Human-like telomeres in *Zostera marina* reveal a mode of transition from the plant to the human telomeric sequences

Vratislav Peska[1†*], Martin Mátl[2†], Terezie Mandákova[2,3], Daniel Vitales[4], Petr Fajkus[1,2], Jiří Fajkus[1,2,5], Sònia Garcia[4*]

[1]Department of Cell Biology and Radiobiology, Institute of Biophysics of the Czech Academy of Sciences, Brno CZ-61265, Czech Republic.

[2]Mendel Centre for Plant Genomics and Proteomics, CEITEC, Masaryk University, Brno CZ-62500, Czech Republic.

[3]Laboratory of Plant Molecular Genetics, Department of Experimental Biology, Faculty of Science, Masaryk University, Brno CZ-61137, Czech Republic.

[4]Institut Botànic de Barcelona (IBB, CSIC-Ajuntament de Barcelona), Passeig del Migdia s/n, Barcelona 08038, Catalonia, Spain.

[5]Laboratory of Functional Genomics and Proteomics, NCBR, Faculty of Science, Masaryk University, Brno CZ-61137, Czech Republic.

†Joint first authors.

*Joint corresponding authors: soniagarcia@ibb.csic.es, vpeska@ibp.cz.

Abstract

A previous study describing the genome of *Zostera marina*, the most widespread seagrass in the Northern hemisphere, revealed some genomic signatures of adaptation to the aquatic environment. Important features related to the 'back-to-the-sea' reverse evolutionary pathway were found, such as the loss of stomatal genes, while other functions like an algal-like cell wall composition were acquired. Beyond these, the genome structure and organization were comparable to the majority of plant genomes sequenced, except for one striking feature that went unnoticed at that time: the presence of human-like instead of the expected plant-type telomeric sequences. By using different experimental approaches including FISH, NGS and Bal31 analysis, we have confirmed its telomeric location in the chromosomes of *Z. marina*. We have also identified its telomerase RNA subunit (TR), confirming the presence of the human-type telomeric sequence in the template region. Remarkably, this region was found to be very variable even in clades with a highly conserved telomeric sequence across their species. Based on this observation, we propose that alternative annealing preferences in the template borders can explain the transition between the plant and human telomeric sequences. The further identification of paralogues of TR in several plant genomes brought us to the hypothesis that plants may keep an increased ability to change their telomeric sequence. We discuss the implications of this occurrence in the evolution of telomeres while introducing a mechanistic model for the transition from the plant to the human telomeric sequences.

Zostera marina, or common eelgrass, belongs to the family Zosteraceae, one of the four Alismatales families (basal monocots) that make up the seagrasses. *Zostera marina* plays a crucial role in coastal ecosystems around the world, providing food and shelter to numerous species. Although it is considered of least concern in terms of conservation status, the population trend is decreasing worldwide (IUCN red list), threatened by processes such as fishing, pollution, and the presence of invasive species[1,2]. Recently, the genome of this species was reported[3], a resource that has already aided in functional ecological studies in seagrass ecosystems under global warming[4,5] or evolutionary research[6,7].

Plants with unusual telomeres

A fundamental trait of a species genome is the composition of its telomeres. Telomeres are terminal chromosomal domains, whose minisatellite sequence is well conserved across large groups of organisms. Basically, we can refer to the plant-type telomere repeat as (TTTAGGG)n, shared by most plants, and to the human-type telomere repeat as (TTAGGG)n, shared by vertebrates and some other animals, while more diverse repeats can be found in fungi, insects, and others. Such typical telomeric sequences are mostly taken for granted in any species from a given group, yet recent experience prevents us from making such generalizations. For the discovery of unusual telomeric sequences in *Cestrum* and *Allium* (well-known genera for their ornamental and crop plants, respectively), which had remained enigmatic for decades, the availability of NGS data and bioinformatics tools have been crucial, together with molecular cytogenetics[8]. The finding of unusual telomeric sequences have had a great scientific impact, for instance, the exceptionally long *Allium* sequence motif[9] was further utilized as bait for the identification of *bona fide* telomerase RNAs (TRs) in plants[10].

Zostera marina has human-type telomeres

Here we provide compelling evidence that the common eelgrass also harbours unusual telomeres, composed exclusively of the human-type telomere DNA sequence (Figure 1). We have done so by (1) analysis of repeats from NGS data, (2) molecular cytogenetics with telomeric probes using fluorescence *in situ* hybridization (FISH) and (3) identification of its telomerase RNA, including its template region for synthesis of the telomeric motifs. More details on the materials and methods used can be found in the Reporting Summary.

The first indication that *Z. marina* could have an uncommon telomeric sequence came from the analysis of genomic data from the Finnish population, reported by Olsen et al. (2016), since information about telomere composition was not present in their genomic assembly. The subsequent screening we performed on this dataset revealed the absence of the expected plant-type and the presence of human-type telomere repeats. Although the quality of raw sequence data was excellent, and extreme care was taken to avoid any possible contamination from small marine animals (J. L. Olsen, personal communication), we provide independent evidence of the presence of human-type telomere repeats, as needed to exclude hypothetical contamination and to confirm this exceptional finding.

