

A bioinformatics approach to identify telomere sequences

Indira Somanathan¹ & Chris Baysdorfer^{*1}

¹Department of Biological Sciences, California State University East Bay, Hayward, CA, USA

BioTechniques 65: 20-25 (July 2018) 10.2144/btn-2018-0057

Keywords: bioinformatics • next-generation sequencing • NGS • sequence repeat finder • SERF • tandem repeats • telomere

Conventional approaches to identify a telomere motif in a new genome are laborious and time-intensive. An efficient new methodology based on next-generation sequencing (NGS), *de novo* sequence repeat finder (SERF) and fluorescence *in situ* hybridization (FISH) is presented. Unlike existing heuristic approaches, SERF utilizes an exhaustive analysis of raw NGS reads or assembled contigs for rapid *de novo* detection of conserved tandem repeats representing telomere motifs. SERF was validated using the NGS data from *Ipheion uniflorum* and *Allium cepa* with known telomere motifs. The analysis program was then used on NGS data to investigate the telomere motifs in several additional plant species and together with FISH proved to be an efficient approach to identify as yet unknown telomere motifs.

Introduction

Eukaryotic linear chromosome ends are characterized by telomeres, which are generally made up of long tracts of minisatellite tandem repeats [1]. Telomeres are maintained by telomerase, a reverse transcriptase-based end maintenance system [2]. Telomeres help to prevent the loss of essential genetic information from the chromosome ends resulting from the incomplete replication of the lagging strand of chromosomal DNA and help to avoid detrimental chromosomal end-to-end fusions by distinguishing the double-stranded DNA (dsDNA) breaks from the native ends of the chromosomes [3]. Telomeres, together with their associated proteins, secure and stop the degradation of the chromosomal ends by various nucleolytic enzymes, as reviewed in de Lange [4]. Thus, telomeres maintain the integrity and stability of chromosomes.

The telomere repeat motif (TTTAGGG)_n, discovered in *Arabidopsis thaliana* [5], is considered to be the consensus telomere motif for most land plants [6]. Similarly, other groups of organisms have shown

conservation of their telomere motifs. However, there are a few exceptions to the ancestral-type consensus telomere motif. *Drosophila melanogaster*, a dipteran, lacks telomerase and the insect-type telomere repeat motif (TTAGG)_n [7,8]. Its telomeres have been replaced by three different types of telomere-specific non-long terminal repeat (non-LTR) retrotransposons [9], namely, *HeT-A* [10], *TART* [11] and *TAHRE* [12], which perform the telomere functions. In various species of algae, telomere repeat motif sequences are seen to differ slightly from the land plant telomere motif (TTTAGGG)_n [13]. For example, in the green alga *Chlamydomonas reinhardtii*, a more T-rich telomere repeat motif (TTTTAGGG)_n has been detected [14]. A telomere repeat motif switch identical to that found in *Chlamydomonas*-type was also seen in a few species of algae from the genus *Klebsormidium* [15,16].

Telomere motif evolution has also been demonstrated across various groups of land plants [17]. The *Arabidopsis*-type telomere is lost in *Cestrum elegans* from the family *Solanaceae* [18] and is replaced

by an unusual 10 base pair (bp) telomere motif (TTTTTTAGGG)_n maintained by telomerase [19]. Yet, many model plant species for telomere studies from this family, such as *Nicotiana tabacum*, *Solanum lycopersicum* and *Solanum tuberosum*, carry the ancestral-type telomere motif (TTTAGGG)_n at their chromosome termini [20–22]. Telomere divergence has also been shown to occur in monocots, especially in the order Asparagales. In the family *Asphodelaceae*, plants of the genus *Aloe* lack the *Arabidopsis*-type (TTTAGGG)_n telomere motif [23] and instead have the human-type (TTAGGG)_n [24]. Different types of telomere motifs are discovered in the family *Amaryllidaceae*. For example, in *Ipheion uniflorum*, the canonical *Arabidopsis*-type telomere motif is replaced by human-type [25], while in *Allium cepa* and related species, a unique 12 base pair telomere motif (CTCGGTATGGG)_n [26] has replaced both the *Arabidopsis*-type (TTTAGGG)_n [27,28], and the human-type (TTAGGG)_n [25,29] telomere motifs. It took more than a decade of extensive research to finally identify the distinct telomere motifs in the genera *Cestrum* and *Allium*.

