072-9

Hu, Z., Killion, P. J., Iyer, V. R. (2007). Genetic reconstruction of a functional transcriptional regulatory network. *Nature Genetics*, *39*(5), 683–687. https://doi.org/10.1038/ng2012

Hull, R. M., Cruz, C., Jack, C. V., Houseley, J. (2017). Environmental change drives accelerated adaptation through stimulated copy number variation. *PLOS Biology*, *15*(6), e2001333. https://doi.org/10.1371/journal.pbio.2001333

Karimian, K., Groot, A., Huso, V., Kahidi, R., Tan, K. T., Sholes, S., Keener, R., McDyer, J. F., Alder, J. K., Li, H., Rechtsteiner, A., Greider, C. W. (2024). Human telomere length is chromosome end–specific and conserved across individuals. *Science*, *384*(6695), 533–539. https://doi.org/10.1126/science.ado0431

Koering, C. E., Fourel, G., Binet-Brasselet, E., Laroche, T., Klein, F., Gilson, E. (2000). Identification of high affinity Tbf1p-binding sites within the budding yeast genome. *Nucleic Acids Res*, *28*(13):2519-26. https://doi.org/10.1093/nar/28.13.2519

Kolberg, L., Raudvere, U., Kuzmin, I., Vilo, J., Peterson, H. (2020). gprofiler2 -- an R package for gene list functional enrichment analysis and namespace conversion toolset. F1000Res 15:9:ELIXIR-709. doi:10.12688/f1000research.24956.2

Kupiec, M. (2014). Biology of telomeres: Lessons from budding yeast. *FEMS Microbiology Reviews*, *38*(2), 144–171. https://doi.org/10.1111/1574-6976.12054

Kuznetsova, V., Grozeva, S., Gokhman, V. (2020). Telomere structure in insects: a review. *Journal of Zoological Systematics and Evolutionary Research*, *58*(1), 127–158. https://doi.org/10.1111/jzs.12332

Li, J., Stenberg, S., Yue, J. X., Mikhalev, E., Thompson, D., Warringer, J., Liti, G. (2023). Genome instability footprint under rapamycin and hydroxyurea treatments. *PLoS Genet*, *19*(11):e1011012. https://doi.org/10.1371/journal.pgen.1011012

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv,* https://doi.org/10.48550/arXiv.1303.3997

Li, H., Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Libuda, D. E., Winston, F. (2006). Amplification of histone genes by circular chromosome formation in Saccharomyces cerevisiae. *Nature*, *443*(7114), 1003–1007. https://doi.org/10.1038/nature05205

Lickwar, C. R., Mueller, F., Hanlon, S. E., McNally, J. G., Lieb, J. D. (2012). Genome-wide protein–DNA binding dynamics suggest a molecular clutch for transcription factor function. *Nature*, *484*(7393),