Thus, in parallel, we sequenced our own materials from Mediterranean populations of *Z. marina* and ran the same genomic screening. We also included the sequencing of a related species, *Z. noltii,* which shares the same habitat. While *Z. marina* genomic telomeric screening yielded exclusively human-type telomere repeats (neglecting reads representing less than 0.001% of the genome portion), *Z. noltii* showed plant-type telomere repeats, also almost exclusively. Based on the total number of reads analysed and the reads containing telomeric sequences, the genome portion of human-type telomeric DNA in *Z. marina* was 0.03-0.3%, while the genome portion of plant-type telomere DNA in *Z. noltii* was 0.06%. The genome proportion of *Z. marina* telomere DNA represented *c.* 3kb of telomere sequence per chromosome arm. This is in accordance with the average telomere length and genome size of *Arabidopsis thaliana*, which is within the known and small genome size range of *Zostera*. However, these are only semi-quantitative estimations, because of the NGS library type, which involves an amplification step biasing the quantitative determination. For more details, see Extended Data (File 1) in the online content. This study was complemented by the search for telomeric motifs in the NGS datasets available for Alismatales species in public databases. We have included species such as the well-known Neptune grass *Posidonia oceanica* (family Posidoneaceae) and other *Zostera* species. The outcome from the tandem repeat analysis, showing the most abundant tandem repeats in all these species, can also be found in the Extended Data (Table 1) in the online content. In all cases, except *Z. marina*, the search yielded the plant-type motif as the most likely telomeric sequence (0.01-7.31% of the genome portion); in comparison, the human-type sequence always represented a much lower proportion of the genome.

Although the analysis of NGS datasets clearly indicated that the *Z. marina* genome unexpectedly harboured these human-type telomeric repeats, cytogenetic confirmation to prove that these ones actually localize to chromosomal ends was needed. To our surprise, no

104 molecular cytogenetic studies of *Zostera* have been reported. Thus, we prepared
105 chromosomes from root tips and analysed them by FISH. The plant-type and human-type
106 probes were independently hybridized to metaphase plates of *Z. marina* and *Z. noltii* (both 2n
107 = 12). Hybridization of *Z. marina* chromosomes with the plant-type probe did not reveal any
108 signal, while, using the human-type probe, weak but clear signals were observed at most
109 chromosome ends. In *Z. noltii*, the opposite FISH profile was obtained: positive hybridization
110 with the plant-type probe, while negative with the human-type probe (Figure 1). As a control
111 for human-type telomere hybridization in plants, we used chromosomes of *Scilla peruviana*
112 (Asparagales), known to possess human-type telomeres[11] and whose chromosomes did label
113 with the human-type probe in the same experiment (data not shown). To further characterize
114 the chromosome structure in both *Zostera* species, rDNA-FISH was performed. In both species,
115 a single locus of terminal 35S and interstitial 5S rDNA was identified. However, the two species
116 differed in the position of the rDNA loci on the chromosomes: for *Z. marina*, 35S and 5S rDNA
117 were located on one arm of one chromosome while for *Z. noltii*, they were located on different
118 chromosomes (Figure 2).

119

120 **The telomere template in *Z. marina* telomerase RNA helps us understand telomere sequence**
121 **shifts in closely related species**

122 Finally, we utilized our recent finding of plant telomerase RNA in plants[10] to obtain
123 independent evidence of the newly found telomeric sequences in *Z. marina*, by analysing its
124 genomic and transcriptomic data. Telomerase RNA is an essential part of the telomerase
125 ribonucleoprotein complex, carrying the template region to produce telomeric DNA by reverse
126 transcription. The template region of the telomerase RNA is usually comprised of two parts.
127 One of them is a complete telomere motif - "*template sequence*" for reverse transcription (Fig.
128 1C and 1F), and the second is usually only a partial one, which serves as an annealing sequence
129 for the existing telomere DNA. We show here, for the first time, that the telomerase RNA
130 sequences of *Z. marina* and *Z. noltii* serve as human-type and plant-type templates,
131 respectively. The predicted human-type template region of *Z. marina* originated, most likely,
132 from the plant-type template by mutation at its borders. The template region of *Z. marina*
133 could, hypothetically, anneal and synthesize the plant telomere sequence, but the prediction
134 of human sequence annealing and synthesis fits better with the identified telomerase RNA
135 from *Z. marina* and, certainly, with the telomeric repeats detected in this species (see Figures 1
136 and 3). All other *Zostera* species and Alismatales analysed (from available online datasets)
137 represent exactly the opposite situation, with the plant-type motif in their predicted

138 templates. The reconstruction of the evolution of *Z. marina*'s template region may give us a
139 clue as to the transition from synthesis of the plant-type to the human-type telomere, driven
140 by a more stable mode of annealing of telomere DNA to the changed border of its template
141 region. Complete putative telomerase RNA sequences, in which the corresponding template
142 region is underlined, and other additional results, are presented in Figure 3 and in the
143 Extended Data (Figure 1 and Table 2) in the online content.

144 Over the last decade, unusual telomeric sequences have been reported for disparate plant
145 groups and in particular, the human-type sequence pops up repeatedly across the green plant
146 phylogeny (Figure 4). Although homoplasy in the evolution of telomere motifs is common, as
147 short/simple motifs like the plant- or the human-type have appeared repeatedly in distant
148 groups[11-14], it has been hypothesized that the ancestral telomere sequence is the human-type
149 since it is the most commonly found across the tree of life and it is present in the branches
150 close to the possible root of eukaryote phylogeny[15]. In the plant monocot order Asparagales,
151 the human variant occurs along with lower abundances of two or more variants of the
152 minisatellite sequences, including the consensus plant-type telomeric sequence , *Bombyx*-type
153 (TTAGG) and *Tetrahymena*-type (TTGGGG)[11], as well as the recently described *Allium*-type
154 telomere sequence (CTCGGTTATGGG)[9]. Genus *Genlisea* (eudicots) displays a similar profile: the
155 plant-type is found in some species while others reveal intermingled sequence variants
156 (TTCAGG and TTTCAGG)[16]. In certain groups of algae (Glaucophyta and Chlorophyta), human-
157 type telomeric sequences have already been detected in some species or genera, while others
158 also present the plant-type and other variants[17]. A similar situation applies to the genus
159 *Zostera* in which both human and plant-type telomeric sequences are found, adding another
160 exception to the increasingly questioned conservation of telomeric sequences throughout
161 plants (and perhaps throughout the tree of life).

