Title: Telomerase-independent survival leads to a mosaic of complex subtelomere rearrangements in *Chlamydomonas reinhardtii*

Frédéric Chaux[1,3], Nicolas Agier[1,*], Clotilde Garrido[1,*], Gilles Fischer[1], Stephan Eberhard[2] and Zhou Xu[1]

Affiliations:

[1]Sorbonne Université, CNRS, UMR7238, Institut de Biologie Paris-Seine, Laboratory of Computational and Quantitative Biology, 75005 Paris, France

[2]Sorbonne Université, CNRS, UMR7141, Institut de Biologie Physico-Chimique, Laboratory of Chloroplast Biology and Light-Sensing in Microalgae, 75005 Paris, France

[3]present address: Laboratory for Marine and Atmospheric Biogeochemistry, Ruđer Bošković Institute, 10000 Zagreb, Croatia

*These authors made equal contributions to this work.

Correspondence: zhou.xu@sorbonne-universite.fr

Keywords: subtelomere, telomere, genome instability, structural variants, *Chlamydomonas reinhardtii*, post-senescence survivor

Abstract

Telomeres and subtelomeres, the genomic regions located at chromosome extremities, are essential for genome stability in eukaryotes. In the absence of the canonical maintenance mechanism provided by telomerase, telomere shortening induces genome instability. The landscape of the ensuing genome rearrangements is not accessible by short-read sequencing. Here, we leverage Oxford Nanopore Technologies long-read sequencing to survey the extensive repertoire of genome rearrangements in telomerase mutants of the model green microalga *Chlamydomonas reinhardtii*. In telomerase mutant strains grown for ~700 generations, most chromosome extremities were capped by short telomere sequences that were either recruited *de novo* from other loci or maintained in a telomerase-independent manner. Other extremities did not end with telomeres but only with repeated subtelomeric sequences. The subtelomeric elements, including rDNA, were massively rearranged and involved in breakage-fusion-bridge cycles, translocations, recombinations and chromosome circularization. These events were established progressively over time and displayed heterogeneity at the subpopulation level. New telomere-capped extremities composed of sequences originating from more internal genomic regions were associated with high DNA methylation, suggesting that *de novo* heterochromatin formation contributes to restore chromosome end stability in *C. reinhardtii*. The diversity of alternative strategies to maintain chromosome integrity and the variety of rearrangements found in telomerase mutants are remarkable and illustrate genome plasticity at short timescales.

1 Introduction

2

3 Protection of chromosome extremities is essential for genome integrity. For most eukaryotes, it is

4 achieved by repeated DNA sequences called telomeres and by telomere-bound factors, which

5 collectively prevent chromosome ends from being processed as DNA damage[1,2]. Telomeres shorten

6 with each round of replication due to the end replication problem and are in general maintained by

7 telomerase, a dedicated reverse transcriptase able to elongate telomeres *de novo*. In its absence, some

8 telomeres eventually reach a critical length that triggers replicative senescence, an arrested state

9 induced by the DNA damage checkpoint. Replicative senescence was shown in some species to

10 increase genome instability due to repair attempts and bypass of the checkpoint arrest through the

11 adaptation to DNA damage process[3-11]. In senescent cells that eventually escape cell cycle arrest, such

12 as some precursor cancer cells, telomeres become dysfunctional and induce further genomic

13 instabilities, a phenomenon termed telomere crisis[12,13].

14 The absence of telomerase therefore generates genome instabilities that stem from telomeres and

15 take many shapes: point mutations, deletions/insertions, translocations, aneuploidy, duplications and

16 even more dramatic rearrangements, such as chromothripsis during telomere crisis[7]. The precise

17 molecular mechanisms underlying these alterations are not all well understood but often involve

18 classical and alternative non-homologous end-joining (c- and a-NHEJ), homology-directed repair

19 (HDR), including homologous recombination and break-induced replication (BIR), together with

20 missegregation of chromosomes, breakage-fusion-bridge (BFB) cycles and other dynamic phenomena

21 that act in cascades over multiple cell divisions[3-5,7,13-17].

22 Subtelomeres are the genomic regions adjacent to telomeres and often contain families of paralogous

23 genes or pseudo-genes, ribosomal DNA (rDNA) arrays, transposable elements and other repeated

24 sequences[18-24]. Subtelomeres are often involved in telomere-associated rearrangements due to their

25 repetitive nature promoting HDR, replication fork stalling and template switching (FoSTeS), and BIR[19,25-]

1   [32]. Consistently, subtelomeres evolve rapidly even in closely related species and within species[18,27,33-]

2   [35]. In some species, in the absence of telomerase, telomeres can be stabilized using the alternative

3   lengthening of telomeres (ALT) pathway, which depends on homologous recombination and uses

4   repeated sequences found in telomeres and subtelomeres as substrates[36-39].

5   Genome alterations, especially structural variations (SVs), initiated by telomere shortening and

6   dysfunction, despite been widely studied in different models including cancer cells, have been difficult

7   to map exhaustively due to the complex nature of the rearrangements and the frequent involvement

8   of repeated sequences such as the ones found in subtelomeres[13,40]. Only recently were long-read

9   sequencing technologies used to enable the resolution of complex rearrangements at chromosome

10   extremities, in response to telomere shortening and dysfunction in *Caenorhabditis elegans* and

11   *Saccharomyces cerevisiae* [41-43].

12   We recently provided a comprehensive map of all 34 subtelomeres of the unicellular green alga

13   *Chlamydomonas reinhardtii* (17 chromosomes as haploid, 111 Mb)[24]. All contain arrays of repeated

14   elements, the most common being the Sultan element, present in 31 out of the 34 chromosome

15   extremities in a haploid strain, arranged in tandem repeats of up to 46 elements. The basic Sultan

16   element has a length of ~850 bp and forms class A subtelomeres. The Sultan element of class B

17   subtelomeres contains additional insertions. Next to most Sultan arrays (29 out of 31), a sequence of

18   ~500 bp called Spacer is unique to each subtelomere and may serve as promoter for downstream non-

19   coding RNA genes. The 3 remaining subtelomeres are entirely composed of rDNA, for a total of

20   approximately 350 copies corresponding to ~3 Mb. Two other repeated elements, called Suber and

21   Subtile, were found next to Sultan elements at 3 subtelomeres, called class C. The Suber element,

22   initially named pTANC[44], contains the most abundant interstitial telomere sequence (ITS) of the

23   genome. We previously found experimental evidence of telomere-associated genome rearrangements

24   potentially involving subtelomeres in telomerase mutants of *C. reinhardtii*, correlated with long-term

25   survival[45]. Indeed, although some telomerase-negative mutant subclones underwent senescence-

1   induced cell death, many managed to survive telomerase absence and must have therefore found a

2   solution to maintain and protect telomeres. In this work, using long-read Oxford Nanopore

3   Technologies ("Nanopore") sequencing able to traverse large repeated regions, we investigated

4   genome instability in telomerase mutants in *C. reinhardtii* and provide an exhaustive view of the

5   landscape of genome rearrangements. Our data show that subtelomeres are massively altered by

6   rearrangements within and between subtelomeres, involving amplification and deletion of repeated

7   elements (or even entire subtelomeres), fusions, BFB cycles, and other events of high complexity.

8   Rearrangements occur progressively over time and create a mosaic population of cells which undergo

9   competition and selection. Telomeres are maintained at a majority of extremities in a telomerase-

10 independent manner and possibly promote methylation of adjacent DNA to help stabilize chromosome

11 extremities.

12

13 **Results**

14

15 **Long-read Nanopore sequencing of prolonged cultures of telomerase mutants**

16 To investigate the genome rearrangements induced by the long-term absence of telomerase, we

17 cultivated two different telomerase mutant strains, *tel-m1* and *tel-m2* (Fig. 1A). The two mutant strains

18 were obtained from a library of random insertion mutants[46] and the insertion of the paromomycin

19 resistance gene in the RNA-binding domain (*tel-m1*) and catalytic domain (*tel-m2*) of *CrTERT*, the gene

20 encoding the catalytic subunit of telomerase, was confirmed previously[45]. We estimated that the

21 mutants underwent ~700 generations between creation of the library of mutants in the Jonikas lab

22 and collection of the samples called *tel-m1-1* and *tel-m2-1*, based on the procedure of transformation,

23 propagation and freezing described in[46] and our subsequent passages. Additionally, for *tel-m1*, we

24 collected an earlier sample cultivated for ~250 generations named *tel-m1-0*, which corresponded to

25 the earliest timepoint we could reasonably obtain upon reception of the strain in our lab. The wild-

5

1   type strain used to make the library of mutants, CC-4355, was also cultivated upon reception alongside

2   the mutants, for at least 450 generations.

3   We optimized a CTAB/phenol/chloroform purification protocol followed by a size selection step to

4   extract and obtain long DNA molecules[47]. We then sequenced the genomic DNA of these 4 samples by

5   long-read Nanopore sequencing and obtained 4 sets of high-quality reads of N50 > 14 kb and depth >

6   25X (Supp. Table ST1). We previously proved that long reads allowed the accurate reconstruction of

7   the highly repeated subtelomere structures in *C. reinhardtii*[24], which is instrumental for the analysis of

8   rearrangements potentially implicating telomeres, subtelomeres and other repeated elements.

9
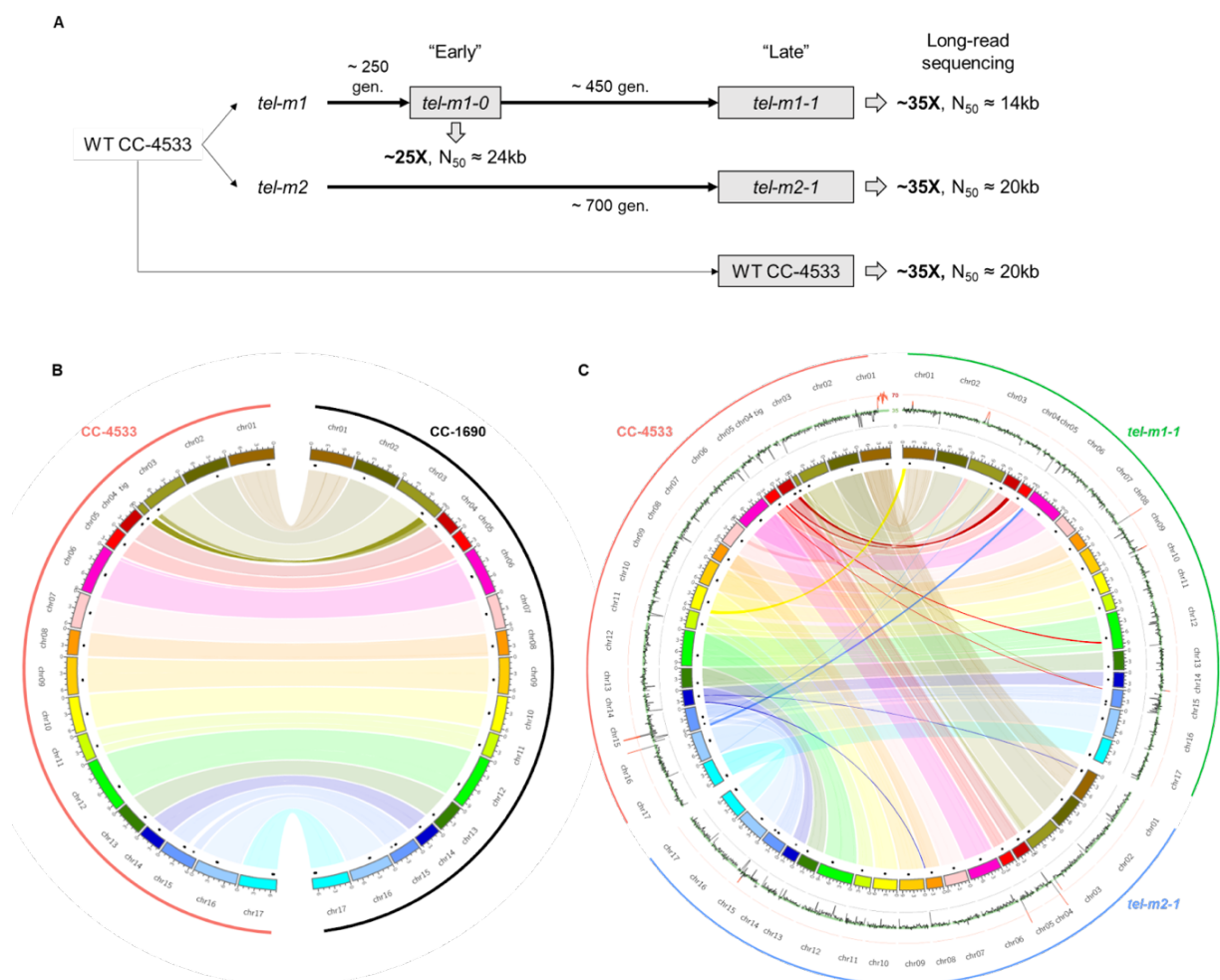
10

11   **Figure 1. Independent genome assemblies of CC-4533 and telomerase mutants *tel-m1-1* and *tel-m2-1* reveal**

12   **few re-arrangements at the genome assembly level.** (A) Experimental setup and main sequencing output. (B-C)

1   Circos plot[48] of (B) chromosome scaffolds from the long-term culture of CC-4533 as compared to CC-1690 and to

2   (C) *tel-m1-1* and *tel-m2-1*, shown as colored boxes, with black marks for centromere clusters (composed of the

3   *Zepp-L1* repeat). Links connect homologous blocks, colored according to CC-4533 chromosomes. Outer plot

4   displays sequencing depth, averaged over 100 kb regions, with 0X, 35X and 70X grid lines shown in black, green

5   and red, respectively. Depth over 50X is highlighted in red and genomic position is indicated in megabases.