43

251–255. https://doi.org/10.1038/nature10985
Lin, J., Smith, D. L., Blackburn, E. H. (2004). Mutant telomere sequences lead to impaired chromosome separation and a unique checkpoint response. *Mol Biol Cell,* 15(4):1623-34. doi: 10.1091/mbc.e03-10-0740
Liti, G., Louis, E. J. (2003). NEJ1 prevents NHEJ-dependent telomere fusions in yeast without telomerase. *Mol Cell,* 11(5):1373-8. doi: 10.1016/s1097-2765(03)00177-1
Liti, G., Barton, D. B. H., Louis, E. J. (2006). Sequence diversity, reproductive isolation and species concepts in Saccharomyces. *Genetics*, *174*(2), 839–850. https://doi.org/10.1534/genetics.106.062166
Louis, E. J. (1995). The chromosome ends of *Saccharomyces cerevisiae*. *Yeast*, *11*(16), 1553–1573. https://doi.org/10.1002/yea.320111604
Louis, E. J. (1998). Whole chromosome analysis. Methods in microbiology, (26), 15-31. https://doi.org/10.1016/S0580-9517(08)70323-4
Lundblad, V., Blackburn, E. H. (1993). An alternative pathway for yeast telomere maintenance rescues est1-senescence. *Cell, 73*(2):347-60. https://doi.org/10.1016/0092-8674(93)90234-h
Lundblad, V., Szostak, J. W. (1989). A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell*, *57*(4), 633–643. https://doi.org/10.1016/0092-8674(89)90132-3
Lyčka, M., Bubeník, M., Závodník, M., Peska, V., Fajkus, P., Demko, M., Fajkus, J., Fojtová, M. (2024). TeloBase: a community-curated database of telomere sequences across the tree of life. *Nucleic Acids Research*, *52*(D1), D311–D321. https://doi.org/10.1093/nar/gkad672
MacIsaac, K. D., Wang, T., Gordon, D. B., Gifford, D. K., Stormo, G. D., Fraenkel, E. (2006). An improved map of conserved regulatory sites for Saccharomyces cerevisiae. *BMC Bioinformatics*, *7*(1), 113. https://doi.org/10.1186/1471-2105-7-113
Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., Zimin, A. (2018). MUMmer4: a fast and versatile genome alignment system. *PLOS Computational Biology*, *14*(1), e1005944. https://doi.org/10.1371/journal.pcbi.1005944
Marcand, S., Gilson, E., Shore, D. (1997). A protein-counting mechanism for telomere length regulation in yeast. *Science*, *275*(5302), 986–990. https://doi.org/10.1126/science.275.5302.986
McEachern, M. J., Blackburn, E. H. (1995). Runaway telomere elongation caused by telomerase RNA gene mutations. *Nature*, *376*(6539), 403–409. https://doi.org/10.1038/376403a0
McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., Flicek, P., Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biology*, *17*(1), 122. https://doi.org/10.1186/s13059-016-0974-4
McNeil, J. B., Dykshoorn, P., Huy, J. N., Small, S. (1990). The DNA-binding protein RAP1 is required for efficient transcriptional activation of the yeast PYK glycolytic gene. *Curr Genet,* 18(5):405-12. doi: 10.1007/BF00309909
Messner, C. B., Demichev, V., Muenzner, J., Aulakh, S. K., Barthel, N., Röhl, A., Herrera-Domínguez, L., Egger, A. S., Kamrad, S., Hou, J., Tan, G., Lemke, O., Calvani, E., Szyrwiel, L., Mülleder, M., Lilley, K. S., Boone, C., Kustatscher, G., Ralser, M. (2023). The proteomic landscape of genome-wide genetic perturbations. *Cell*, *186*(9), 2018-2034.e21. https://doi.org/10.1016/j.cell.2023.03.026
Mozzachiodi, S., Krogerus, K., Gibson, B., Nicolas, A., Liti, G. (2022). Unlocking the functional potential of polyploid yeasts. *Nature Communications*, *13*(1), 2580. https://doi.org/10.1038/s41467-022-30221-x
Muenzner, J., Trébulle, P., Agostini, F., Zauber, H., Messner, C. B., Steger, M., Kilian, C., Lau, K., Barthel, N., Lehmann, A., Textoris-Taube, K., Caudal, E., Egger, A. S., Amari, F., De Chiara, M., Demichev, V., Gossmann, T. I., Mülleder, M., Liti, G., Schacherer, J., Selbach, M., Berman, J., Ralser, M. (2024). Natural proteome diversity links aneuploidy tolerance to protein turnover. *Nature*, *630*(8015), 149–157. https://doi.org/10.1038/s41586-024-07442-9
O'Donnell, S., Yue, J. X., Saada, O. A., Agier, N., Caradec, C., Cokelaer, T., De Chiara, M., Delmas, S., Dutreux, F., Fournier, T., Friedrich, A., Kornobis, E., Li, J., Miao, Z., Tattini, L., Schacherer, J., Liti, G., Fischer, G. (2023). Telomere-to-telomere assemblies of 142 strains characterize the genome structural landscape in Saccharomyces cerevisiae. *Nature Genetics*, *55*(8), 1390–1399. https://doi.org/10.1038/s41588-023-01459-y
Okonechnikov, K., Golosova, O., Fursov, M., the UGENE team. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, *28*(8), 1166–1167. https://doi.org/10.1093/bioinformatics/bts091
Papp, B., Pál, C., Hurst, L. D. (2003). Dosage sensitivity and the evolution of gene families in yeast. *Nature*, *424*(6945):194-7. https://doi.org/10.1038/nature01771
Pavelka, N., Rancati, G., Zhu, J., Bradford, W. D., Saraf, A., Florens, L., Sanderson, B. W., Hattem, G. L., Li, R. (2010). Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature*, *468*(7321), 321–325. https://doi.org/10.1038/nature09529
Perez-Sepulveda, B. M., Heavens, D., Pulford, C. V., Predeus, A. V., Low, R., Webster, H., Dykes, G. F., Schudoma, C., Rowe, W., Lipscombe, J., Watkins, C., Kumwenda, B., Shearer, N., Costigan, K., Baker, K. S., Feasey, N. A., Hinton, J. C. D., Hall, N., The 10KSG consortium, … Wilson, C. (2021). An accessible, efficient and global approach for the large-scale sequencing of bacterial genomes. *Genome Biology*, *22*(1), 349. https://doi.org/10.1186/s13059-021-02536-3