162 What are the evolutionary implications of this finding? The cases of telomere shifts in closely
163 related species have huge evolutionary significance, as these not only affect the terminal
164 chromosomal DNA sequence but, if successfully maintained, must be associated with profound
165 changes in the protein components of telomeres, which recognize this telomere DNA in a
166 sequence-specific way. Our data show that the telomere system is more dynamic than
167 previously thought and that evolution tends to converge to a few successful types, since such
168 changes are "evolutionarily demanding". This may explain the several independent findings of
169 the human-type motif (one of the least complex telomeres) in some plants.

170 Moreover, a combination of the available genomic data from Alismatales with the recent
171 breakthrough of the discovery of plant telomerase RNA[10,18,19] have led us to the astonishing

conclusion that the template region in plant telomerase RNA is a very dynamic element, with unexpectedly variable borders, even in species with the same telomere sequence. The finding of two distinct telomerase RNA paralogs within several of the Alismatales genomes analysed, including the *Z. marina* genome assembly, offers an answer to how plants are able to change their telomeres. In the genome assembly of *Z. marina* we found two telomerase RNA gene candidates, however, only one with the human-like template region was detected in the transcriptomic analysis (Extended Data Table 2). Hence, some plants can use mutations in the template regions (particularly at its borders) of the telomerase RNA paralogs to escape from telomere sequence uniformity. Thus, our detailed analysis of the *Zostera marina* genome has revealed the first model for the transition from the plant to the human telomere sequence.

Online content Methods along with any additional Extended Data items are available in the online version of the paper; references unique to these sections appear only in the online paper.

Reporting summary Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability Raw NGS data files have been deposited in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI). The genomic skimming whole-genome shotgun sequences for *Zostera marina* and *Z. noltii* are accessible in bioproject PRJNA594842. The raw transcriptomic data for *Z. marina* are accessible in SRR10664371.

References

1 Borum, J., Monitoring & Seagrasses, M. o. E. *European Seagrasses: An Introduction to Monitoring and Management*. (M&MS Project, 2004).

2 Infantes, E., Crouzy, C. & Moksnes, P. O. Seed Predation by the Shore Crab Carcinus maenas: A Positive Feedback Preventing Eelgrass Recovery? *Plos One* **11**, doi:10.1371/Journal.Pone.0168128 (2016).

3 Olsen, J. L. *et al.* The genome of the seagrass Zostera marina reveals angiosperm adaptation to the sea. *Nature* **530**, 331-335, doi:10.1038/nature16548 (2016).

4 Lamb, J. B. *et al.* Seagrass ecosystems reduce exposure to bacterial pathogens of humans, fishes, and invertebrates. *Science* **355**, 731-733, doi:10.1126/science.aal1956 (2017).

5 Yin, R. H. & Ulm, R. How plants cope with UV-B: from perception to response. *Curr Opin Plant Biol* **37**, 42-48, doi:10.1016/j.pbi.2017.03.013 (2017).

6 Brawley, S. H. *et al.* Insights into the red algae and eukaryotic evolution from the genome of Porphyra umbilicalis (Bangiophyceae, Rhodophyta). *P Natl Acad Sci USA* **114**, E6361-E6370, doi:10.1073/pnas.1703088114 (2017).

7 Chater, C. C. *et al.* Origin and function of stomata in the moss Physcomitrella patens. *Nat Plants* **2**, doi:10.1038/Nplants.2016.179 (2016).

8 Peska, V. *et al.* Characterisation of an unusual telomere motif (TTTTTTAGGG)(n) in the plant Cestrum elegans (Solanaceae), a species with a large genome. *Plant J* **82**, 644-654, doi:10.1111/tpj.12839 (2015).

9 Fajkus, P. *et al.* Allium telomeres unmasked: the unusual telomeric sequence (CTCGGTTATGGG)(n) is synthesized by telomerase. *Plant J* **85**, 337-347, doi:10.1111/tpj.13115 (2016).

10 Fajkus, P. *et al.* Telomerase RNAs in land plants. *Nucleic Acids Res* **47**, 9842-9856, doi:10.1093/nar/gkz695 (2019).

11 Sykorova, E. *et al.* Telomere variability in the monocotyledonous plant order Asparagales. *P Roy Soc B-Biol Sci* **270**, 1893-1904, doi:10.1098/rspb.2003.2446 (2003).

12 Peska, V. & Garcia, S. Origin, Diversity and Evolution of Telomere Sequences in Plants. *Front Plant Sci*, doi:10.3389/fpls.2020.00117 (2020).

13 Puizina, J. *et al.* Karyotype analysis in Hyacinthella dalmatica (Hyacinthaceae) reveals vertebrate-type telomere repeats at the chromosome ends. *Genome* **46**, 1070-1076, doi:10.1139/g03-078 (2003).

14 Weiss, H. & Scherthan, H. Aloe spp.--plants with vertebrate-like telomeric sequences. *Chromosome res* **10**, 155-164, doi:10.1023/a:1014905319557 (2002).

15 Fulneckova, J. *et al.* A broad phylogenetic survey unveils the diversity and evolution of telomeres in eukaryotes. *Genome Biol Evol* **5**, 468-483, doi:10.1093/gbe/evt019 (2013).

16 Tran, T. D. *et al.* Centromere and telomere sequence alterations reflect the rapid genome evolution within the carnivorous plant genus Genlisea. *Plant J* **84**, 1087-1099, doi:10.1111/tpj.13058 (2015).

17 Fulneckova, J. *et al.* Dynamic Evolution of Telomeric Sequences in the Green Algal Order Chlamydomonadales. *Genome Biol Evol* **4**, 248-264, doi:10.1093/gbe/evs007 (2012).

18 Dew-Budd, K., Cheung, J., Palos, K., Forsythe, E. S. & Beilstein, M. A. Evolutionary and biochemical analyses reveal conservation of the Brassicaceae telomerase ribonucleoprotein complex. *bioRxiv*, 760785, doi:10.1101/760785 (2019).