6

7   The long-read datasets were processed by 3 assemblers, allowing us to generate *de novo* genome

8   assemblies of CC-4533 and each telomerase mutant (Fig. 1B and C). We reached large chromosome-

9   scale scaffolds similar to the most contiguous reference genome of CC-1690[49], to which we compared

10   our new assembly of CC-4533. Sequencing depth (Fig. 1C, outer circles) indicated a partial duplication

11   of chromosome 1 in CC-4533 not present in previous sequencing[50], suggesting that a duplicated mini-

12   chromosome 1 emerged during the long-term culture in our laboratory, a phenomenon that is quite

13   common since similar duplications were observed in other laboratory strains[51]. The copy number

14   variation (CNV) on chromosome 15 and 16 were due to frequently misassembled regions known to

15   contain repeated sequences[52]. No other chromosome-scale difference was found between CC-4533

16   and CC-1690 (Fig. 1B). When compared to CC-4533, *tel-m1-1* displayed a number of large-scale SVs

17   and CNVs, corresponding to translocations and duplications of chromosome extremities (Fig. 1C). The

18   other mutant, *tel-m2-1*, showed fewer large-scale rearrangements in the assembly, but CNVs were

19   detected at several locations, suggesting genome alterations. We then used the SV caller MUM&Co[53]

20   to detect insertions, deletions, duplications, inversions and translocations compared to CC-1690 and

21   we found similar numbers for CC-4533 and the telomerase mutants assemblies, with about half of SVs

22   shared by all 3 strains (n = 370 out of a total of 746) (Supp. Fig. S1A and B). The assemblies of the

23   telomerase mutants displayed a moderate number of unique SVs (n = 52 and 49 for *tel-m1-1* and *tel-*

24   *m2-1*, respectively), involving sequences of median size < 600 bp.

25   However, we found that the quality of read mapping to subtelomeres was lower than those

26   corresponding to the core genome both in CC-4533 and in the telomerase mutants (Supp. Fig. S1C),

1   presumably because the repeated elements found at subtelomeres present a challenge for genome

2   assembly even using long-read sequencing data[54]. Consistently, the assembled genomes reached only

3   25, 2 and 19 telomeres in CC-4533, *tel-m1-1* and *tel-m2-1*, respectively. We thus suspected that the

4   assembly-level analysis would miss genome rearrangements affecting chromosome extremities.

5   Additionally, assemblers were not designed to handle genome heterogeneity at the subpopulation

6   level, which is likely to be the case in telomerase mutants. We therefore directly looked for genome

7   alterations at the level of the sequencing reads instead of the assemblies in the rest of this work.

8   **Telomere shortening and loss in telomerase mutants**

9   We were first interested in investigating changes related to telomeres. To detect all telomere

10   sequences at the level of individual reads, we developed a new method called TeloReader that scores

11   8-mers with respect to their level of identity to any of the canonical 8-mers (TTTTAGGG/CCCTAAAA

12   and circular permutations) in sliding windows and uses thresholds for the average score to find the

13   boundaries (see Methods section). TeloReader allowed us to detect all telomere sequences of at least

14   16 bp, whether they were at a chromosome extremity (terminal) or not (interstitial telomere

15   sequence, ITS), in the read datasets.

16   We measured the length of all terminal telomere sequences in the wild-type and the mutants reads

17   and found shorter telomeres in the mutants (CC-4533: 293 ± 126 bp, *tel-m1-0*: 167 ± 109 bp, *tel-m1-*

18   *1*: 162 ± 120 bp, *tel-m2-1*: 181 ± 88 bp; Fig. 2A). Telomeres were already short in the *tel-m1-0* sample

19   and their length did not further decrease in the *tel-m1-1* sample, suggesting the stabilization of

20   telomere length by telomerase-independent mechanisms or *de novo* telomere formation after

21   complete loss. Because we suspected that the aggregate length distribution measurement would mask

22   telomere length dynamics at each extremity, we assigned the reads containing a terminal telomere to

23   a specific chromosome arm based on the presence of specific and unique Sultan elements in the reads

24   and measured telomere length distribution for each Sultan-associated extremity (Supp. Fig. S2). Out

25   of the 27 studied class A and B extremities, 8 showed telomeres further shortening between *tel-m1-0*

8

1 and *tel-m1-1*, with large variations in the rate of shortening. Interestingly, they were stabilized or even

2 slightly increased in length between *tel-m1-0* and *tel-m1-1* in 16 cases, indicating a telomerase-

3 independent mechanism of telomere maintenance or *de novo* formation. In the remaining 3

4 extremities, the very low or even absence of reads in *tel-m1-0* and/or *tel-m1-1* prevented comparison

5 and suggested the loss of the entire telomere or subtelomere sequence. Telomere length distribution

6 at each extremity also displayed variations in *tel-m2-1*.

7 To test if a subset of telomeres were completely lost, we computed the total number of telomere-

8 containing reads normalized to the sequencing depth of the nuclear genome and found that the

9 average number of telomeres per cell was smaller in the mutants: 22 in *tel-m1-0*, 18 in *tel-m1-1* and

10 25 in *tel-m2-1* vs 36 in CC-4533 (which has an extra mini-chromosome 1 in addition to the 17

11 chromosomes) (Fig. 2B).

12 Overall, telomeres shortened as expected in telomerase mutants but our data revealed the

13 heterogeneity of shortening pattern at each specific telomere and suggested that telomeres were

14 maintained or newly formed in *tel-m1-1* as compared to *tel-m1-0* in a telomerase-independent

15 manner. In the next sections, we investigated changes and chromosomal rearrangements occurring at

16 subtelomeric regions.

17
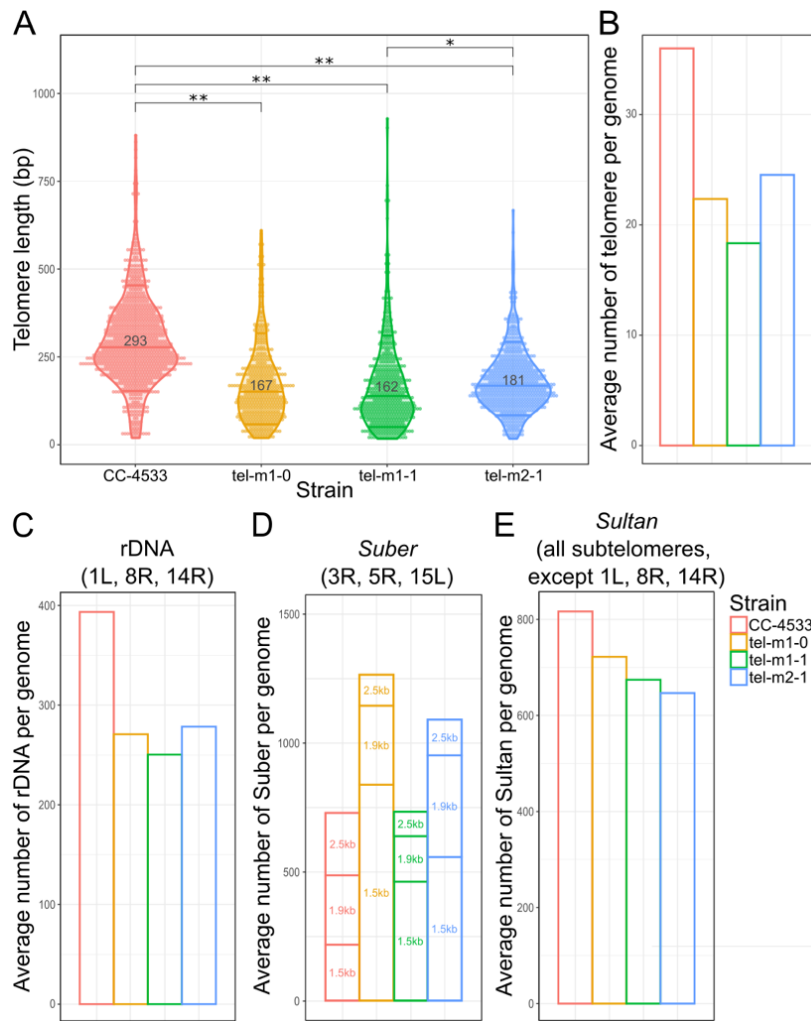
Figure 2. Telomere length distribution and CNVs of subtelomeric elements. (A) Telomere length distribution of terminal telomere sequences detected by TeloReader from the reads dataset in CC-4533 and 3 telomerase mutant samples. * and ** indicate statistically significant difference with p-value = 6.8×10^{-9} and p-value $< 2.2 \times 10^{-16}$, respectively, using a non-parametric Kruskal-Wallis test. (B) Number of telomeres in reads normalized by the sequencing depth. (C-E) CNVs for repeated elements normally found at subtelomeres including (C) the rDNA, (D) the Suber elements subdivided into 3 types according to their length ("1.5kb", "1.9kb" and "2.5kb"), and (E) the Sultan elements.

Copy number variations of subtelomeric elements in telomerase mutants

As a first insight into the stability of repeated subtelomeric elements (rDNA, Suber and Sultan), we computed their CNVs in telomerase mutants compared to CC-4533.

10

1   We used the rDNA sequence as a query to find all reads containing rDNA. We observed an overall

2   decrease of 31%, 36% and 29% in rDNA copy number in *tel-m1-0*, *tel-m1-1* and *tel-m2-1*, respectively,

3   compared to CC-4533 (Fig. 2C). We also found that, while in the wild-type, it was exclusively found at

4   3 subtelomeres as expected[24], rDNA could be mapped at several additional regions of the genome in

5   telomerase mutants, implicating rDNA in genome rearrangements (see below). The two major rDNA

6   clusters at subtelomeres 8R and 14R likely lost a large fraction of their rDNA copies.

7   The same analysis was performed with the Suber element, which is found as 3 main subtypes of sizes

8   ~1.5 kb, ~1.9 kb and ~2.5 kb[24,44], and showed changes specific to each subtype (Fig. 2D). The copy

9   number of the 1.5 kb element was greatly increased in the telomerase mutants compared to wild-type,

10 especially in *tel-m1-0*, in contrast to the 2.5 kb element which decreased in copy number in all mutants.

11 The number of 1.9-kb elements only decreased in *tel-m1-1*. Given the organization of Subers in arrays

12 containing only one subtype[24], the different changes in copy number depending on the subtype likely

13 reflected duplicated, deleted and potentially rearranged arrays.

14 Finally, the Sultan element, the most abundant and widespread repeated element in subtelomeres,

15 was decreased in copy number in both telomerase mutants (Fig. 2E). In the next sections, we

16 investigate in more detail what this decrease entails at the level of each chromosome extremity.

17 Overall, all the main subtelomeric repeated elements showed CNVs in the telomerase mutants,

18 indicating that subtelomeres are involved in telomere-induced chromosome rearrangements.

19

20 Contraction, erosion and expansion of Sultan arrays

21 To analyze subtelomere structure, we looked for the presence of the Spacer element in all reads

22 because this element is present as a single copy in 29 out of 34 subtelomeres in the wild-type strain,

23 with sufficient sequence differences to uniquely identify the corresponding subtelomere[24]. Reads

24 containing a Spacer element were thus grouped according to their Spacer (Supp. Dataset 1).