METHOD SUMMARY

Here we introduce a bioinformatics-based methodology for the rapid *de novo* detection of the most probable telomere motifs in genomes of any size. Raw unassembled genomic data from next-generation sequencers are analyzed using the newly developed SERF *de novo* sequence analysis algorithm to identify a culled set of putative telomere motifs based on their biological characteristics. The cytological validation of these motifs using fluorescence *in situ* hybridization followed by its detection using confocal microscopy can expedite the identification of true telomere sequence with fewer wet-lab iterations.

A number of specific approaches to quantify human telomere length [30] from whole genome sequencing (WGS) data use the known telomere motif (TTAGGG)_n. The few *de novo* repeat detection tools [31–33] available are not equipped with a probable telomere motif detection algorithm from the NGS sampling data or WGS data. Consequently, efforts to detect telomere motifs employ time-intensive laboratory protocols and complex computational methods [26,34].

Future efforts to identify as yet unknown telomere motifs in a variety of genomes can be expedited using bioinformatics. We present a simple and efficient methodology that utilizes NGS data in conjunction with sequence repeat finder (SERF) algorithms, which is tailored towards identifying the probable conserved telomere repeat motifs and their subsequent cytological validation using fluorescence *in situ* hybridization (FISH) [35] (Figure 1).

Materials & methods

Plant materials

Plant materials used in this research include *Ipheion uniflorum*, *Allium cepa*, *Narcissus jonquilla* and *Muscari comosum*, obtained from commercial sources.

Isolation of total genomic DNA

Total genomic DNA was isolated from young leaves of these plants using the Qiagen DNeasy plant mini kit (Hilden, Germany) as per manufacturer's protocols. Approximately 100 mg (wet weight) of the leaf tissue was used for each of the isolations.

Next-generation sequencing

Genome sampling data was collected in multiple runs from *A. cepa* (GenBank Sequence Read Archive [SRA] accession numbers: SRX3381812, SRX3381811 and SRX3381814) and *N. jonquilla* (GenBank SRA accession numbers: SRX3381816 and SRX3381815) and in single runs from *I. uniflorum* (GenBank SRA accession

number: SRX3381813) and *M. comosum* (GenBank SRA accession number: SRX3381817) using the Ion Torrent next-generation sequencing (NGS) platform (Life Technologies, CA, USA), as per manufacturer's instructions. *De novo* assemblies for *A. cepa* and *I. uniflorum* were performed using the software SeqMan NGen (www.dnastar.com). Previously sampled plant genomes from more than 20 species belonging to the families *Solanaceae*, *Liliaceae*, *Smilacaceae* and *Velloziaceae* (Unpublished Data) were also used to validate the telomere motif identification methodology (see Table 4).

SERF *de novo* telomere sequence analysis

Consensus sequences from the whole genome assemblies in FASTA format and the raw NGS reads data in FASTQ format were separately analyzed by the sequence repeat finder (SERF) analysis platform (bioserf.org) to *de novo* identify and extract contiguous conserved tandem repeats varying in size from 5–15 base pairs. Genomes of *I. uniflorum* with human-type (TTAGGG)_n [25] and *A. cepa* with a unique 12 base pair (CTCGGTTATGGG)_n [26] telomere repeat motifs were used as positive controls for the initial validation of the SERF algorithm.

Homopolymeric nucleotide tracts and supersets of repeats were filtered out from the analysis results. This analysis looked for a pattern of organization (contiguous versus scattered), positional distribution of the repeats within the contigs or reads, and the frequency of the conserved repeat motifs in these sequences. The quality metric concept employed in SERF reflected the percentage of the individual read or contig length occupied by the identified contiguous repeats. A similar approach was used in *A. cepa* telomere motif identification [26]. Permutative combinations were analyzed, and repeats were clustered based on the size, position and type of base pair. Based on the assumptions that telomeres are poorly represented in genome sampling, and that the reads generated using the Ion

Torrent are generally short in length, an initial protocol of 2-5-15-90-TG-contiguous (minimum number of contiguous repeats–minimum base pair size–maximum base pair size–90% quality metric–TG richness–contiguous mode) was employed. Emphasis was given to the quality metric rather than the number of tandem telomere motifs to analyze the assembled contigs in *I. uniflorum*. An identical protocol successfully identified the respective telomere motifs from unassembled NGS reads of *I. uniflorum* and *A. cepa* with a high quality metric. The same protocol was used on all other species tested. Additional information on SERF can be found in the Supplementary material (Figure S1) and from bioserf.org.