19 Song, J. R. *et al.* The conserved structure of plant telomerase RNA provides the missing link for an evolutionary pathway from ciliates to humans. *P Natl Acad Sci USA* **116**, 24542-24550, doi:10.1073/pnas.1915312116 (2019).

Acknowledgements We are thankful to Jeanine Olsen who kindly discussed several aspects on this paper; Javier Romero and Francesca Rossi for their careful indications on where to find the studied *Zostera* and to Jordi Rull, Amelia Gómez and Joan Pere-Pascual Díaz who assisted in plant collection and DNA extraction. This work was supported by: ERDF [project SYMBIT, reg. no. CZ.02.1.01/0.0/0.0/15_003/0000477], Czech Science Foundation [17-09644S and 19-03442S], Ministry of Education, Youth and Sports of the Czech Republic [project CEITEC 2020 (LQ1601)], and by the Spanish [CGL2016-75694-P (AEI/FEDER, UE)] and Catalan [grant number 2017SGR1116] governments. VP benefited from an EMBO Short-Term fellowship (grant no. 7368) and SG is the holder of a Ramón y Cajal contract (RYC-2014-16608). We also acknowledge the CF Genomics CEITEC MU supported by the NCMG research infrastructure (LM2015091 funded by MEYS CR). Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum (https://metavo.metacentrum.cz/en/index.html) provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042), is greatly appreciated.

Author Contributions VP and SG initiated the project. VP, MM and SG performed the bioinformatic search of telomere motifs in SRA datasets. VP, SG and DV obtained and processed *Z. marina* and *Z. noltii* material for DNA extraction and chromosome preparation. TM performed the molecular cytogenetics experiments. PF and JF searched the TR subunit in the NGS datasets. SG and VP wrote the manuscript and all authors participated in the manuscript revision.

Competing interests The authors declare no competing interests.

METHODS

Plant material collection, DNA and RNA extraction. Young plants of *Zostera marina* and *Z. noltii* were collected in the wild (Location: Étang de Thau, beach next to the Musée de l'Étang de Thau, Sète, France. Collectors: Sònia Garcia, Amelia Gómez, Vratislav Peska, Jordi Rull & Daniel Vitales. Date: 20.02.2018). Vouchers are deposited at the Herbarium of the Institut Botànic de Barcelona with the numbers BC973540 (*Z. noltii*) and BC973541 (*Z. marina*). The leaves were carefully cleaned and inspected for the absence of other organisms. After that, they were wiped by paper tissue and frozen on dry ice. They were stored at -80 °C until usage. Genomic DNA was extracted by a CTAB method[20] and the quality checked with Qubit Fluorometric Quantification (ThermoFisher Scientific, 128 Waltham, Massachusetts, USA). Actively growing young roots were harvested from collected plants, and subsequently split into two aliquotes. The first aliquote was pre-treated with 0.05% aqueous colchicine for 2h 30 min at room temperature, fixed in ethanol/acetic acid (3:1, v/v) fixative for 24 h at 4 °C and stored at -20 °C until further use. The second aliquote was frozen on dry ice and stored at -80 °C until usage in the RNA extraction by TRI Reagent® (Sigma-Aldrich) according to the manufacturer's instructions. The extracted RNA was checked on an Agilent 2200 TapeStation system (Agilent Technologies) using RNA ScreenTape® (Agilent Technologies) and RNA concentration was determined by a Qubit 2.0 fluorometer using a Qubit™ RNA BR Assay Kit. Ribosomal RNAs were depleted from 5 µg of total RNA of the sample by Ribo-Zero rRNA Removal Kit (Seed, Root) (Illumina®). After rRNA depletion, samples were diluted in 10 µl of RNase-free water and 5 µl of the sample were used for RNA library construction using NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina® (NEB) according to the manufacturer's instructions. Library was treated by a fragment size selection in range from c. 200 to 500 bp using AGENCOURT®AMPURE®XP magnetic beads (Beckman Coulter). Library quality was checked by an Agilent 2200 TapeStation system using High Sensitivity D1000 ScreenTape® (Agilent Technologies).

Illumina sequencing. Extracted genomic DNA was sent to the NGS provider (BGI Genomics, Shenzen, China) and skimming genomic sequencing on Hiseq4000 was ordered for each sample. Approximately 15 million of 2 × 150 bp PE reads were obtained from each species. The cDNA Library (total RNA, rRNA depleted, and converted to cDNA) was sequenced in CEITEC MU Genomics Core Facility on a NextSeq500 platform (Illumina®) using NextSeq 500 v2.5., mid Output 150 cycles kits (Illumina®) for 2 × 75 bp PE reads. Raw reads are deposited in GenBank under the accession numbers SRR10664371 (RNAseq from *Z. marina*), SRR10664317 (genomic NGS data for *Z. marina*) and SRR10664318 (genomic NGS data for *Z. noltii*).

NGS data analysis. The telomere sequences from Alismatales were searched in the genomic skimming data obtained from BGI and GenBank by Tandem Repeats Finder (TRFi) tool (https://tandem.bu.edu/trf/trf.html): *Alocasia odora* (Araceae) - SRR7121940, *Halophila ovalis* (Hydrocharitaceae) - SRR5877255, *Lemna gibba* (Araceae) - SRR074103, *Lemna minor* (Araceae) - SRR2882980, *Posidonia oceanica* (Posidoniaceae) - SRR2315671, *Spirodela polyrhiza* (Araceae) - SRR7548932, *Wolffia australiana* (Araceae) - SRX3579183, *Valisneria spinulosa* (Hydrocharitaceae) - SRR6038670, *Zostera muelleri* (Zosteraceae) - SRR1714574, *Zostera marina* (Zosteraceae) - SRR3926352. *Zostera marina* (Zosteraceae) - SRR10664317, our dataset, *Zostera noltii* (Zosteraceae) - SRR10664318, our dataset. Subsequent analysis with custom made scripts were performed according to Peska et al. (2017)[21]. Briefly, the TRFi script was set to look for the short tandems with the unit size of 4-50 nt in at least pentamerous array. Motifs were sorted in descending order in each species. Human and plant telomere motifs were manually found in the final output. A summarized output of the most abundant tandem repeats, including telomere motifs, is presented in the Extended Data (Table 1). RNA-seq *de novo* assembly of data published in GenBank (SRR10664371) was done using Trinity-v2.7.0 (https://github.com/trinityrnaseq/trinityrnaseq/wiki). The assembly was performed as stranded RNA-seq with paired-end fastq datasets. Putative TR from *Z. marina* were found out using blast with published orthologs[10,18,19].