44

Peska, V., Garcia, S. (2020). Origin, diversity, and evolution of telomere sequences in plants. *Frontiers in Plant Science*, *11*, 117. https://doi.org/10.3389/fpls.2020.00117

Peska, V., Fajkus, P., Bubeník, M., Brázda, V., Bohálová, N., Dvořáček, V., Fajkus, J., Garcia, S. (2021). Extraordinary diversity of telomeres, telomerase RNAs and their template regions in Saccharomycetaceae. *Sci Rep, 11*(1):12784. https://doi.org/10.1038/s41598-021-92126-x

Peter, J., De Chiara, M., Friedrich, A., Yue, J. X., Pflieger, D., Bergström, A., Sigwalt, A., Barre, B., Freel, K., Llored, A., Cruaud, C., Labadie, K., Aury, J. M., Istace, B., Lebrigand, K., Barbry, P., Engelen, S., Lemainque, A., Wincker, P., Liti, G., Schacherer, J. (2018). Genome evolution across 1,011 Saccharomyces cerevisiae isolates. *Nature*, *556*(7701), 339–344. https://doi.org/10.1038/s41586-018-0030-5

Platt, J. M., Ryvkin, P., Wanat, J. J., Donahue, G., Ricketts, M. D., Barrett, S. P., Waters, H. J, Song, S., Chavez, A., Abdallah, K. O., Master, S. R., Wang, L. S., Johnson, F. B. (2013). Rap1 relocalization contributes to the chromatin-mediated gene expression profile and pace of cell senescence. *Genes Dev*, *27*(12):1406-20. https://doi.org/10.1101/gad.218776.113

Pompei, S., Cosentino Lagomarsino, M. (2023). A fitness trade-off explains the early fate of yeast aneuploids with chromosome gains. *Proceedings of the National Academy of Sciences*, *120*(15), e2211687120. https://doi.org/10.1073/pnas.2211687120

Quinlan, A. R., Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Rancati, G., Pavelka, N., Fleharty, B., Noll, A., Trimble, R., Walton, K., Perera, A., Staehling-Hampton, K., Seidel, C. W., Li, R. (2008). Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. *Cell*, *135*(5), 879–893. https://doi.org/10.1016/j.cell.2008.09.039

Ribaud, V., Ribeyre, C., Damay, P., Shore, D. (2012). DNA-end capping by the budding yeast transcription factor and subtelomeric binding protein Tbf1: DNA-end capping by yeast Tbf1. *The EMBO Journal*, *31*(1), 138–149. https://doi.org/10.1038/emboj.2011.349

Rojas, J., Hose, J., Dutcher, H. A., Place, M., Wolters, J. F., Hittinger, C. T., Gasch, A. P. (2024). Comparative modeling reveals the molecular determinants of aneuploidy fitness cost in a wild yeast model. *Cell Genomics*, *4*(10), 100656. https://doi.org/10.1016/j.xgen.2024.100656