1   While for all Spacers, reads from CC-4533 were consistent with a subtelomeric structure identical or

2   near-identical to the one we inferred from the genome assembly of CC-1690[24] (Fig. 3, Supp. Dataset

3   1), reads from the telomerase mutants revealed many differences compared to CC-4533, including a

4   striking variety of rearrangements (Table 1 and Supp. Dataset 1).

5

	tel-m1-0	tel-m1-1	tel-m2-1
Similar to WT	9	7	9
Absence of Spacer	1	1	3
Contraction of stabilized Sultan array	3	1	5
Unstable Sultan array	3	3	4
Expansion of stabilized Sultan array	1	0	1
Fusion between Sultan arrays (total):	13	17	3
- Same orientation	3	5	1
- Inverted orientation	10	12	2
Recruitment of rDNA	4	5	1
Recruitment of other sequences	3	4	2
BFB	4	5	0
Chromosome circularization	1	1	0
Heterogeneous subpopulations	7	9	4

6

7   **Table 1. Summary of rearrangements classified by types, found at Spacer-containing subtelomeres.** The

8   number of the described rearrangements/events found at Spacer-containing subtelomeres, compared to CC-

9   4533, is indicated. For subtelomeres with reads that represented multiple subpopulations ("Heterogeneous

10   subpopulations"), the rearrangements/events of each subpopulation were counted. Only rearrangements

11   supported by at least two reads were counted.

12

13   We first focused on subtelomeres composed, in CC-4533, of an array of Sultan elements as the only

14   repeated element, which are the vast majority[24] (class A and B, 27 subtelomeres out of 34). In CC-4533,

15   at each class A/B Sultan subtelomere, the reads were consistent with a single population of

16   subtelomeres with a fixed number of Sultans and capped by telomeres, as supported by a majority of

17   reads (Fig. 3A and D). Reads that did not reach the extremity of the chromosome were due to DNA

18   molecules physically broken during the extraction and preparation procedure and would be expected

19   to terminate at random position in the Sultan array and within a Sultan element, which we confirmed

1 (Fig. 3B and E). Additionally, in CC-4533, all reads containing a specific Sultan element and telomere

2 sequence could be retrieved and their number was often close to the number of Spacer-containing

3 reads (Supp. Fig. S3A), consistent the fact that Spacer-containing reads did not always reach a telomere

4 sequence because of the physical breakage of the extracted DNA molecule. Interestingly, both

5 telomerase mutants frequently showed reads that also supported a fixed number of Sultans (Fig. 3A

6 and C). We call such subtelomeres "stabilized Sultan arrays". However, the number of Sultan elements

7 could differ from CC-4533 (Fig. 3C): in *tel-m1-1* and *tel-m2-1* combined, we observed a decreased

8 number of Sultans in 6 cases by 2-32 repeats, and in 1 case on the contrary, the number of Sultan

9 element was slightly increased by 1. For the rest of the cases (n = 11), no variation of the number of

10 Sultans was observed. We found that stabilized Sultan arrays in telomerase mutants were capped by

11 short telomere sequences (Fig. 3A and Supp. Fig. S3A), suggesting that these telomeres were

12 maintained in a telomerase-independent manner.

13 We then asked whether at the extremities of stabilized Sultan arrays, the telomeres could be

14 established *de novo* after complete loss or whether Sultan elements were excised or amplified

15 internally independently of the initially present telomere. To address this point, we identified the

16 junction between the telomere sequence and the last Sultan element and compared it between CC-

17 4533 and the mutants (Fig. 3B and Supp. Fig. S3B). In 4 out of 18 stabilized Sultan arrays in the two

18 mutants combined, we found that telomeres transitioned into the Sultan element at a different

19 position within the Sultan compared to CC-4533, indicating that the telomeres were added *de novo* on

20 truncated Sultans (Supp. Fig. S3A and B). In the rest (14 out of 18), telomeres were found at the same

21 position of the last Sultan in CC-4533 and the mutants, suggesting that changes in the number of Sultan

22 elements of the array might also be due to excision of internal Sultan elements, without affecting the

23 last Sultan.

24 At some subtelomeres (3 for *tel-m1-1* and 6 for *tel-m2-1*), the reads were not consistent with a

25 stabilized number of Sultan elements, and the vast majority had different counts, suggesting a

1   heterogeneous population of cells in which that specific extremity shortened progressively. These

2   Sultan arrays were qualified as "unstable". This situation is exemplified by subtelomere 2R in *tel-m2-1*

3   (Fig. 3D). Consistent with a progressive erosion of the extremity, the last Sultan element ended at

4   different positions of the Sultan element in different reads and was often not capped by telomeres

5   (Fig. 3E and Supp. Fig. S3A and B). To rule out that these reads containing a variable number of Sultan

6   elements were not simple due to physically broken DNA molecules, we searched for all reads

7   containing the same Sultan elements and telomeres and compared their number with the number of

8   reads containing the corresponding Spacer (Supp. Fig. S3A). In contrast to CC-4533, unstable Sultan

9   arrays in the mutants showed fewer telomere-containing reads than Spacer-containing ones,

10   indicating telomere loss and supporting the idea that these extremities truly ended with Sultan

11   elements. The progressive loss of telomere repeats and Sultan element was consistent with the overall

12   decrease in the number of telomeres and Sultan copy number (Fig. 2B and E). Erosion of Sultan arrays

13   would eventually lead to their complete loss, a situation that we could clearly evidence for

14   subtelomere 7L (Supp. Fig. S3C): in *tel-m1-0*, the subtelomere contained only 3 Sultans, compared to

15   8 in CC-4533, but in *tel-m1-1* the extremity was eroded beyond the Spacer and a new telomere was

16   formed with the additional translocation, in a subpopulation of reads, of a sequence from chromosome

17   11 which had a microhomology to the telomere sequence (Supp. Fig. S3D and E). Interestingly, an

18   analogous comparison between CC-4533, *tel-m1-0* and *tel-m1-1* for the class C subtelomere 15L

19   tended to suggest that Suber arrays could also be progressively eroded until completely lost (Supp. Fig.

20   S3F).

21
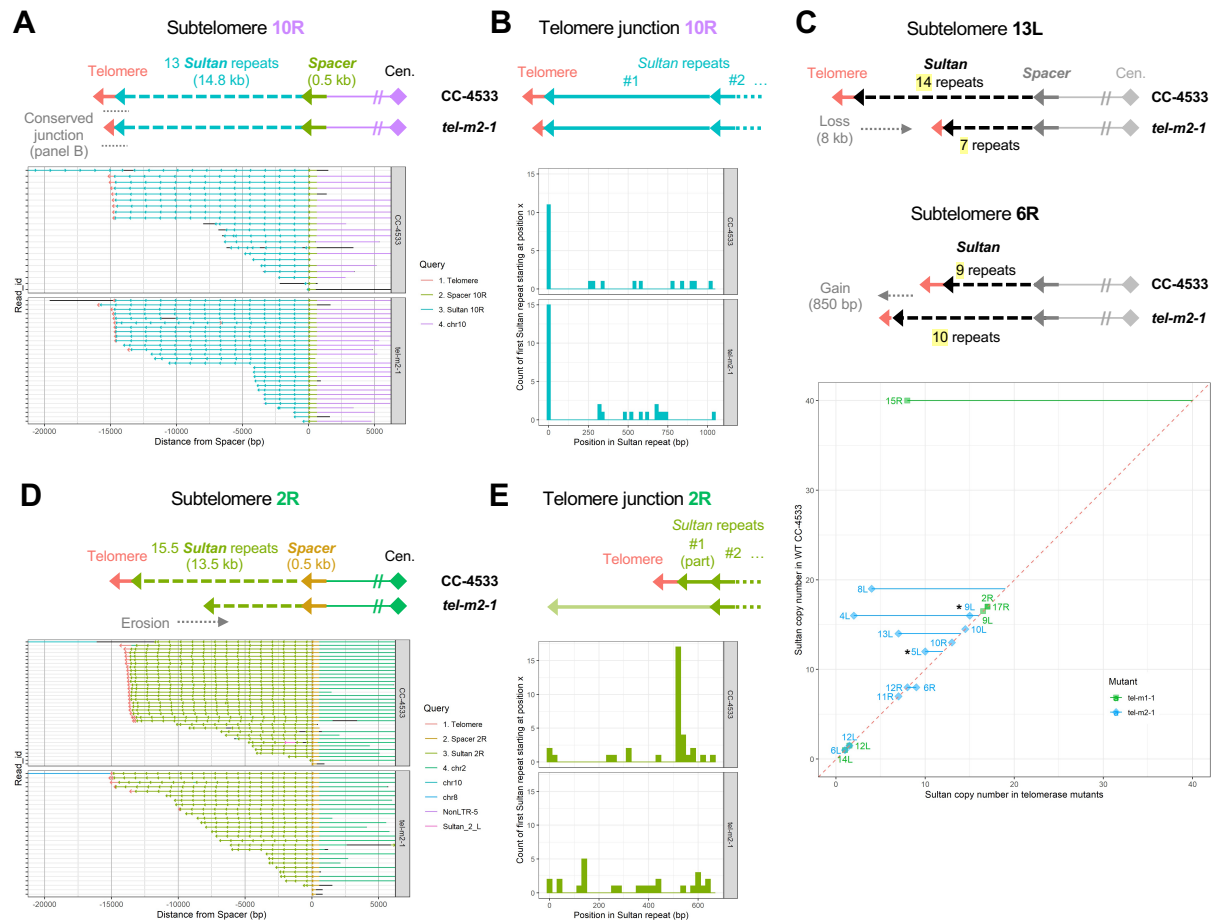
Figure 3. Contraction, erosion and expansion of Sultan arrays. (A) The stabilized 10R Sultan array of *tel-m2-1* capped by a telomere sequence is depicted in comparison to CC-4533. Raw Nanopore reads, colored based on the element (as indicated in the legend; black segments correspond to sequences that did not align to any of the selected queries) and anchored at the Spacer, are shown. (B) For each read, the starting position of the Sultan element adjoining the telomere sequence is recorded and the overall distribution of counts over the length of a Sultan element is represented. (C) Schematic representation of contraction and expansion of stabilized Sultan arrays (upper part) and overall changes in Sultan number in stabilized Sultan arrays for *tel-m1-1* and *tel-m2-1* (lower part). (D) Same as (A) with the unstable 2R Sultan array of *tel-m2-1* mostly uncapped by telomeres. (E) Same as (B) but for the unstable 2R Sultan array of *tel-m2-1*.

Complex rearrangements within and between subtelomeres

In telomerase mutants, we observed a striking fraction of telomere-containing reads with unusual combinations of subtelomeric elements (Table 1 and Supp. Dataset 1). To better characterize the

1   structure of these rearranged subtelomeres, we again used the Spacer-based analysis to help anchor

2   the reads to a given chromosome extremity.

3

4   *Fusion between subtelomeric elements*

5   Fusions of elements from different subtelomeres were the most frequent genome rearrangement

6   occurring at subtelomeres (Table 1). The simplest type corresponded to the juxtaposition of two or

7   more arrays of different Sultan elements, either in the same or in reverse orientations, such as

8   subtelomere 6R in *tel-m1-0* (Fig. 4A) or subtelomere 15R in *tel-m2-1* (Supp. Fig. S4A). Multiple

9   subtelomeric elements of different types, such as rDNA and Sultan in subtelomere 6L of *tel-m1-0* and

10   *tel-m1-1*, were also observed (Fig. 4B and Table 1). Because rDNA and Sultan do not share sequence

11   homology, we speculate their fusion stemmed from NHEJ-dependent translocation events or end-to-

12   end fusion.

13   The complexity of the observed genome rearrangements suggested that they formed through a

14   multistep process. To test this hypothesis, we compared the rearrangements found in *tel-m1-0* to

15   those in *tel-m1-1*, as these two samples were collected with a ~450-generation delay. We could find

16   intermediate rearrangement states in the earlier *tel-m1-0* sample, suggesting that at least some

17   complex rearrangements were formed in multiple steps over time (Fig. 4B and C; Supp. Fig. S3C and

18   S4B). For example, at subtelomere 6L, the already complex structure found in *tel-m1-0* further changed

19   in *tel-m1-1* and diverged into two structurally distinct subpopulations (Fig. 4B). As another example,

20   subtelomere 17L displayed a reduced number of Sultan elements in *tel-m1-0* compared to wild-type

21   with a subpopulation of reads capped by telomeres, representing a short stabilized Sultan array, but

22   contained additional sequences from a $(CA)_n$ microsatellite, from chromosome 15 and from 5R Sultan

23   elements, fused to the remaining 17L Sultans in *tel-m1-1* (Fig. 4C).