Probe synthesis for FISH on *N. jonquilla*

11-dUTP biotinylated probes were synthesized using PCR with *N. jonquilla* genomic DNA for (A) the unique 15 bp (GAGGGTGCCTCAG)_n and (B) human-type (TTAGGG)_n repeats identified by SERF analysis on its raw NGS reads. The following primers were used: for repeat (A): 5' ACTCACGGAGGGTGCCTCAG 3' (for), 5' CGTGAGTGCACCTCCGTGAGT 3' (rev), and for repeat (B): 5' TTAGGGTTAGGGTTAGGGTTAGGG 3' (for) and 5' CCCTAACCTAACCTAACCTAA 3' (rev). The lengths of the probes were between 200 and 400 bp.

Preparation of metaphase chromosome spreads of *N. jonquilla*

Young root tips (3–5 days old) from *N. jonquilla* bulbs were arrested in metaphase by ice cold water treatment, fixed in freshly prepared Carnoy's solution I, and spread by squashing [36]. Metaphase chromosome spreads on slides were immediately frozen at -80°C for 5–10 min before they were fixed in 4% formaldehyde in 2XSSC (saline sodium citrate) for 10 min. Slides were washed twice (5 min each) in 2XSSC followed by 95% ethanol fixation for another 10 min. Slides were air-dried, and those with good metaphase spreads were selected for performing FISH.

In situ hybridization on *N. jonquilla* metaphase chromosome spreads

A slightly modified version of the FISH protocol by Fajkus *et al.* [26] was followed. First, 30 µl/slide of hybridization mix, which included 50% (v/v) deionized formamide (Applied Biosystems, USA), 10% (w/v) dextran sulphate in 2×SSC, 200–300 ng



Figure 1. Workflow: NGS–SERF-based approach. Putative telomere repeat motif analysis by SERF using the genomic NGS data as raw reads (FASTQ) or contig assembly (FASTA) followed by cytological validation with FISH. FISH: Fluorescence *in situ* hybridization; NGS: Next-generation sequencing; SERF: Sequence repeat finder.

of the biotinylated probe, 50× salmon sperm carrier DNA (stock 10 mg/ml; 5 Prime → 3 Prime, Inc., CO, USA) was prepared. The chromosomal DNA on the slides, together with the probe in the hybridization mixture, was denatured in a thermocycler for 3 min at 80°C. The slides were promptly moved to a moist chamber kept at 35°C in a hybridization chamber and hybridized overnight for 12–16 h. The cover slip was carefully removed by dipping in 2×SSC at 42°C. Post-hybridization washes were comprised of 3 × 5 min each at 42°C in 40% formamide in 2×SSC, pH 7; 1 × 5 min in 2×SSC at 42°C; 1 × 10 min in 2×SSC at room temperature. Slides were incubated with streptavidin 680-I dye conjugate (Abnova, Taiwan) at 1:200 dilution in 2×SSC for 40 min in a 37°C moist chamber and excess conjugate mix was washed off in 2×SSC at 45°C for 10 min. Finally, the slides were air-dried and counter-stained with ProLong® Diamond Antifade Mountant with DAPI (4',6-diamidino-2 phenylindole) by Life Technologies. Hybridization signals were imaged using a Leica TCS SP8 Confocal Microscope (Wetzlar, Germany).

Results & discussion

A SERF protocol of 2-5-15-90-TG-contiguous was used to validate the human-type telomere motif from the unassembled NGS reads and contig assembly of *I. uniflorum*. SERF analysis successfully identified its telomere motif despite the fact that only 1.4% (0.13 Gbp/9.44 Gbp) of the genome was available for analysis (Figure 2).

Table 1 shows the identification of the human-type telomere motif from the contig assembly of *I. uniflorum*. The analysis retrieved a single short contig of 59 base pairs in length consisting of 14 reads and harboring nine of the human-type (TTAGGG)_n telomere repeat motifs in tandem with a high quality metric of 91%. SERF analysis on the unassembled NGS reads of *I. uniflorum* has also detected an identical repeat motif output from multiple reads with high quality metric (data not shown).