Sanchez, S. E., Gu, Y., Wang, Y., Golla, A., Martin, A., Shomali, W., Hockemeyer, D., Savage, S. A., Artandi, S. E. (2024). Digital telomere measurement by long-read sequencing distinguishes healthy aging from disease. *Nature Communications*, *15*(1), 5148. https://doi.org/10.1038/s41467-024-49007-4

Schmidt, T. T., Tyer, C., Rughani, P., Haggblom, C., Jones, J. R., Dai, X., Frazer, K. A., Gage, F. H., Juul, S., Hickey, S., Karlseder, J. (2024). High resolution long-read telomere sequencing reveals dynamic mechanisms in aging and cancer. *Nature Communications*, *15*(1), 5149. https://doi.org/10.1038/s41467-024-48917-7

Sepsiova, R., Necasova, I., Willcox, S., Prochazkova, K., Gorilak, P., Nosek, J., Hofr, C., Griffith, J. D., Tomaska, L. (2016). Evolution of telomeres in Schizosaccharomyces pombe and its possible relationship to the diversification of telomere binding proteins. *PLOS ONE*, *11*(4), e0154225. https://doi.org/10.1371/journal.pone.0154225

Sharp, N. P., Sandell, L., James, C. G., Otto, S. P. (2018). The genome-wide rate and spectrum of spontaneous mutations differ between haploid and diploid yeast. *Proceedings of the National Academy of Sciences*, *115*(22). https://doi.org/10.1073/pnas.1801040115

Sheltzer, J. M., Amon, A. (2011). The aneuploidy paradox: costs and benefits of an incorrect karyotype. *Trends in Genetics*, *27*(11), 446–453. https://doi.org/10.1016/j.tig.2011.07.003

Shen, W., Le, S., Li, Y., Hu, F. (2016). SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLOS ONE*, *11*(10), e0163962. https://doi.org/10.1371/journal.pone.0163962

Shen, X.-X., Opulente, D. A., Kominek, J., Zhou, X., Steenwyk, J. L., Buh, K. V., Haase, M. A. B., Wisecaver, J. H., Wang, M., Doering, D. T., Boudouris, J. T., Schneider, R. M., Langdon, Q. K., Ohkuma, M., Endoh, R., Takashima, M., Manabe, R., Čadež, N., Libkind, D., … Rokas, A. (2018). Tempo and Mode of Genome Evolution in the Budding Yeast Subphylum. *Cell*, *175*(6), 1533-1545.e20. https://doi.org/10.1016/j.cell.2018.10.023

Sholes, S. L., Karimian, K., Gershman, A., Kelly, T. J., Timp, W., & Greider, C. W. (2022). Chromosome-specific telomere lengths and the minimal functional telomere revealed by nanopore sequencing. *Genome Research*, *32*(4), 616–628. https://doi.org/10.1101/gr.275868.121

Sikorski, R. S., & Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* 122(1):19-27. doi: 10.1093/genetics/122.1.19

Song, S., Perez, J. V., Svitko, W., Ricketts, M. D., Dean, E., Schultz, D., Marmorstein, R., Johnson, F. B. (2020). Rap1-mediated nucleosome displacement can regulate gene expression in senescent cells without impacting the pace of senescence. *Aging Cell*, *19*(1), e13061. https://doi.org/10.1111/acel.13061

Stanway, C. A., Chambers, A., Kingsman, A. J., Kingsman, S. M. (1989). Characterization of the

45

transcriptional potency of sub-elements of the UAS of the yeast PGK gene in a PGK mini-promoter. *Nucleic Acids Res,* 17(22):9205-18. doi: 10.1093/nar/17.22.9205

Tattini, L., Tellini, N., Mozzachiodi, S., D'Angiolo, M., Loeillet, S., Nicolas, A., Liti, G. (2019). Accurate tracking of the mutational landscape of diploid hybrid genomes. *Molecular Biology and Evolution*, *36*(12), 2861–2877. https://doi.org/10.1093/molbev/msz177

Teixeira, M. T., Gilson, E. (2005). Telomere maintenance, function and evolution: The yeast paradigm. *Chromosome Research*, *13*(5), 535–548. https://doi.org/10.1007/s10577-005-0999-0