24

1   *Signature of BFB events*

2   The read coverage mapped to the assembly of *tel-m1-1* genome identified a duplicated region at the

3   extremity of chromosome 9R (Fig. 4D, upper part). Closer inspection of the reads anchored by their

4   Spacer element revealed a pattern of 2 inverted Sultan arrays fused to each other in a head-to-head

5   manner (Fig. 4D, middle part). All Sultan elements could be attributed to subtelomere 9R

6   unambiguously. Each array was adjacent to its own 9R Spacer element as well as downstream 9R-

7   specific sequences. Overall, this configuration was indicative of a fusion between sister chromatids

8   followed by a breakage in anaphase, an event that initiated a canonical breakage-fusion-bridge cycle

9   until the extremities were stabilized. We then looked for the new 9R extremity in *tel-m1-1* and

10   identified the reads that mapped close to the 1X/2X coverage boundary. Strikingly, reads that mapped

11   only to the 2X side continued with at least two rDNA units over >15 kb (Fig. 4D, lower part). The rDNA

12   sequence disrupted a RTEX-1 long interspersed nuclear element (LINE) retrotransposon framed by two

13   arrays of MSAT-4B satellites, as shown in Fig. 4D. What lay beyond the rDNA array remained unknown,

14   in particular whether the extremity was capped by additional telomere sequences. In another example,

15   the structure of the 11R extremity of *tel-m1-0* was consistent with at least two cycles of BFB, based on

16   the multiple inverted repeats (Supp. Fig. S4C). We found a total of 5 events consistent with BFB in *tel-

17   m1-1* (Table 1), indicating that BFB cycles were a common mechanism at play once telomeres became

18   too short or were lost.

19

20   *Circularization of chromosome 4*

21   We report the head-to-head fusion of subtelomeres 4L and 4R in *tel-m1-1*, after complete loss of the

22   telomere sequences and partial loss of Sultan elements (from 15 to 2 for 4L and from 21 to 7 for 4R)

23   (Fig. 4E). The *de novo* assembly of chromosome 4 in *tel-m1-1* without relying on scaffolding with a

24   reference sequence indicated that chromosome 4 was made of a single contig with no extremity and

17

1   had therefore a circular structure (Supp. Fig. S4D-E). Circularization of the chromosome bypasses the

2   need for functional telomeres to stabilize the extremities of linear chromosomes.

3
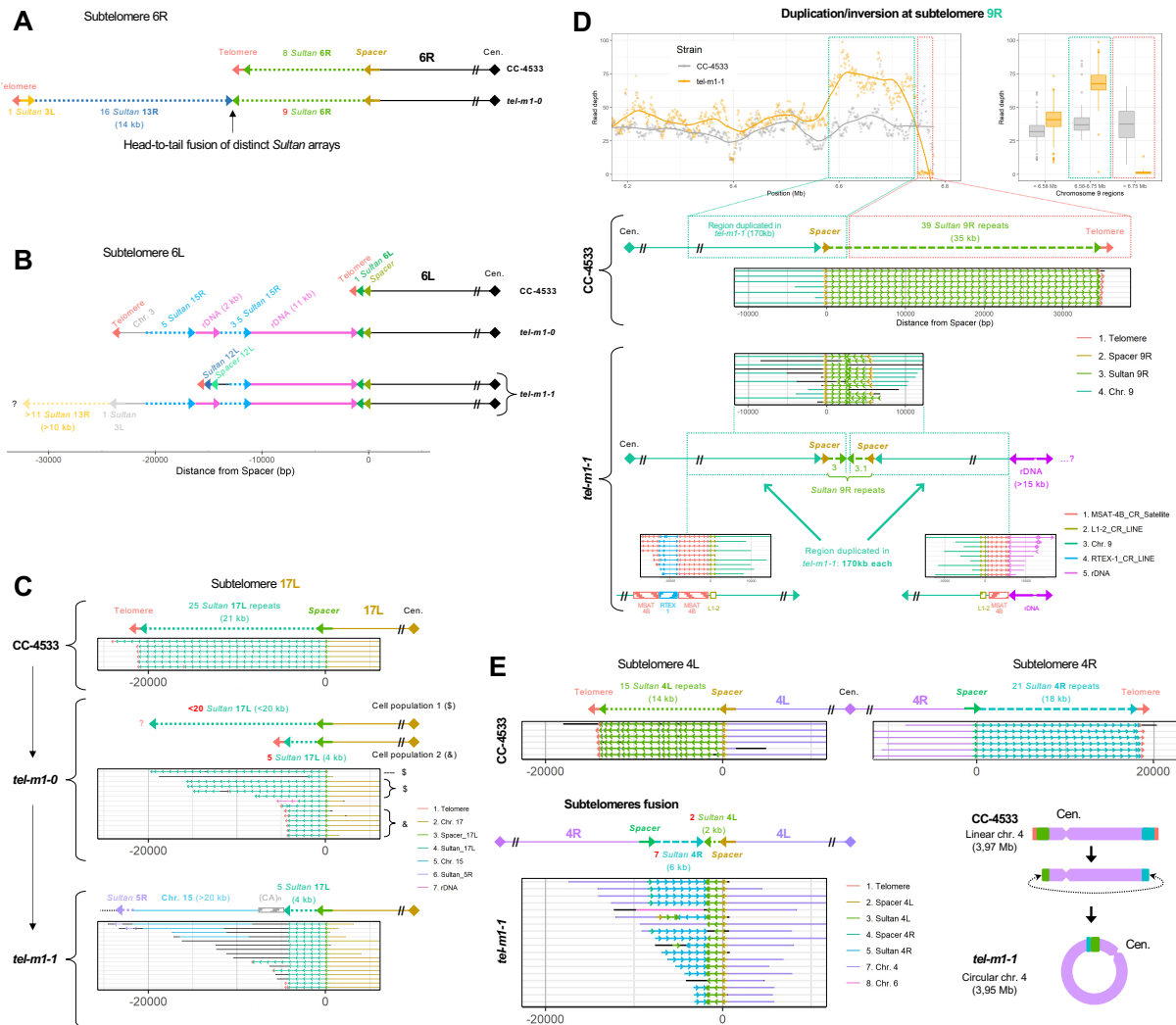
4

5   **Figure 4. Complex genome rearrangements at subtelomeres.** (A) Schematic representation of subtelomere 6R

6   where, in *tel-m1-0*, 2 additional Sultan arrays were fused to the initial Sultan array in inverted orientation. (B)

7   Schematic representation of complex rearrangements at subtelomere 6L in *tel-m1-0* and *tel-m1-1* involving

8   Sultan elements from different subtelomeres and rDNA sequences. In *tel-m1-1*, 2 subpopulations of reads reveal

9   2 distinct structures stemming from the initial rearrangement found in *tel-m1-0*. (C) Schematic representation of

10   the dynamics of rearrangements at subtelomere 17L between CC-4533, *tel-m1-0* and *tel-m1-1*. The structure of

11   the subtelomere in *tel-m1-1* appeared to have evolved from the structure in cell population 2 of *tel-m1-0*. (D)

12   Signature of a BFB event at subtelomere 9R in *tel-m1-1*. (*Upper part*) Duplication of the last 170 kb of the

18

1   chromosome end. (*Middle part*) Individual reads supporting the loss of telomeres and 36 Sultan elements, and

2   the end-to-end fusion of the 9R sister chromatids. (*Lower part*) Reads showing the recruitment of > 15 kb of rDNA

3   sequences at the new 9R extremity, disrupting an array of MSTA-4B satellite sequences. Black segments

4   correspond to sequences that are not aligned to any of the selected queries. (E) Representation of reads

5   supporting the fusion of subtelomeres 4L and 4R in *tel-m1-1*, after complete loss of telomeres and a total of 27

6   Sultan elements. (Lower right) Scheme of the inferred circularization of chromosome 4, also supported by *de*

7   *novo* genome assembly (Supp. Fig. S4D-E). Black segments correspond to sequences that are not aligned to any

8   of the selected queries.

9

10   **New telomeres outside canonical subtelomere regions lead to drastic karyotype**

11   **alterations**

12   We next investigated terminal-telomere-containing reads not associated with the canonical

13   subtelomeric elements, with the aim of revealing the formation of new extremities outside their

14   normal subtelomeric contexts. Using TeloReader, we gathered all reads containing telomere

15   sequences that were not associated to known subtelomeric elements (*Sultan*, *Spacer*, *Suber*, *Subtile* or

16   rDNA). We mapped these reads on the genome of CC-4533 and found such subtelomere-less telomeres

17   at 3 and 2 locations in *tel-m1-1* and *tel-m2-1*, respectively. Consistently with the genome assembly, we

18   also detected the new telomere extremity of the duplicated mini-chromosome 1 from CC-4533, which

19   transitioned directly into the chromosome 1 sequence through a microhomology and without

20   subtelomeric element nor other exogenous sequence (Supp. Fig. S5A).

21   At subtelomeres 14L in *tel-m2-1* and 7L in *tel-m1-1* (Fig. 5A, Supp. Fig. S3C-E), telomeres connected to

22   the chromosome arm several kilobases away from the expected telomere location in CC-4533, and the

23   corresponding arrays of *Sultan* repeats were lost, but without loss of annotated genes. At these novel

24   telomere junctions, the chromosome arm showed no sequence homology with telomeres, suggesting

25   that the telomere was recruited at these sites by non-homologous end joining rather than homology-

19

1   mediated mechanisms. At subtelomere 7L in *tel-m1-1*, a minor population of reads showed an

2   additional fragment from chromosome 11 next to the new telomere, with a 4-bp microhomology at

3   the junction (Supp. Fig. S3D). The resulting interstitial telomeric sequence was shorter than the

4   telomeres observed in the main population of reads (Supp. Fig. S3E).

5   The other new telomeres were involved in more complex rearrangements. The subtelomere 3L of *tel-*

6   *m1-0* had 13 Sultan elements remaining after loss of the telomere sequence and 4 Sultan elements,

7   followed by a short array of 13R Sultan elements and telomere sequence (Fig. 5B). In a subpopulation

8   of *tel-m1-0* reads and in *tel-m1-1*, this structure was further altered by the loss of the telomere and

9   the fusion of 500 kb of duplicated sequence originally located at 2.1 Mb from the end of the left arm

10   of chromosome 7, and capped by a new telomere (Fig. 5B). Junctions to telomeres (junction A) and to

11   the 13R Sultan array (junction B) involved micro-homologies of 2-4 bp (Supp. Fig. S5B and C), while the

12   Sultans from 13R and 3L were connected with the insertion of a short sequence of unknown origin

13   (junction C) (Supp. Fig. S5D), but without homology nor truncation of the neighboring Sultan elements.

14   On chromosome 12 of *tel-m1-1*, the 2-Mb distal part of the right arm was translocated to the

15   subtelomere 5L and the new right end of chromosome 12 (at 7.7 Mb from its initial extremity) carried

16   telomeres preceded by a tandem duplication consisting of a fragment from chromosome 12, 9

17   telomeric repeats forming an interstitial array and a short Sultan array from subtelomere 7L (Supp. Fig.

18   S5E).

19   In *tel-m2-1*, new telomeres were found at two close locations around the centromere of chromosome

20   7, at the extremities of a 90-kb region displaying a 2X sequencing depth, indicating that the centromere

21   was duplicated and the two arms of chromosome 7 split into two telocentric chromosomes (Fig. 5C),

22   thus constituting a Robertsonian fission, the reciprocal event of a Robertsonian fusion whereby two

23   telocentric chromosomes are combined through their long arm with the loss of the very short arms

24   and one centromere[55,56]. On the left end of the right arm, the new telomere capped an array of Sultan

1   7R lacking Spacer and a 65-bp ITS. On the right end of the left arm, the new telomere directly

2   transitioned into the centromere.

3   Overall, duplications, non-reciprocal translocations, deletions and more complex rearrangements

4   created new extremities outside canonical regions and altered the karyotype of the telomerase mutant

5   strains. For each new extremity, stability seemed to be ensured by newly recruited telomere

6   sequences.