SERF analysis detected and retrieved a single contig of 39 base pairs from *A. cepa*'s contig assembly of the pooled read sets with the known *Allium* telomere repeat motif (Table 2). The identified telomere motif from the assembled contig has a high quality metric of 92% and is rich in T and G as well. The telomere repeat motif in its various permutative forms was also located in many

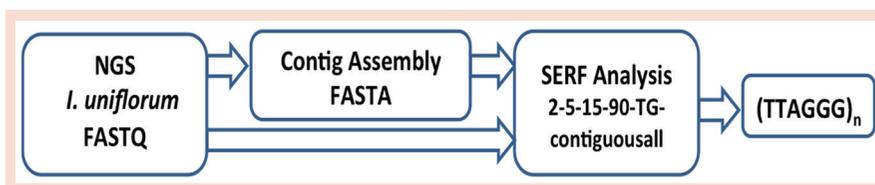


Figure 2. Telomere motif detection of *Ipheion uniflorum* with NGS–SERF-based approach. Detection of the human-type (TTAGGG)_n telomere repeat motif in *I. uniflorum* from its raw NGS reads and contig assembly using SERF.

NGS: Next-generation sequencing; SERF: Sequence repeat finder.

reads of single *A. cepa* NGS runs using the same protocol (Table 3). Depending on the number of NGS reads that harbor these extended repeats, assembly may not always yield a contig with these repeat motifs, which makes analysis on unassembled NGS data more desirable. WGS can overcome this difficulty and can generate a sufficient number of reads to form contig representing the telomere, enabling the measurement of the telomere content in the genome [30,37].

Sample output of the SERF analysis on one of the raw NGS sampling data sets of *A. cepa* using the protocol 2-5-15-90-TG-contiguous for a selected quality metric of 90% (Figure S2), FISH on its metaphase chromosome spreads (Figure S3) and the FISH methods are included in the Supplementary Material.

An extensive study of telomere repeat motif identification was also undertaken on the unassembled NGS data from genomic sampling of over 20 plant species from diverse families such as *Amaryllidaceae*, *Hyacinthaceae*, *Liliaceae*, *Solanaceae*, *Smilacaceae* and *Velloziaceae* (Table 4). Many of the plants analyzed have previously identified or probable consensus telomere repeat motifs of their respective families or genera. SERF analysis on *F. imperialis* predominantly detected the presence of *Arabidopsis*-type telomere repeat motifs, but combination repeats such as TTTAGGGTTAGGG and TTTAGGGTTAGGGG were also detected, but with a lower quality metric (44 and 54%). SERF analysis also identified both human-type (TTAGGG)_n and *Bombyx*-type (TTAGG)_n putative telomere motifs from a single read in *F. pudica* (41 and 48%) and *Arabidopsis*-type and *Bombyx*-type as putative telomere motifs in *T. nana* (100 and 98%). Analysis of *Xerophyta humilis* from the family *Velloziaceae* identified (TTTAGGG)_n and (TTTTAGG)_n as probable telomere repeat motifs with a high-quality metric (100 and 84%). The probable telomere repeat motifs (Table 4, superscript 5) identified by SERF analysis from various species require further

cytological confirmation. Most species that were analyzed had unassembled NGS data equivalent to 50–150 Mbp and were processed by SERF in a few hours.

In *N. jonquilla* and *M. comosum* from the order *Asparagales*, the *Arabidopsis*-type and human-type telomere repeat motifs were originally shown to be absent from their genome [38]. Later studies, however, did suggest the presence of these two types of telomere repeat motifs with an abundance of the human-type in *Narcissus triandrus* and *Muscari neglectum* telomeres [39]. SERF analysis on *N. jonquilla*'s raw NGS data sets identified a 15 bp unique conserved tandem repeat motif (GAGGGT-GCACTCACG) on a read with a high quality metric (90%) in addition to the human-type telomere repeat motif (TTAGGG) with 90% (Table 5). Higher probability of human-type motif as *N. jonquilla* telomere motif is evident from its abundance in the PCR amplification products (Figure S4). Analysis of the raw NGS data of *M. comosum* also revealed the abundance of human-type telomere repeat motif in multiple reads (Table 4). A few *Arabidopsis*-type repeats in tandem were also retrieved, but with a low quality metric.

The human-type and unique repeats identified by SERF analysis were subjected to cytological validation on *N. jonquilla* metaphase chromosome spreads using FISH. FISH with probes synthesized using the repeat motif (GAGGGT-GCACTCACG)_n did not yield any telomere-specific signals (data not shown), whereas FISH performed using probes specific to the human-type telomere repeat motif (TTAGGG)_n generated clear signals from the telomeres on the metaphase chromosome spreads of *N. jonquilla* (Figure 3).

Telomere repeat motifs identified by SERF from the analysis of plant genomes (Table 4) are deposited in the SERFDDB database (bioserf.org). Some motifs already have cytological validation from previous work, while others require such validation.

assistance was utilized in the production of this manuscript.