Tham, C. Y., Poon, L., Yan, T., Koh, J. Y. P., Ramlee, M. K., Teoh, V. S. I., Zhang, S., Cai, Y., Hong, Z., Lee, G. S., Liu, J., Song, H. W., Hwang, W. Y. K., Teh, B. T., Tan, P., Xu, L., Koh, A. S., Osato, M., Li, S. (2023). High-throughput telomere length measurement at nucleotide resolution using the PacBio high fidelity sequencing platform. *Nature Communications*, *14*(1), 281. https://doi.org/10.1038/s41467-023-35823-7

Thorvaldsdottir, H., Robinson, J. T., Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, *14*(2), 178–192. https://doi.org/10.1093/bib/bbs017

Vowinckel, J., Hartl, J., Marx, H., Kerick, M., Runggatscher, K., Keller, M. A., Mülleder, M., Day, J., Weber, M., Rinnerthaler, M., Yu, J. S. L., Aulakh, S. K., Lehmann, A., Mattanovich, D., Timmermann, B., Zhang, N., Dunn, C. D., MacRae, J. I., Breitenbach, M., Ralser, M. (2021). The metabolic growth limitations of petite cells lacking the mitochondrial genome. *Nat Metab*, *3*(11):1521-1535. https://doi.org/10.1038/s42255-021-00477-6

Wagih, O., Galardini, M., Busby, B. P., Memon, D., Typas, A., Beltrao, P. (2018). A resource of variant effect predictions of single nucleotide variants in model organisms. *Molecular Systems Biology*, *14*(12), e8430. https://doi.org/10.15252/msb.20188430

Wellinger, R. J., Zakian, V. A. (2012). Everything you ever wanted to know about Saccharomyces cerevisiae telomeres: beginning to end. *Genetics*, *191*(4), 1073–1105. https://doi.org/10.1534/genetics.111.137851

Whittemore, K., Vera, E., Martínez-Nevado, E., Sanpera, C., Blasco, M. A. (2019). Telomere shortening rate predicts species life span. *Proceedings of the National Academy of Sciences*, *116*(30), 15122–15127. https://doi.org/10.1073/pnas.1902452116

Yona, A. H., Frumkin, I., Pilpel, Y. (2015). A relay race on the evolutionary adaptation spectrum. *Cell*, *163*(3):549-59. https://doi.org/10.1016/j.cell.2015.10.005

Yona, A. H., Manor, Y. S., Herbst, R. H., Romano, G. H., Mitchell, A., Kupiec, M., Pilpel, Y., Dahan, O. (2012). Chromosomal duplication is a transient evolutionary solution to stress. *Proceedings of the National Academy of Sciences*, *109*(51), 21010–21015. https://doi.org/10.1073/pnas.1211150109

Yue, J.-X., Liti, G. (2018). Long-read sequencing data analysis for yeasts. *Nature Protocols*, *13*(6), 1213–1231. https://doi.org/10.1038/nprot.2018.025

Zackrisson, M., Hallin, J., Ottosson, L.-G., Dahl, P., Fernandez-Parada, E., Ländström, E., Fernandez-Ricaud, L., Kaferle, P., Skyman, A., Stenberg, S., Omholt, S., Petrovic, U., Warringer, J., Blomberg, A. (2016). Scan-o-matic: High-Resolution Microbial Phenomics at a Massive Scale. *G3 (Bethesda),* 8;6(9):3003-14. doi: 10.1534/g3.116.032342

Zappulla, D. C., Cech, T. R. (2004). Yeast telomerase RNA: a flexible scaffold for protein subunits. *Proceedings of the National Academy of Sciences*, *101*(27), 10024–10029. https://doi.org/10.1073/pnas.0403641101

Zetka, M. C., Müller, F. (1996). Telomeres in nematodes. *Seminars in Cell & Developmental Biology*, *7*(1), 59–64. https://doi.org/10.1006/scdb.1996.0009

Zonneveld, B. J. M., Van Der Zanden, A. L. (1995). The red *ade* mutants of *Kluyveromyces lactis* and their classification by complementation with cloned *ADE1* or *ADE2* genes from *Saccharomyces cerevisiae*. *Yeast*, *11*(9), 823–827. https://doi.org/10.1002/yea.320110904

46