Telomerase RNA search analysis. Telomerase RNA (TR) orthologs were identified in publicly available datasets, going from TRs published from Alismatales using novel TR candidates in BLAST searches. The final set of TR presented in this work was aligned in Geneious 8.1.9 (https://www.geneious.com) using MAFFT alignment (Algorithm: E-INS-I; Scoring matrix: 200PAM/k = 2; gap open penalty: 1.53).

Chromosome preparation. Chromosome spreads from root tips were prepared according to a published protocol[22]. Briefly, selected root tips were rinsed in distilled water (twice for 5 min) and sodium citrate buffer (10 mM sodium citrate, pH 4.8; twice for 5 min), and digested in 0.3% (w/v) cellulase, cytohelicase and pectolyase (all Sigma-Aldrich, St Louis, MO, USA) in citrate buffer at 37 °C for 90 min. After digestion, individual root tips were dissected on a microscope slide in approximately 10 µl acetic acid and covered with a cover slip. The cell material was then spread evenly by tapping, thumb pressing and gentle flame-heating. Finally, the slide was quickly frozen in liquid nitrogen and the cover slip flicked off with a razor blade. The slide was fixed in ethanol-acetic acid (3:1) and air-dried. Ready-to-use chromosome spreads were checked under a phase contrast microscope for suitable mitotic chromosome figures and the amount of cytoplasm. When appropriate, preparations were treated with 100

342 µg/ml RNase (AppliChem) in 2× sodium saline citrate (SSC; 20× SSC: 3 M sodium chloride, 300
343 mM trisodium citrate, pH 7.0) for 60 min and 0.1 mg/ml pepsin (Sigma) in 0.01 M HCl at 37 °C
344 for 5 min, then post-fixed in 4% formaldehyde in 2× SSC for 10 min, washed in 2× SSC twice for
345 5 min, and dehydrated in an ethanol series (70%, 90%, and 100%, 2 min each).

346 **DNA probes preparation.** The plant-type (TTTAGGG)n and human-type (TTAGGG)n telomere
347 repeat probes were prepared according to protocol for non-template PCR[23]. The *Arabidopsis*
348 *thaliana* BAC clone T15P10 (AF167571) bearing 35S rRNA gene repeats was used for in situ
349 localization of nucleolar organizer regions (NORs), and the *A. thaliana* clone pCT4.2 (M65137),
350 corresponding to a 500 bp 5S rDNA repeat, was used for localization of 5S rDNA loci. All DNA
351 probes were labeled with biotin- or digoxigenin-dUTP by nick translation as described by
352 Mandáková and Lysak (2016b)[24]. Labeled probes were ethanol precipitated, desiccated and
353 dissolved in 20 µl of 50% formamide and 10% dextran-sulfate in 2× SSC for 3 h.

354 ***In situ* hybridization and microscopy.** For FISH, 20 µl of the labeled probe were pipetted on a
355 chromosome-containing spread and immediately denatured on a hot plate at 80 °C for 2 min.
356 Hybridization was carried out in a moist chamber at 37°C overnight, followed by post-
357 hybridization washing in 20% formamide in 2× SSC at 42 °C. The immunodetection of hapten-
358 labeled probes was performed as described by Mandáková and Lysak (2016b)[24]. Chromosomes
359 were counterstained with DAPI (2 µg/ml) in Vectashield. Fluorescence signals were analysed
360 and photographed using a Zeiss Axioimager Z2 epifluorescence microscope and a CoolCube
361 camera (MetaSystems). Images were acquired separately for all three fluorochromes using
362 appropriate excitation and emission filters (AHF Analysentechnik). The monochromatic images
363 were pseudocolored and merged using Photoshop CS software (Adobe Systems).

364

365

REFERENCES

20 Doyle, J. J. & Dickson, E. E. Preservation of Plant-Samples for DNA Restriction Endonuclease Analysis. *Taxon* **36**, 715-722, doi:Doi 10.2307/1221122 (1987).

21 Peska, V., Sitova, Z., Fajkus, P. & Fajkus, J. BAL31-NGS approach for identification of telomeres de novo in large genomes. *Methods* **114**, 16-27, doi:10.1016/j.ymeth.2016.08.017 (2017).

22 Mandakova, T. & Lysak, M. A. Chromosome Preparation for Cytogenetic Analyses in Arabidopsis. *Curr Protoc Plant Biol* **1**, 43-51, doi:10.1002/cppb.20009 (2016).

23 Ijdo, J. W., Wells, R. A., Baldini, A. & Reeders, S. T. Improved telomere detection using a telomere repeat probe (TTAGGG)n generated by PCR. *Nucleic Acids Res* **19**, 4780, doi:10.1093/nar/19.17.4780 (1991).