Open access

This work is licensed under the Creative Commons Attribution 4.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.future-science.com/doi/suppl/10.2144/btn-2018-0057

References

- Blackburn EH, Gall JG. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. *J. Mol. Biol.* 120(1), 33–53 (1978).
- Greider CW, Blackburn EH. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* 43(2), 405–413 (1985).
- Greider CW. Telomeres, telomerase and senescence. *Bioessays* 12(8), 363–369 (1990).
- de Lange T. Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev.* 19(18), 2100–2110 (2005).
- Richards EJ, Ausubel FM. Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* 53(1), 127–136 (1988).
- Frydrychova RC, Mason JM. Telomeres: their structure and maintenance. In: *The Mechanisms of DNA Replication*. David Stuart (Ed.). InTech Publishing, London, UK, 423–443 (2013).
- Okazaki S, Tsuchida K, Maekawa H, Ishikawa H, Fujiwara H. Identification of a pentanucleotide telomeric sequence, (TTAGG)_n, in the silkworm *Bombyx mori* and in other insects. *Mol. Cell. Biol.* 13(3), 1424–1432 (1993).
- Sahara K, Marec F, Traut W. TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosome Res.* 7(6), 449–460 (1999).
- Pardue M-L, DeBaryshe PG. Retrotransposons that maintain chromosome ends. *Proc. Natl Acad. Sci. USA* 108(51), 20317–20324 (2011).
- Bliessmann H, Mason JM, Ferry K *et al.* Addition of telomere-associated HeT DNA sequences ‘heals’ broken chromosome ends in *Drosophila*. *Cell* 61(4), 663–673 (1990).
- Levis RW, Ganesan R, Houtchens K, Tolar LA, Sheen F. Transposons in place of telomeric repeats at a *Drosophila* telomere. *Cell* 75(6), 1083–1093 (1993).
- Abad JP, de Pablos B, Osoegawa K *et al.* TAHRE, a novel telomeric retrotransposon from *Drosophila melanogaster*, reveals the origin of *Drosophila* telomeres. *Mol. Biol. Evol.* 21(9), 1620–1624 (2004).
- Procházková Schruppová P, Schořová Š, Fajkus J. Telomere- and telomerase-associated proteins and their functions in the plant cell. *Front. Plant Sci.* 7, 851 (2016).
- Petracek ME, Lefebvre PA, Silflow CD, Berman J. Chlamydomonas telomere sequences are A + T-rich but contain three consecutive G-C base pairs. *Proc. Natl Acad. Sci. USA* 87(21), 8222–8226 (1990).
- Fulnecková J, Sevcíková T, Fajkus J *et al.* A broad phylogenetic survey unveils the diversity and evolution of telomeres in eukaryotes. *Genome Biol. Evol.* 5(3), 468–483 (2013).
- Fulnečková J, Ševčíková T, Lukešová A, Sýkorová E. Transitions between the *Arabidopsis*-type and the human-type telomere sequence in green algae (clade *Caulodivuloxa*, *Chlamydomonadales*). *Chromosoma* 125(3), 437–451 (2016).
- Fajkus J, Sýkorová E, Leitch AR. Telomeres in evolution and evolution of telomeres. *Chromosome Res.* 13(5), 469–479 (2005).
- Sykorova E, Lim KY, Chase MW *et al.* The absence of *Arabidopsis*-type telomeres in *Cestrum* and closely related genera *Vestia* and *Sessea* (Solanaceae): first evidence from eudicots. *Plant J.* 34(3), 283–291 (2003).
- Peška V, Fajkus P, Fojtová M *et al.* Characterisation of an unusual telomere motif (TTTTT-TAGG)_n in the plant *Cestrum elegans* (Solanaceae), a species with a large genome. *Plant J.* 82(4), 644–654 (2015).
- Broun P, Ganai MW, Tanksley SD. Telomeric arrays display high levels of heritable polymorphism among closely related plant varieties. *Proc. Natl Acad. Sci. USA* 89(4), 1354–1357 (1992).
- Fajkus J, Novotná M, Ptacek J. Analysis of chromosome termini in potato varieties. *Rostl. Vyroba* 48(11), 477–479 (2002).
- Ganai MW, Lapitan NL, Tanksley SD. Macrostructure of the tomato telomeres. *Plant Cell* 3(1), 87–94 (1991).