7
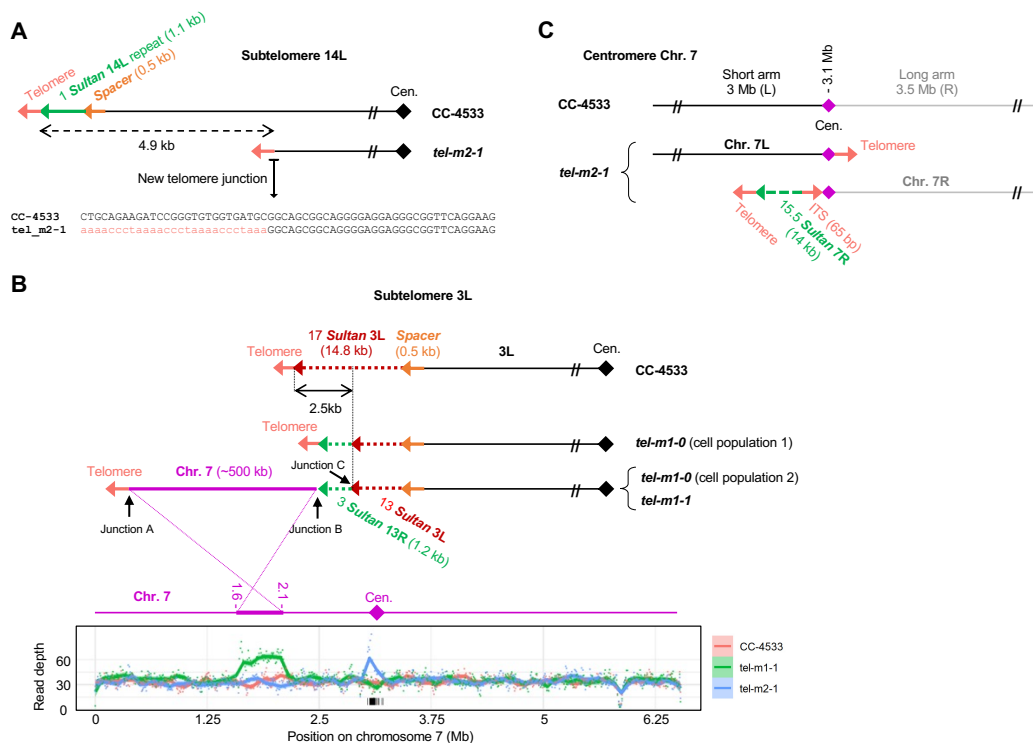
8

9   **Figure 5. New extremities without canonical subtelomeric elements.** (A) Scheme of the loss of 4.9 kb

10   of sequence and formation of a new telomere at subtelomere 14L of *tel-m2-1*. The new junction

11   sequence is shown. (B) Scheme of the complex new 3L extremity in *tel-m1-0* and *tel-m1-1*. The

12   sequencing depth for chromosome 7 is shown, revealing the duplicated 500-kb region in *tel-m1-1*. The

13   sequences of Junction A, B and C are shown in Supp. Fig. S5B-D. (C) Scheme of the Robertsonian fission

14   that occurred on chromosome 7 of *tel-m2-1*. The new centromere-proximal extremities were stabilized

15   by telomeres, with in addition a Sultan array and an ITS for the 7R long arm.

21

1

2   **DNA methylation is maintained at chromosome extremities and at displaced Sultan arrays**

3   We previously showed that Sultan arrays and to a lesser degree, Suber arrays, were hypermethylated,

4   whereas the two major rDNA clusters were hypomethylated except for a few telomere-proximal

5   repeats[24]. Sultan subtelomeres are also associated with the heterochromatin mark H3K9me1[57]. Since

6   subtelomeres in telomerase mutants were heavily rearranged, we asked whether DNA methylation

7   remained associated with the Sultan elements that were no longer close to chromosome extremities

8   and whether new extremities would acquire hypermethylation.

9   We thus basecalled 5-methylcytosines (5mC) at CpG sites using Nanopolish[58] and focused on

10   rearranged loci. We first investigated the complex rearrangement found at extremity 3L of *tel-m1-1*

11   where a duplicated 500 kb of chromosome 7 capped with telomere formed the new extremity and the

12   original 3L Sultan elements were thus located much more internally (> 500 kb) in the chromosome (Fig.

13   5B). To analyze 5mC content of the new extremity, we selected reads that aligned to the 500 kb of

14   chromosome 7 and at the same time contained the terminal telomere sequence, and in a second

15   control group, reads that aligned to the 500 kb region but also continued beyond on chromosome 7

16   (Fig. 6A and Supp. Fig. S6A). Strikingly, the new 3L extremity was highly methylated over a length of

17   ~12 kb, whereas the same sequence on its original locus on chromosome 7 was unmethylated, and so

18   was the same locus on the wild-type strain (Supp. Fig. S6A). The 3L Sultan elements as well as the

19   neighboring 13R Sultans in *tel-m1-1* maintained their high methylation status despite being now > 500

20   kb away from the extremity.

21   That new extremities were hypermethylated even though they did not contain canonical subtelomeric

22   sequences was confirmed at the truncated 14L extremity of *tel-m2-1* (Fig. 5A and 6B) and 7L of *tel-m1-*

23   *1* (Supp. Fig. S6B). Other examples of Sultan elements no longer located at subtelomeres but still

24   methylated included the 4L and 4R Sultan elements of the circularized chromosome 4 in *tel-m1-1* (Fig.

25   4E and 6C) and the head-to-head fused 9R Sultan elements, which resulted from the BFB event we

1   described in a previous section (Fig. 4D and 6D). In the latter example, the Sultan elements were

2   located at least 185 kb away from the new extremity. The reads mapping to the rDNA sequence at the

3   9R subtelomere showed that the rDNA was methylated (Fig. 6D), which in CC-1690 was only true for

4   the rDNA of subtelomere 1L and a few telomere-proximal rDNA sequences at 8R and 14R

5   subtelomeres, suggesting that the rDNA might constitute the new subtelomere of 9R in *tel-m1-1*.

6   Subtelomere 16R in *tel-m2-1* provided another example of translocated rDNA, now directly between

7   a terminal telomere sequence and Sultan elements, which was also hypermethylated (Supp. Fig. S6C).

8   Overall, the Sultan elements that were no longer near chromosome extremities maintained their

9   hypermethylation level. This observation was verified in all cases in *tel-m1-1* and *tel-m2-1*, for Sultan

10   arrays separated from the extremity by up to > 500 kb of other non-methylated sequences. Conversely,

11   sequences that became capped by telomeres and formed new extremities acquired a

12   hypermethylation pattern.

13

1
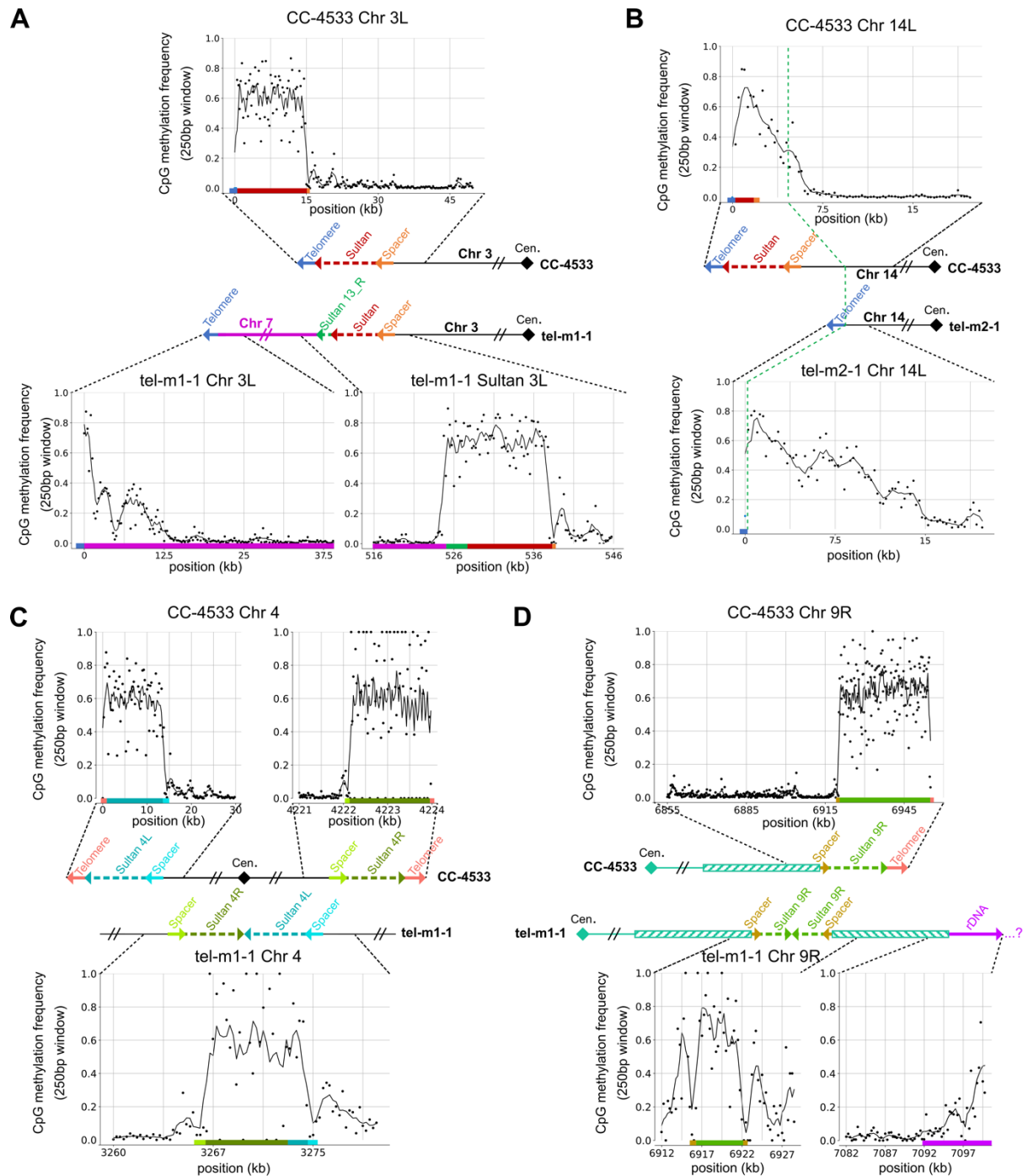
Figure 6. Methylation frequency of new extremities and displaced Sultan elements. (A-D) Analysis of 5mC frequency at CpG sites using reads that unambiguously spanned the indicated regions, in different illustrative cases. (A) At the rearranged 3L chromosome arm of *tel-m1-1*. Methylation frequency was plotted over different regions of subtelomere 3L in *tel-m1-1* and in CC-4533 as depicted. See also Supp. Fig. S6A for the methylation frequency at chromosome 7 in CC-4533 and *tel-m1-1*. (B) At the broken 14L subtelomere of *tel-m2-1*, where internal genomic sequence was then capped by a telomere, compared to CC-4533. (C) At the fused 4L and 4R

24

Sultan arrays of *tel-m1-1*, compared to the two arrays in CC-4533. (D) At the head-to-head fused 9R Sultan arrays and the transition to rDNA in *tel-m1-1*, compared to the 9R extremity of CC-4533.

Discussion

Telomeres and subtelomeres are extensively implicated in genome instability induced by telomere shortening or dysfunction, as exemplified in post-telomere-crisis tumors[13]. Many types of telomere-related rearrangements have been studied in experimental systems of different model organisms and in tumor samples from patients. However, a direct assessment and comprehensive picture of the repertoire of genome rearrangements that functional telomeres protect from have rarely been achieved. Here, we took advantage of long-read Nanopore sequencing to access telomere-induced SVs that would have escaped detection from short-read sequencing or at least would have been difficult to identify, particularly in repeated regions. We chose to sequence heterogeneous populations of telomerase-negative cells containing a complex mixture of subclonal genome rearrangements and to perform our analyses mainly at the level of reads, revealing mosaic rearrangement patterns that would have been missed in assembled clonal genomes.