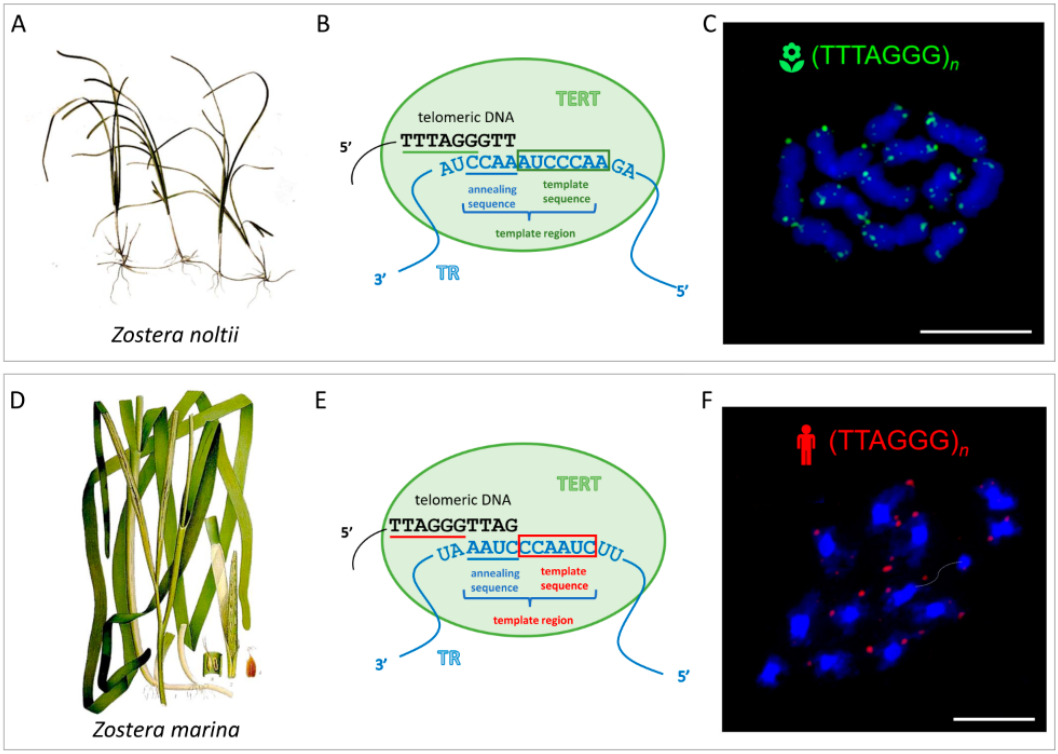
24 Mandakova, T. & Lysak, M. A. Painting of Arabidopsis Chromosomes with Chromosome-Specific BAC Clones. *Curr Protoc Plant Biol* **1**, 359-371, doi:10.1002/cppb.20022 (2016).

25 Byng, J. W. *et al.* An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc* **181**, 1-20, doi:10.1111/boj.12385 (2016).

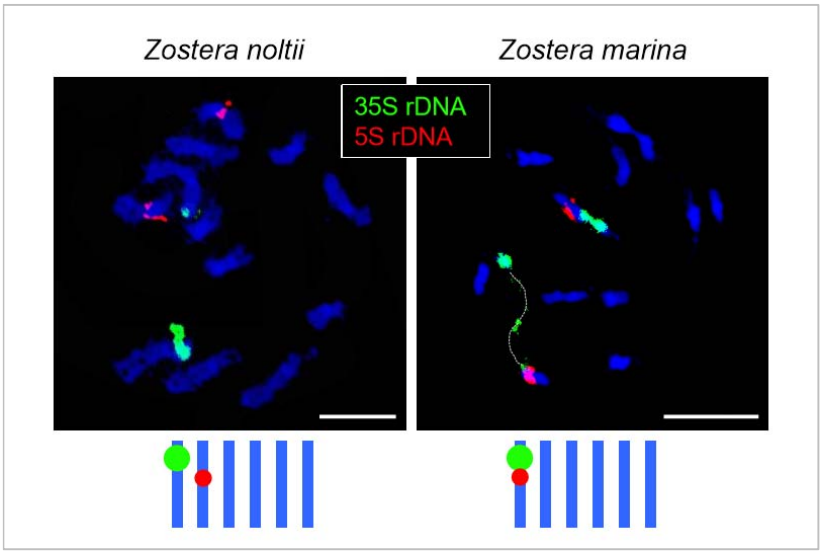
26 Leebens-Mack, J. H. *et al.* One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**, 679-+, doi:10.1038/s41586-019-1693-2 (2019).

27 Kumar, S., Stecher, G., Suleski, M. & Hedges, S. B. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Mol Biol Evol* **34**, 1812-1819, doi:10.1093/molbev/msx116 (2017).

Figure 1. Morphological appearances in *Zostera noltii* and *Z. marina*, respectively - images from Wikipedia - public domain (A, D), schematic model of telomerase RNAs (TR) and telomerase reverse transcriptase (TERT) indicating the annealing and template regions in the synthesis of telomeric DNA (B, E); fluorescence *in situ* hybridization of telomere probes on mitotic chromosomes and (C, F; scale bars = 10 µm).



403 **Figure 2.** Karyotypes of *Zostera noltii* and *Z. marina*. Both species bear a single terminally
404 located 35S rDNA locus and a single interstitial 5S rDNA locus. While 35S and 5S are situated
405 on different chromosomes in *Z. noltii*, they occupy the same chromosome in *Z. marina*. Note,
406 highly decondensed 35S rDNA site in *Z. marina* (dotted line). The plate of *Z. marina* is
407 incomplete, showing only eleven chromosomes. Scale bars = 10 µm.



408

409

Figure 3. Alignment of telomerase RNAs of several Alismatales species showing the putative template regions (underlined) in which the borders are variable even for plants with the same telomere sequence (e.g. *L. minor, P. oceanica*, and *Z. capensis*) or for paralogs from the same species (e.g., TR1 and TR2 from *Z. muelleri*) (A). The template region may expand, shrink or shift according to the sequence mutation at its borders and adjacent surroundings, as suggested in our model for human- and plant-type telomeric sequence synthesis in *Z. noltii* and *Z. marina*. A more stable mode of annealing of telomere DNA to the changed border of the template region drives the transition from plant- to human-type sequence synthesis in *Z. marina*. (B). The annealing part of template region is highlighted in bold font, the newly synthetized DNA is in red.