- Adams SP, Leitch IJ, Bennett MD, Leitch AR. *Aloe L.* – a second plant family without (TTTAGGG)_n telomeres. *Chromosoma* 109(3), 201–205 (2000).
- Weiss H, Scherthan H. *Aloe* spp. – plants with vertebrate-like telomeric sequences. *Chromosome Res.* 10(2), 155–164 (2002).
- Sýkorová E, Fajkus J, Mezníková M *et al.* Minisatellite telomeres occur in the family *Alliaceae* but are lost in *Allium*. *Am. J. Bot.* 93(6), 814–823 (2006).
- Fajkus P, Peška V, Sitová Z *et al.* Allium telomeres unmasked: the unusual telomeric sequence (CTCGGTTATGGG)_n is synthesized by telomerase. *Plant J.* 85(3), 337–347 (2016).
- Pich U, Fuchs J, Schubert I. How do *Alliaceae* stabilize their chromosome ends in the absence of TTTAGGG sequences? *Chromosome Res.* 4(3), 207–213 (1996).
- Fuchs J, Brandes A, Schubert I. Telomere sequence localization and karyotype evolution in higher plants. *Plant Syst. Evol.* 196(3–4), 227–241 (1995).
- Sýkorová E, Lim KY, Kunická Z *et al.* Telomere variability in the monocotyledonous plant order Asparagales. *Proc. R. Soc. Lond., B, Biol. Sci.* 270(1527), 1893–1904 (2003).
- Lee M, Napier CE, Yang SF, Arthur JW, Reddel RR, Pickett HA. Comparative analysis of whole genome sequencing-based telomere length measurement techniques. *Methods* 114, 4–15 (2017).
- Girgis HZ. Red: an intelligent, rapid, accurate tool for detecting repeats *de-novo* on the genomic scale. *BMC Bioinformatics* 16, 227 (2015).
- Wirawan A, Kwok CK, Hsu LY, Koh TH. INVERTER: INtegrated Variable numbER Tandem rEpeat findeR. In: *Computational Systems-Biology and Bioinformatics*. Chan JH, Ong YS, Cho SB (Eds). Springer, Berlin, Heidelberg, Germany, 151–164 (2010).
- Novák P, Neumann P, Pech J, Steinhaisl J, Macas J. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* 29(6), 792–793 (2013).
- Peška V, Sitová Z, Fajkus P, Fajkus J. BAL31-NGS approach for identification of telomeres *de novo* in large genomes. *Methods* 114, 16–27 (2017).
- Somanathan I. *Methods and Analytics to Identify Telomere Repeats in Genomes*. M.S. Thesis. Biological Sciences Department, California State University East Bay, Hayward, CA, USA (2016).
- Mirzaghaderi G. Simple metaphase chromosome preparation from meristematic root tip cells of wheat for karyotyping or *in situ* hybridization. *Afr. J. Biotechnol.* 9(3), 314–318 (2010).
- Feuerbach L, Sieverling L, Deeg K *et al.* TelomereHunter: telomere content estimation and characterization from whole genome sequencing data. *bioRxiv* 065532 (2016).
- de La Herrán R, Cuñado N, Navajas-Pérez R *et al.* The controversial telomeres of lily plants. *Cytogenet. Genome Res.* 109(1–3), 144–147 (2005).
- Shibata F, Hizume M. Survey of *Arabidopsis*- and human-type telomere repeats in plants using fluorescence *in situ* hybridisation. *Cytologia* (Tokyo, Japan) 76(3), 353–360 (2011).
- Williamson JR, Raghuraman MK, Cech TR. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* 59(5), 871–880 (1989).
- Simonsson T. G-quadruplex DNA structures – variations on a theme. *Biol. Chem.* 382(4), 621–628 (2001).
- Adams SP, Hartman TPV, Lim KY *et al.* Loss and recovery of *Arabidopsis*-type telomere repeat sequences 5’-(TTTAGGG)_n-3’ in the evolution of a major radiation of flowering plants. *Proc. R. Soc. Lond., B, Biol. Sci.* 268(1476), 1541–1546 (2001).
- Fujimoto S, Ito M, Matsunaga S, Fukui K. An upper limit of the ratio of DNA volume to nuclear volume exists in plants. *Genes Genet. Syst.* 80(5), 345–350 (2005).

First draft submitted: 9 December 2018; Accepted for publication: 24 May 2018

Address correspondence to: Chris Baysdorfer; Department of Biological Sciences, California State University East Bay, Hayward, CA, USA; E-mail: chris.baysdorfer@csueastbay.edu

To purchase reprints of this article contact: s.cavana@future-science.com