Alternative telomere maintenance

We previously demonstrated that telomeres in *C. reinhardtii* are maintained by telomerase and that mutants of *CrTERT*, the gene encoding its catalytic subunit, displayed an "ever shorter telomere" phenotype[45]. We evidenced cell death consistent with replicative senescence in some telomerase mutant cultures but also that other long-term cultures failed to show any growth defect, suggesting that cells with alternative telomere maintenance mechanisms were selected. In this work, analysis of chromosome extremities in the long-term cultures of telomerase mutants showed that overall ~60%

25

1   were still capped by short telomeres, reminiscent of type I post-senescent survivors in *S. cerevisiae*

2   [37,59]. Interestingly, for at least a subset of them, the telomere sequence was likely recruited *de novo* on

3   a truncated subtelomere by non-reciprocal translocation, as shown before for cancer cell lines[60]. In all

4   cases, telomeres were maintained, albeit at a short length distribution, which implied that an

5   alternative maintenance mechanism was used, relying for example on homologous recombination or

6   BIR.

7   A number of chromosome extremities completely lacked telomere sequences and their lost telomeres

8   might have been translocated to the new telomere sites we described. Individual read analysis of such

9   telomere-less extremities showed that the Sultan array composing the subtelomere displayed a length

10   distribution at the population level consistent with progressive loss of sequence. This observation leads

11   to the intriguing possibility that the array of Sultan elements might directly function in end protection

12   instead of telomeres. Consistently, in end-to-end fusion events involving Sultan arrays (e.g. Fig. 4D and

13   E), fewer than 10 Sultan elements were left when fusion occurred, which might suggest that larger

14   Sultan arrays could confer some partial telomere protection. The Sultan element must then ensure

15   telomere protection by for example binding a yet unknown factor, which might be achieved through

16   the short telomere-like sequence at the beginning of the element[24] or through another intrinsic binding

17   site. Since Sultan elements are associated with heterochromatin, indirect binding of factors on

18   modified histones might also provide sufficient protection, as proposed for some telomerase-negative

19   *S. pombe* survivors[61]. Consistently, even when the chromosome extremities are not composed of

20   Sultan elements, but of rDNA repeats or even other sequences, the DNA sequence is highly methylated

21   (see discussion below).

22   Alternatively, *C. reinhardtii* might transiently tolerate unprotected or partially unprotected telomeres,

23   as a result of the lack of detectable DNA damage checkpoint kinases Chk1 and Chk2 orthologs or at

24   least differences in DNA damage response machineries compared to other Viridiplantae and to

25   Opisthokonts[62,63].

1   Another strategy to maintain chromosome integrity without telomerase consists in chromosome

2   circularization, which we observed for chromosome 4 in *tel-m1-1*. Although we encountered only one

3   such circularization event, the fact that we did not detect another subpopulation with another

4   structure for chromosome 4 suggests that this circular chromosome conferred some selective

5   advantage (e.g. telomerase-independent genome integrity) or at least was not counter-selected. A

6   similar circularization strategy in response to telomerase absence or disruption of telomere protection

7   was also described in other eukaryotes[38,64-67] and therefore appears to be relatively well-tolerated

8   across evolution.

9   Overall, although they came at the cost of wide-spread genome rearrangements, the diversity of

10   alternative strategies to maintain chromosome integrity in telomerase mutants found in a single

11   organism is remarkable and illustrates genome plasticity at short timescales.

12

13   **Mechanisms of telomere-shortening-induced genome rearrangements**

14   Because the approach used in this work consists in a direct visualization of chromosome sequences on

15   a large scale with as little source of bias as possible, we were able to document a striking variety of

16   rearrangements. The fact that most rearrangements involved telomeres and subtelomeres suggested

17   that critically short or dysfunctional telomeres preferentially induced local instabilities, which can

18   propagate through various mechanisms such as BFB cycles involving dicentrics, as reported in yeast

19   and human tumors[4,5,13,68-70].

20   Short or dysfunctional telomeres can lead to end-to-end fusions as observed across evolution[14,71-75],

21   although telomere sequences themselves may be completely absent at fusion sites in telomerase-

22   negative cells[3,4,38,64]. Despite many instances of end-to-end fusions, we found no remaining telomere

23   sequences at fusion sites in our read datasets, suggesting that in *C. reinhardtii*, even very short

24   telomeres are sufficient to inhibit fusions. Instead, we detected many subtelomere-subtelomere

25   fusions, the most frequent involving Sultan arrays. They can occur between different subtelomeres, as

1   in the case of chromosome 4 circularization, or between sister chromatid subtelomeres. Sister

2   chromatid fusions were followed by cycles of BFB, propagating genome instability over multiple

3   divisions and over more internal genomic regions, as for extremity 9R in *tel-m1-1*. Of note, we did not

4   find any signature of chromothripsis, resulting from the aberrant repair of the fragmentation of a

5   chromosomal region, despite chromothripsis being associated with BFB cycles and telomere crisis in

6   cancer[76,77].

7   Interestingly, most rearrangements implicated the repeated elements that are found at the

8   subtelomeres, which could promote homology-mediated mechanisms for amplification, deletion and

9   translocation. When different types of elements were found juxtaposed, the transition sequence often

10   involved microhomologies of a few base pairs (e.g. Fig. 5B and Supp. Fig. S5B-C), suggesting that

11   mechanisms such as microhomology-mediated break-induced replication (MMBIR)[78,79] or

12   microhomology-mediated end-joining (MMEJ)[80] might be involved. Even when no homology or

13   microhomology was detected at junctions, canonical subtelomeric elements were still frequently

14   found in rearrangements. We speculate that these subtelomeric elements were frequently excised

15   from their native locus due to terminal instability and were therefore available in subsequent fusion

16   events. Nevertheless, other internal genomic regions were sometimes duplicated and translocated to

17   chromosome ends, supporting the idea of a genome-wide increase in genome instability in

18   telomerase-negative cells[3,4,6,8,9].

19

20   **New chromosome extremities establish heterochromatin**

21   We found that new telomere-capped chromosome extremities formed by sequences originating from

22   internal genomic regions showed high levels of 5mC. We therefore speculate that telomeres

23   themselves are sufficient to establish new DNA methylation domains in *C. reinhardtii*, likely to be

24   associated with heterochromatin, stabilizing the chromosome end and protecting it from degradation

25   and repair. While the heterochromatic nature of telomeres and subtelomeres is conserved in

1 eukaryotes (except for *A. thaliana* where it is less well established[81-88]), whether and how

2 heterochromatin forms at a new telomere is less well known across organisms. The detection of new

3 extremities in our work thus provides an additional piece of evidence for heterochromatin formation

4 spreading from the telomere sequence being an evolutionarily conserved property of chromosome

5 extremities. Since many chromosome extremities were not capped by telomere sequences but directly

6 by the hypermethylated Sultan arrays, one can speculate that heterochromatin might be the primary

7 functional feature of a protected extremity, even more important than telomere sequences

8 themselves.

9 Interestingly, Sultan elements that are displaced from their normal subtelomeric location retain a high

10 level of 5mC, even when found at > 500 kb of the extremity of the chromosome. We previously showed

11 that Sultan elements are exclusively found at subtelomeres in wild-type strains. We thus propose that

12 their sequence coevolved with their function at subtelomeres, so that they might recruit

13 methyltransferases to actively contribute to heterochromatin maintenance and spread. This property

14 would allow then to remain hypermethylated even when displaced from subtelomeres. Interestingly,

15 in most cases, the Spacer sequence still acted as the boundary of the hypermethylated domain even

16 when not located at subtelomeres.

17 To conclude, our results show that *C. reinhardtii* telomerase-negative cells use a variety of strategies

18 to protect chromosome extremities, including the establishment of DNA methylation, and undergo

19 diverse and complex rearrangements, highlighting a remarkable plasticity of the genome. This work

20 demonstrates the potential of long-read sequencing to provide a comprehensive view of complex

21 genome rearrangements even in mosaic populations of cells at the level of subclones, an analysis that

22 could be applied to tumor genome heterogeneity.

23

24

1 Material and Methods

2

3 **Strains and growth conditions**

4 Wild-type CC-4533 and *tel-m1* and *tel-m2* mutant strains were obtained from the Jonikas lab[46]. They

5 were maintained on plates in TAP medium[89] at 25°C under low light (5 $\mu E.m^{-2}.s^{-1}$) and restreaked in

6 bulk without subcloning. Before DNA extraction, cells were grown in 200 mL liquid TAP medium until

7 they reached ~$2x10^7$ cell/mL and collected by centrifugation at 5000 g for 5 min.

8

9 **DNA extraction and size selection**

10 To extract genomic DNA while preserving high-molecular-weight fragments, we developed a modified

11 version of a Joint Genome Institute protocol (https://www.pacb.com/wp-

12 content/uploads/2015/09/DNA-extraction-chlamy-CTAB-JGI.pdf), as described here[47]. The extraction

13 is followed by a clean-up using magnetic beads and size selection based on solid-phase reversible

14 immobilization[90] using the SRE kit (Circulomics).

15

16 **Library preparation and sequencing**

17 Sequencing libraries were prepared from SRE-treated high-molecular-weight DNA (except for *tel-m1-*

18 *1* which was sequenced in two runs, one without SRE treatment and one with), following Oxford

19 Nanopore Technologies (ONT) protocols for genomic DNA without pre-amplification. Kits LSK109 and

20 LSK110 were obtained from ONT and Companion Module NebNext from New England Biolabs (NEB).

21 As a minor modification, we started the preparation with 3 μg of DNA. The second *tel-m1-1* library was

22 barcoded using barcode NB07 from kit NBD104 (ONT).

1   DNA libraries were sequenced on R9.4 or R10.4 Nanopore flow cells in a MinION Mk1C sequencer with

2   default parameters on the MinKNOW operating software, except for the basecalling with Guppy which

3   was set to real-time high accuracy and for the MUX scan which was decreased to 1 hr. For each run,

4   500 ng of DNA were loaded and sequenced for 8-12 hrs. The flow cell was then washed using the Wash

5   kit (ONT) and reloaded with the same sample, up to 4 times.

6

7   **Genome assembly**

8   Reads in FASTQ files and with quality >7 were processed using Porechop with default parameters

9   (https://github.com/rrwick/Porechop), for removing adapters/barcodes and splitting artefactual

10   chimeras. Reads were assembled using Canu[91], SMARTdenovo[92] and NextDenovo

11   (https://github.com/Nextomics/NextDenovo) using default parameters, and all chromosomes were

12   successfully represented by one main contig. These draft assemblies were polished using Racon[93]

13   (https://github.com/lbcb-sci/racon) and Medaka (ONT; https://github.com/nanoporetech/medaka)

14   with the same reads, then scaffolded on the CC-1690 reference genome using RagTag[94]

15   (https://github.com/malonge/RagTag). These assemblies were compared to each other and to

16   references (CC-1690, CC-4532) using D-genies[95] (https://github.com/genotoul-bioinfo/dgenies) to

17   assess contiguity. To generate a genome model for each strain, we chose for each chromosome the

18   assembly giving the most colinear chromosome with CC-1690.

19   To test if chromosome 4 of *tel-m1-1* was circular at the assembly level without relying on a reference

20   genome for scaffolding, we used the assembler Flye[96] (https://github.com/fenderglass/Flye), as it

21   generated contigs long enough for this purpose. Visualization of the circular chromosome 4 was done

22   using Bandage[97] (https://rrwick.github.io/Bandage/).

23

1   Telomere sequence detection

2   To detect telomere sequences in individual Nanopore reads, we developed TeloReader, a Python

3   (v3.9.12) script with only 3 dependencies (pandas (v1.4.3), numpy (v1.19.5) and matplotlib (v3.5.1)).

4   TeloReader scans each read in both directions, from 5' to 3' and from 3' to 5', to search for the C-rich

5   telomere motif and the G-rich one, respectively. In the first step of TeloReader, the DNA sequence is

6   transformed into a series of scores corresponding to the best alignment score (using pairwise2 from

7   the package biopython (v1.79)) of each 8-mer against all 8 circular permutations of the telomere motif

8   (CCCTAAAA for the C-rich and TTTTAGGG for the G-rich). This score is between 0 and 8, but all scores

9   lower than 4 are replaced by 4, to reduce the impact of sequencing errors in the following step. The

10   second step defines the sequence in a sliding window of size 15 bp (size_window) as telomeric if the

11   average score is greater or equal to 7 (min_mean_window). In a third step, consecutive overlapping

12   telomeric sliding windows are merged to form a single telomere sequence. We added other constraints

13   and rules to ensure the specificity and sensitivity of TeloReader: the minimal length of a telomere

14   sequence is set at 16 bp (min_len); a telomere sequence must contain at least 8 scores of 8; a

15   configuration where a non-telomeric sequence is found between 2 telomere sequences (as defined

16   after the third step) can be considered as a telomere sequence is the non-telomeric sequence is shorter

17   than 20 bp (max_size_gap) and if the average score of the whole sequence encompassing the 2

18   telomere sequences and the non-telomeric sequence in-between is greater than 6.5 (min_mean_telo).