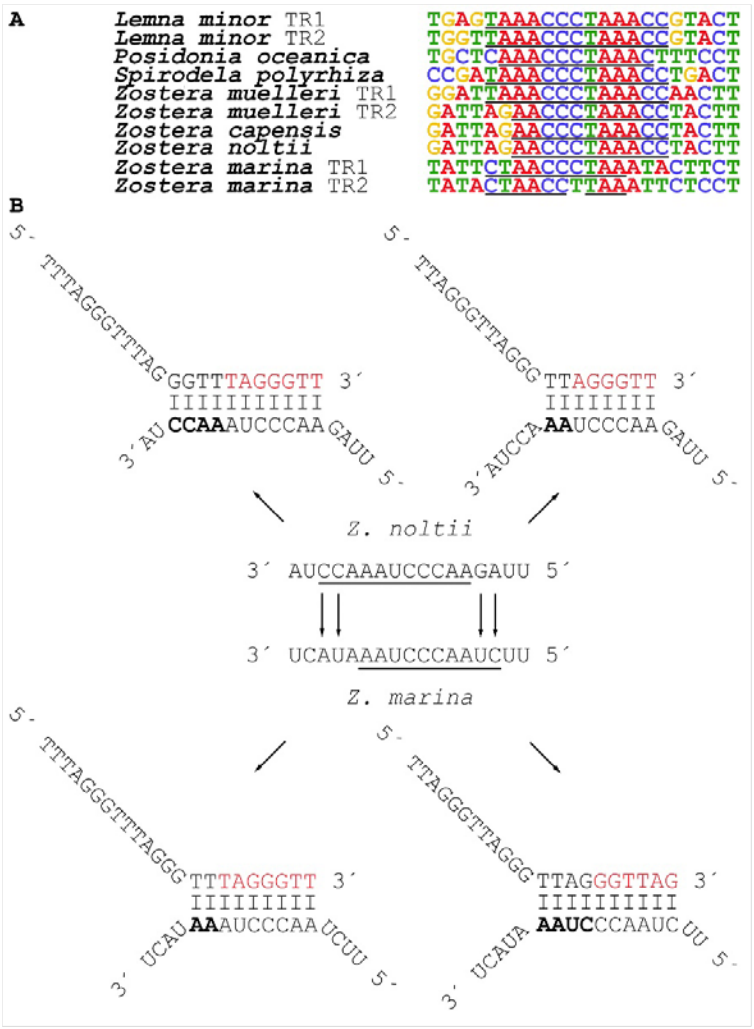
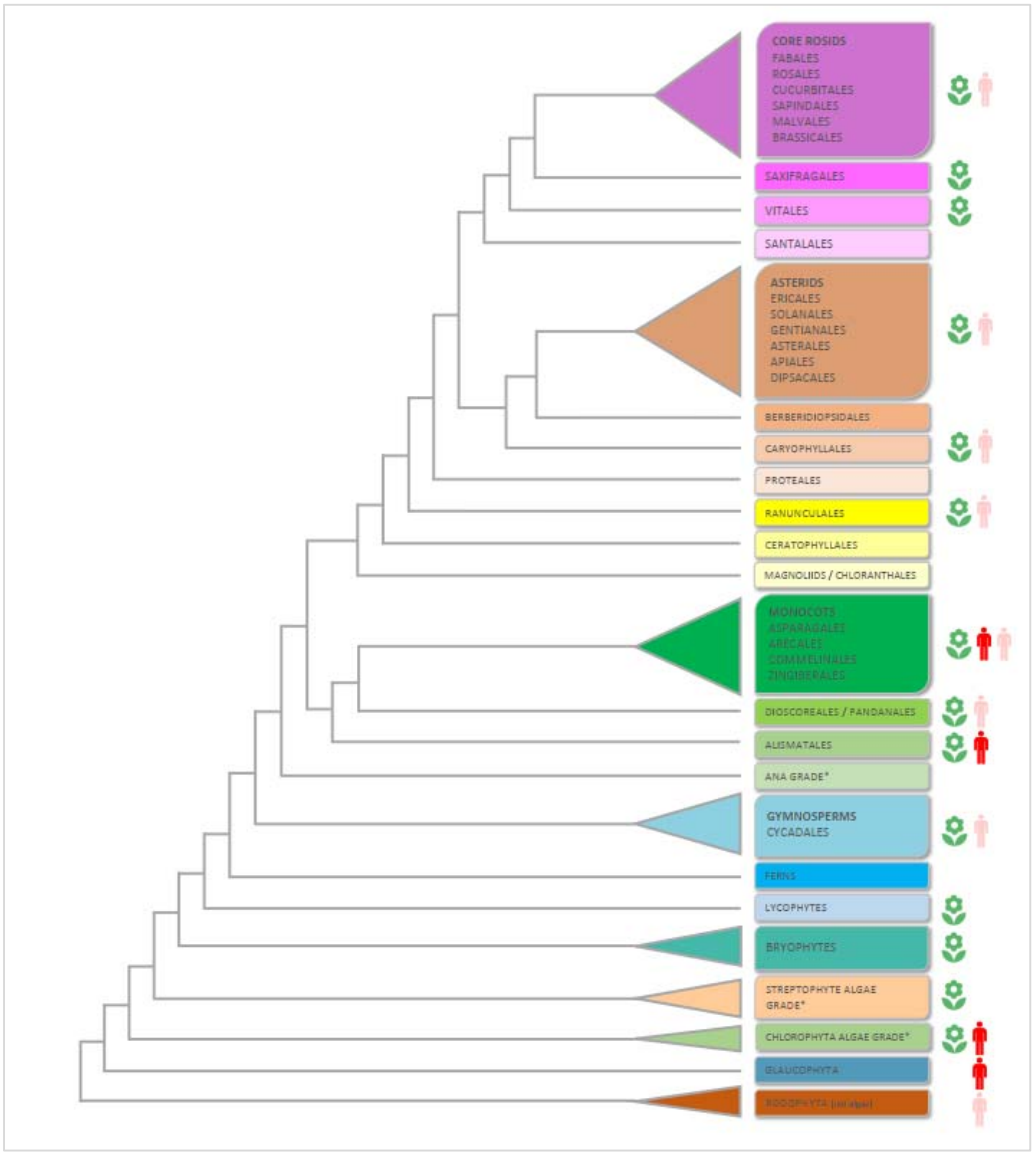
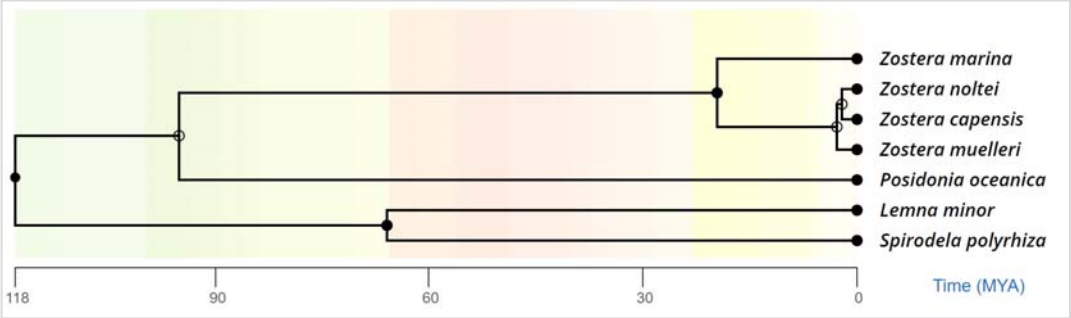