19   A telomere sequence is considered as terminal if it is found at < 50 bp from the extremities of the input

20   sequence (3' extremity for a C-rich telomere and 5' extremity for a G-rich telomere). The code for

21   TeloReader   is   available   at: https://github.com/Telomere-Genome-

22   Stability/Telomere_2023/tree/main/TELOREADER.

23

1   **5-Methylcytosine detection**

2   To detect 5mC in a CpG context, we used nanopolish[58] (https://github.com/jts/nanopolish) on the

3   FAST5 files and their associated FASTQ files. Reads were aligned to their corresponding genome

4   assembly using Minimap2[98] (https://github.com/lh3/minimap2). Alternatively, for detection of 5mC in

5   subpopulations, reads were first selected using fast5_subset

6   (https://github.com/nanoporetech/ont_fast5_api) and aligned to a consensus sequence obtained

7   through a multiple alignment using Clustal W[99]. The frequency of methylation for each base was

8   obtained using nanopolish call-methylation and the script calculate_methylation_frequency.py.

9

10   **Bioinformatic analysis**

11   For genome-to-genome comparison, genome models were aligned against references genomes (CC-

12   1690 and CC-4533) using Minimap2[98] and visualized with Circos plots[48]. SV caller MUM&Co[53]

13   (https://github.com/SAMtoBAM/MUMandCo) was used for structural variant detection and

14   classification. For read analysis, reads were mapped to genomes using Minimap2 with parameters: -a

15   -x map-ont -K 5M -t 3. We used Integrative Genome Viewer (IGV)[100] (https://igv.org/) and Tablet[101]

16   (https://ics.hutton.ac.uk/tablet/) to visualize read mapping. To calculate sequencing depth, the

17   genomes were divided into windows using 'makewindows' from BEDTools[102]

18   (https://bedtools.readthedocs.io/en/latest/index.html), then depth was averaged over the windows

19   using 'bedcov' from Samtools[103] (https://www.htslib.org/) with parameters: -g SUPPLEMENTARY with

20   or without -q 60.

21   For read level analysis, subtelomeric elements (Spacer, Sultan, rDNA and Suber) were searched in reads

22   using blastn[104] with parameters: -max_target_seqs 10000 -evalue 0.001 -outfmt "6 qaccver qlen

23   saccver slen pident length mismatch gapopen qstart qend sstart send evalue bitscore". Multiple hits

24   at close positions were filtered using a custom bash script: hits were sorted by decreasing order of

25   bitscore (*sort* -rgk 14) and overlapping matches were eliminated. Reads containing subtelomeric

33

1   elements or telomeres (detected by TeloReader) were further blasted against a library of repeated

2   elements[105] and against reference genome, both of which were then filtered as above to exclude

3   multiple matches. Annotations from all blast hits were then plotted on reads using R. Plots were

4   created using an R script based on dplyr, tidyr and ggplot2 libraries. All lists of filtered blast hits lists

5   were merged with 'bind_rows', reads of interest were identified (e.g. match to a given Spacer or

6   Sultan) with 'filter' and all hits on this read subset were extracted with 'semi_join'. To order these

7   reads, first, an anchor was selected, most often a specific Spacer sequence, allowing to set an origin

8   and the orientation of each read. From these two pieces of information, new coordinates were re-

9   calculated for each hit along the reads and for read extremities. Hits were then plotted with distinct

10   reads along the y-axis and coordinates along the x-axis (e.g. Fig. 3A, 3D, 4C-E). Due to the error rate of

11   Nanopore sequencing at the raw read level together with rare artefactual chimeric reads[106], we only

12   describe rearrangements structurally supported by at least two reads.

13   Computations were mostly run on the cluster of the French Institute of Bioinformatics

14   (https://www.france-bioinformatique.fr/en/home/).

15

Acknowledgement

23

1   Author contributions

2   Investigation: FC, NA, CG, SE; Conceptualization: FC, GF, SE, ZX; Methodology: FC, NA, CG; Software:

3   FC, CG; Formal analysis: FC, NA, CG, SE; Supervision: GF, ZX; Writing – Original Draft: FC, ZX; Writing –

4   Review & Editing: all authors.

5

6   Conflicts of interest

7   None declared.

8

9   References

10   1   Jain, D. & Cooper, J. P. Telomeric strategies: means to an end. *Annual Review of Genetics* **44**,
11   243-269, doi:10.1146/annurev-genet-102108-134841 (2010).
12   2   de Lange, T. Shelterin-Mediated Telomere Protection. *Annu. Rev. Genet.* **52**, 223-247,
13   doi:10.1146/annurev-genet-032918-021921 (2018).
14   3   Blasco, M. A. *et al.* Telomere shortening and tumor formation by mouse cells lacking
15   telomerase RNA. *Cell* **91**, 25-34 (1997).
16   4   Hackett, J. A., Feldser, D. M. & Greider, C. W. Telomere dysfunction increases mutation rate
17   and genomic instability. *Cell* **106**, 275-286. (2001).
18   5   Hackett, J. A. & Greider, C. W. End resection initiates genomic instability in the absence of
19   telomerase. *Mol. Cell. Biol.* **23**, 8450-8461 (2003).
20   6   Coutelier, H. *et al.* Adaptation to DNA damage checkpoint in senescent telomerase-negative
21   cells promotes genome instability. *Genes Dev.* **32**, 1499-1513, doi:10.1101/gad.318485.118
22   (2018).
23   7   Maciejowski, J., Li, Y., Bosco, N., Campbell, P. J. & de Lange, T. Chromothripsis and Kataegis
24   Induced by Telomere Crisis. *Cell* **163**, 1641-1654, doi:10.1016/j.cell.2015.11.054 (2015).
25   8   Chin, L. *et al.* p53 deficiency rescues the adverse effects of telomere loss and cooperates with
26   telomere dysfunction to accelerate carcinogenesis. *Cell* **97**, 527-538 (1999).
27   9   Lee, H. W. *et al.* Essential role of mouse telomerase in highly proliferative organs. *Nature* **392**,
28   569-574 (1998).
29   10   Artandi, S. E. *et al.* Telomere dysfunction promotes non-reciprocal translocations and
30   epithelial cancers in mice. *Nature* **406**, 641-645, doi:10.1038/35020592 (2000).
31   11   Henninger, E. & Teixeira, M. T. Telomere-driven mutational processes in yeast. *Curr. Opin.*
32   *Genet. Dev.* **60**, 99-106, doi:10.1016/j.gde.2020.02.018 (2020).
33   12   Artandi, S. E. & Depinho, R. A. Telomeres and telomerase in cancer. *Carcinogenesis*, doi:bgp268
34   (2009).
35   13   Maciejowski, J. & de Lange, T. Telomeres in cancer: tumour suppression and genome
36   instability. *Nat Rev Mol Cell Biol* **18**, 175-186, doi:10.1038/nrm.2016.171 (2017).
37   14   Capper, R. *et al.* The nature of telomere fusion and a definition of the critical telomere length
38   in human cells. *Genes Dev.* **21**, 2495-2508 (2007).

15 Jones, R. E. *et al.* Escape from telomere-driven crisis is DNA ligase III dependent. *Cell Rep* **8**, 1063-1076, doi:10.1016/j.celrep.2014.07.007 (2014).

16 McClintock, B. The Stability of Broken Ends of Chromosomes in Zea Mays. *Genetics* **26**, 234-282, doi:10.1093/genetics/26.2.234 (1941).

17 Davoli, T., Denchi, E. L. & de Lange, T. Persistent telomere damage induces bypass of mitosis and tetraploidy. *Cell* **141**, 81-93, doi:S0092-8674(10)00066-8 (2010).

18 Yue, J. X. *et al.* Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat. Genet.* **49**, 913-924, doi:10.1038/ng.3847 (2017).

19 Corcoran, L. M., Thompson, J. K., Walliker, D. & Kemp, D. J. Homologous recombination within subtelomeric repeat sequences generates chromosome size polymorphisms in P. falciparum. *Cell* **53**, 807-813, doi:10.1016/0092-8674(88)90097-9 (1988).

20 Louis, E. J. The chromosome ends of Saccharomyces cerevisiae. *Yeast* **11**, 1553-1573 (1995).

21 Kim, J. M., Vanguri, S., Boeke, J. D., Gabriel, A. & Voytas, D. F. Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete Saccharomyces cerevisiae genome sequence. *Genome Res.* **8**, 464-478 (1998).

22 Fabre, E. *et al.* Comparative genomics in hemiascomycete yeasts: evolution of sex, silencing, and subtelomeres. *Mol. Biol. Evol.* **22**, 856-873 (2005).

23 Richard, M. M. *et al.* The Subtelomeric khipu Satellite Repeat from Phaseolus vulgaris: Lessons Learned from the Genome Analysis of the Andean Genotype G19833. *Front Plant Sci* **4**, 109, doi:10.3389/fpls.2013.00109 (2013).

24 Chaux-Jukic, F. *et al.* Architecture and evolution of subtelomeres in the unicellular green alga Chlamydomonas reinhardtii. *Nucleic Acids Res* **49**, 7571-7587, doi:10.1093/nar/gkab534 (2021).

25 Takikawa, M., Tarumoto, Y. & Ishikawa, F. Fission yeast Stn1 is crucial for semi-conservative replication at telomeres and subtelomeres. *Nucleic Acids Res* **45**, 1255-1269, doi:10.1093/nar/gkw1176 (2017).

26 Chen, N. W. G. *et al.* Common Bean Subtelomeres Are Hot Spots of Recombination and Favor Resistance Gene Evolution. *Front Plant Sci* **9**, 1185, doi:10.3389/fpls.2018.01185 (2018).

27 Kim, C. *et al.* Long-read sequencing reveals intra-species tolerance of substantial structural variations and new subtelomere formation in C. elegans. *Genome Res.* **29**, 1023-1035, doi:10.1101/gr.246082.118 (2019).

28 Louis, E. J. & Haber, J. E. Mitotic recombination among subtelomeric Y' repeats in Saccharomyces cerevisiae. *Genetics* **124**, 547-559. (1990).

29 Linardopoulou, E. V. *et al.* Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. *Nature* **437**, 94-100 (2005).

30 Rudd, M. K. *et al.* Elevated rates of sister chromatid exchange at chromosome ends. *PLoS Genet* **3**, e32, doi:10.1371/journal.pgen.0030032 (2007).

31 Kuo, H. F., Olsen, K. M. & Richards, E. J. Natural variation in a subtelomeric region of Arabidopsis: implications for the genomic dynamics of a chromosome end. *Genetics* (2006).

32 Maestroni, L. *et al.* Eroded telomeres are rearranged in quiescent fission yeast cells through duplications of subtelomeric sequences. *Nat Commun* **8**, 1684, doi:10.1038/s41467-017-01894-6 (2017).

33 Otto, T. D. *et al.* Long read assemblies of geographically dispersed Plasmodium falciparum isolates reveal highly structured subtelomeres. *Wellcome Open Res* **3**, 52, doi:10.12688/wellcomeopenres.14571.1 (2018).

34 Anderson, J. A., Song, Y. S. & Langley, C. H. Molecular population genetics of Drosophila subtelomeric DNA. *Genetics* **178**, 477-487, doi:10.1534/genetics.107.083196 (2008).

35 Young, E., Abid, H. Z., Kwok, P. Y., Riethman, H. & Xiao, M. Comprehensive Analysis of Human Subtelomeres by Whole Genome Mapping. *PLoS Genet* **16**, e1008347, doi:10.1371/journal.pgen.1008347 (2020).

36 Cesare, A. J. & Reddel, R. R. Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet* **11**, 319-330, doi:10.1038/nrg2763 (2010).

37   Lundblad, V. & Blackburn, E. H. An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell* **73**, 347-360 (1993).

38   Nakamura, T. M., Cooper, J. P. & Cech, T. R. Two modes of survival of fission yeast without telomerase. *Science* **282**, 493-496 (1998).

39   Zellinger, B., Akimcheva, S., Puizina, J., Schirato, M. & Riha, K. Ku suppresses formation of telomeric circles and alternative telomere lengthening in Arabidopsis. *Mol. Cell* **27**, 163-169 (2007).

40   Ho, S. S., Urban, A. E. & Mills, R. E. Structural variation in the sequencing era. *Nat Rev Genet* **21**, 171-189, doi:10.1038/s41576-019-0180-9 (2020).

41   Kockler, Z. W., Comeron, J. M. & Malkova, A. A unified alternative telomere-lengthening pathway in yeast survivor cells. *Mol. Cell*, doi:10.1016/j.molcel.2021.02.004 (2021).

42   Sholes, S. L. *et al.* Chromosome-specific telomere lengths and the minimal functional telomere revealed by nanopore sequencing. *Genome Res.*, doi:10.1101/gr.275868.121 (2021).

43   Kim, E., Kim, J., Kim, C. & Lee, J. Long-read sequencing and de novo genome assemblies reveal complex chromosome end structures caused by telomere dysfunction at the single nucleotide level. *Nucleic Acids Res* **49**, 3338-3353, doi:10.1093/nar/gkab141 (2021).

44   Hails, T., Jobling, M. & Day, A. Large arrays of tandemly repeated DNA sequences in the green alga Chlamydomonas reinhardtii. *Chromosoma* **102**, 500-507, doi:10.1007/BF00357106 (1993).

45   Eberhard, S. *et al.* Molecular characterization of Chlamydomonas reinhardtii telomeres and telomerase mutants. *Life science alliance* **2**, doi:10.26508/lsa.201900315 (2019).

46   Li, X. *et al.* An Indexed, Mapped Mutant Library Enables Reverse Genetics Studies of Biological Processes in Chlamydomonas reinhardtii. *Plant Cell* **28**, 367-387, doi:10.1105/tpc.15.00465 (2016).

47   Chaux-Jukic, F., Agier, N., Eberhard, S. & Xu, Z. Extraction and selection of high-molecular-weight DNA for long-read sequencing from Chlamydomonas reinhardtii. *bioRxiv*, doi:https://doi.org/10.1101/2022.08.16.504088 (2022).

48   Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639-1645, doi:10.1101/gr.092759.109 (2009).

49   O'Donnell, S., Chaux, F. & Fischer, G. Highly Contiguous Nanopore Genome Assembly of Chlamydomonas reinhardtii CC-1690. *Microbiol Resour Announc* **9**, doi:10.1128/MRA.00726-20 (2020).

50   Gallaher, S. D., Fitz-Gibbon, S. T., Glaesener, A. G., Pellegrini, M. & Merchant, S. S. Chlamydomonas Genome Resource for Laboratory Strains Reveals a Mosaic of Sequence Variation, Identifies True Strain Histories, and Enables Strain-Specific Studies. *Plant Cell* **27**, 2335-2352, doi:10.1105/tpc.15.00508 (2015).

51   Flowers, J. M. *et al.* Whole-Genome Resequencing Reveals Extensive Natural Variation in the Model Green Alga Chlamydomonas reinhardtii. *Plant Cell* **27**, 2353-2369, doi:10.1105/tpc.15.00492 (2015).

52   Craig, R. J. *et al.* The Chlamydomonas Genome Project, version 6: reference assemblies for mating type plus and minus strains reveal extensive structural mutation in the laboratory. *Plant Cell*, doi:10.1093/plcell/koac347 (2022).

53   O'Donnell, S. & Fischer, G. MUM&Co: accurate detection of all SV types through whole-genome alignment. *Bioinformatics* **36**, 3242-3243, doi:10.1093/bioinformatics/btaa115 (2020).

54   Filloramo, G. V., Curtis, B. A., Blanche, E. & Archibald, J. M. Re-examination of two diatom reference genomes using long-read sequencing. *BMC Genomics* **22**, 379, doi:10.1186/s12864-021-07666-3 (2021).

55   Robertson, W. R. B. Chromosome studies. I. Taxonomic relationships shown in the chromosomes of tettigidae and acrididae: V-shaped chromosomes and their significance in acrididae, locustidae, and gryllidae: Chromosomes and variation. *Journal of Morphology* **27**, 179-331, doi:https://doi.org/10.1002/jmor.1050270202 (1916).

56   Jones, K. Robertsonian fusion and centric fission in karyotype evolution of higher plants. *The Botanical Review* **64**, 273-289, doi:10.1007/BF02856567 (1998).

57   Strenkert, D., Schmollinger, S. & Schroda, M. Heat shock factor 1 counteracts epigenetic silencing of nuclear transgenes in Chlamydomonas reinhardtii. *Nucleic Acids Res* **41**, 5273-5289, doi:10.1093/nar/gkt224 (2013).

58   Simpson, J. T. *et al.* Detecting DNA cytosine methylation using nanopore sequencing. *Nat Methods* **14**, 407-410, doi:10.1038/nmeth.4184 (2017).

59   Teng, S. C. & Zakian, V. A. Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **19**, 8083-8093 (1999).

60   Sabatier, L., Ricoul, M., Pottier, G. & Murnane, J. P. The loss of a single telomere can result in instability of multiple chromosomes in a human tumor cell line. *Mol Cancer Res* **3**, 139-150, doi:10.1158/1541-7786.MCR-04-0194 (2005).

61   Jain, D., Hebden, A. K., Nakamura, T. M., Miller, K. M. & Cooper, J. P. HAATI survivors replace canonical telomeres with blocks of generic heterochromatin. *Nature* **467**, 223-227, doi:nature09374 (2010).

62   Cross, F. R. & Umen, J. G. The Chlamydomonas cell cycle. *Plant J.* **82**, 370-392, doi:10.1111/tpj.12795 (2015).

63   Fauser, F. *et al.* Systematic characterization of gene function in the photosynthetic alga Chlamydomonas reinhardtii. *Nat. Genet.* **54**, 705-714, doi:10.1038/s41588-022-01052-9 (2022).

64   Naito, T., Matsuura, A. & Ishikawa, F. Circular chromosome formation in a fission yeast mutant defective in two ATM homologues. *Nat. Genet.* **20**, 203-206 (1998).

65   McEachern, M. J., Iyer, S., Fulton, T. B. & Blackburn, E. H. Telomere fusions caused by mutating the terminal region of telomeric DNA. *PNAS* **97**, 11409-11414 (2000).

66   Baumann, P. & Cech, T. R. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* **292**, 1171-1175 (2001).

67   Wu, Z. J. *et al.* Cdc13 is predominant over Stn1 and Ten1 in preventing chromosome end fusions. *eLife* **9**, doi:10.7554/eLife.53144 (2020).

68   Beyer, T. & Weinert, T. Ontogeny of Unstable Chromosomes Generated by Telomere Error in Budding Yeast. *PLoS Genet* **12**, e1006345, doi:10.1371/journal.pgen.1006345 (2016).

69   Pobiega, S. & Marcand, S. Dicentric breakage at telomere fusions. *Genes Dev.* **24**, 720-733, doi:24/7/720 (2010).

70   Umbreit, N. T. *et al.* Mechanisms generating cancer genome complexity from a single cell division error. *Science* **368**, doi:10.1126/science.aba0712 (2020).

71   Chan, S. W. & Blackburn, E. H. Telomerase and ATM/Tel1p protect telomeres from nonhomologous end joining. *Mol. Cell* **11**, 1379-1387 (2003).

72   Mieczkowski, P. A., Mieczkowska, J. O., Dominska, M. & Petes, T. D. Genetic regulation of telomere-telomere fusions in the yeast Saccharomyces cerevisae. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 10854-10859 (2003).

73   Heacock, M., Spangler, E., Riha, K., Puizina, J. & Shippen, D. E. Molecular analysis of telomere fusions in Arabidopsis: multiple pathways for chromosome end-joining. *EMBO J.* **23**, 2304-2313 (2004).

74   Pardo, B. & Marcand, S. Rap1 prevents telomere fusions by nonhomologous end joining. *EMBO J.* (2005).

75   Lowden, M. R., Meier, B., Lee, T. W., Hall, J. & Ahmed, S. End joining at Caenorhabditis elegans telomeres. *Genetics* **180**, 741-754 (2008).

76   Li, Y. *et al.* Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia. *Nature* **508**, 98-102, doi:10.1038/nature13115 (2014).

77   Garsed, D. W. *et al.* The architecture and evolution of cancer neochromosomes. *Cancer Cell* **26**, 653-667, doi:10.1016/j.ccell.2014.09.010 (2014).

78   Payen, C., Koszul, R., Dujon, B. & Fischer, G. Segmental duplications arise from Pol32-dependent repair of broken forks through two alternative replication-based mechanisms. *PLoS Genet* **4**, e1000175, doi:10.1371/journal.pgen.1000175 (2008).

79   Hastings, P. J., Ira, G. & Lupski, J. R. A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet* **5**, e1000327, doi:10.1371/journal.pgen.1000327 (2009).

80   Sfeir, A. & Symington, L. S. Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends Biochem. Sci.* **40**, 701-714, doi:10.1016/j.tibs.2015.08.006 (2015).

81   Elgin, S. C. & Reuter, G. Position-effect variegation, heterochromatin formation, and gene silencing in Drosophila. *Cold Spring Harb Perspect Biol* **5**, a017780, doi:10.1101/cshperspect.a017780 (2013).

82   Matsuda, A. *et al.* Highly condensed chromatins are formed adjacent to subtelomeric and decondensed silent chromatin in fission yeast. *Nat Commun* **6**, 7753, doi:10.1038/ncomms8753 (2015).

83   Koering, C. E. *et al.* Human telomeric position effect is determined by chromosomal context and telomeric chromatin integrity. *EMBO Rep* **3**, 1055-1061 (2002).

84   Pedram, M. *et al.* Telomere position effect and silencing of transgenes near telomeres in the mouse. *Mol. Cell. Biol.* **26**, 1865-1878 (2006).

85   Gottschling, D. E., Aparicio, O. M., Billington, B. L. & Zakian, V. A. Position effect at S. cerevisiae telomeres: reversible repression of Pol II transcription. *Cell* **63**, 751-762 (1990).

86   Baur, J. A., Zou, Y., Shay, J. W. & Wright, W. E. Telomere position effect in human cells. *Science* **292**, 2075-2077 (2001).

87   Vrbsky, J. *et al.* siRNA-mediated methylation of Arabidopsis telomeres. *PLoS Genet* **6**, e1000986, doi:10.1371/journal.pgen.1000986 (2010).

88   Vaquero-Sedas, M. I., Gamez-Arjona, F. M. & Vega-Palas, M. A. Arabidopsis thaliana telomeres exhibit euchromatic features. *Nucleic Acids Res* **39**, 2007-2017, doi:10.1093/nar/gkq1119 (2011).

89   Harris, E. H. *The Chlamydomonas Sourcebook*. (Elsevier/Academic Press, 2009).

90   Stortchevoi, A., Kamelamela, N. & Levine, S. S. SPRI Beads-based Size Selection in the Range of 2-10kb. *J Biomol Tech* **31**, 7-10, doi:10.7171/jbt.20-3101-002 (2020).

91   Koren, S. *et al.* Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722-736, doi:10.1101/gr.215087.116 (2017).

92   Liu, H., Wu, S., Li, A. & Ruan, J. SMARTdenovo: a de novo assembler using long noisy reads. *Gigabyte* **2021** (2021).

93   Vaser, R., Sovic, I., Nagarajan, N. & Sikic, M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* **27**, 737-746, doi:10.1101/gr.214270.116 (2017).

94   Alonge, M. *et al.* Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol* **23**, 258, doi:10.1186/s13059-022-02823-7 (2022).

95   Cabanettes, F. & Klopp, C. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. *PeerJ* **6**, e4958, doi:10.7717/peerj.4958 (2018).

96   Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* **37**, 540-546, doi:10.1038/s41587-019-0072-8 (2019).

97   Wick, R. R., Schultz, M. B., Zobel, J. & Holt, K. E. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* **31**, 3350-3352, doi:10.1093/bioinformatics/btv383 (2015).

98   Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100, doi:10.1093/bioinformatics/bty191 (2018).

99   Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948, doi:10.1093/bioinformatics/btm404 (2007).

100   Robinson, J. T., Thorvaldsdottir, H., Wenger, A. M., Zehir, A. & Mesirov, J. P. Variant Review with the Integrative Genomics Viewer. *Cancer Res.* **77**, e31-e34, doi:10.1158/0008-5472.CAN-17-0337 (2017).

101   Milne, I. *et al.* Using Tablet for visual exploration of second-generation sequencing data. *Brief Bioinform* **14**, 193-202, doi:10.1093/bib/bbs012 (2013).

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

103   Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

104   Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410, doi:10.1016/S0022-2836(05)80360-2 (1990).

105   Craig, R. J., Hasan, A. R., Ness, R. W. & Keightley, P. D. Comparative genomics of Chlamydomonas. *Plant Cell*, doi:10.1093/plcell/koab026 (2021).

106   Delahaye, C. & Nicolas, J. Sequencing DNA with nanopores: Troubles and biases. *PLoS One* **16**, e0257521, doi:10.1371/journal.pone.0257521 (2021).