Figure 4. Occurrences of the plant-type and human-type telomere motifs in Archaeplastida (plants in the broad sense), based the APG IV (The Angiosperm Phylogeny Group 2016)[25] and on the One Thousand Plant Transcriptomes Initiative (2019)[26] Branch lengths do not express real time scales. For simplicity reasons, some groups have been collapsed or represented by a single branch. Some minor orders (Acorales, Petrosaviales, Trochodendrales, Buxales, Gunnerales, Dilleniales, Zygophyllales, Crossosomatales, Picramniales, Huerteales, Icacinales, Metteniusales, Garryales, Escalloniales, Bruniales and Paracryphilales) not depicted on the tree. The green flower represents the plant-type telomere sequence (TTTAGGG) and the red man the human-type telomere sequence (TTAGGG). The light red man represents cases in which there is some evidence (FISH, slot blot of from NGS data only) that would point to the presence of the human-type sequence, but further work is needed to confirm it.



436 **Extended Data Figure 1.** Tree showing phylogenetic relationships between selected
437 Alismatales based on TimeTree.org[27]. *Zostera marina*, which is the only one with human-like
438 telomeres, stands outside of the group formed by the other *Zostera* species analysed in this
439 work.



440

441

442

443 **Extended Data Table 1.** Report on the most abundant tandem repeats from TRFi analysis in
444 several Alismatales species. Plant-type telomere sequences indicated in green and human-type
445 telomere sequences indicated in blue.

446 (see independent file Extended Data Table 1.pdf)

Extended Data Table 2. Predicted telomerase RNA sequences of Alismatales species. The sequences of the predicted telomerase RNAs were obtained from or ABG – Applied Bioinformatics Group (http://appliedbioinformatics.com.au/index.php/Seagrass_Zmu_Genome), NCBI - National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/), and CoGe – Comparative Genomics (https://genomevolution.org/coge/).

TR1 and TR2 indicate two possible paralogs for telomerase RNA. TR1a and TR2 from *Z. marina* are genomic sequences. TR1a includes complete promoter and terminator non-transcribed sequences, while TR2 from *Z. marina* lacks several promoter elements. The TR1b from *Z. marina* was assembled for this study from reads obtained in RNA-Seq from root tips. Underlined nucleotides represent the predicted template regions. Note, the red dinucleotide of CT in *Z. marina* TR1b, which as an evolutionary novelty changes the template annealing potential to the preferred human-like telomere sequence. On the other hand, the potential for the plant telomere motif synthesis has not been completely lost in *Z. marina* template motif (green highlighting). In the non-transcribed TR2 of *Z. marina*, there is an additional mutation (C --> T) in the template region, which would potentially lead to (TTAGG)n (known from insects, blue highlighting) telomere sequence. From this point of view, the *Z. marina* TR represents transitional state producing human-type sequence, with a trace of plant-type synthesis.

***Lemna minor* (TR1) (Lemnaceae)** Sequence ID: lminor_contig_4989, Coordinates: 12166-12483, Source: CoGe, Genome ID:27408
AATCCCACATCGGAAAATCAGAGAAATCAAGATTTAATATATTTGCAAGTCTGCTAAAAATAGTAATTATTCCAGGGGAAGGGTG AGAGAGGTTAGAGAGAGAAACTGCTACTGAGTAAACCCTAAACCGTACTCTTAATTGAGGAATCTACCGGGCTTGATAGTGGGC TGTTTGTCCGGCGTTTGAGCTCCCGGGTTGAAAGGCCCAATGAAATGCCGATGCACGCGGCTTTCTCTCTCTAAACCATTGGAAG AGGCTTGTAGGGGGCAAACTGATTTGCCGATTTCCCTCGTCCCTCCCAAAATCCCCTGTTTTCTTTC
***Lemna minor* (TR2) (Lemnaceae)** Sequence ID: lminor_contig_4989, Coordinates: 10338-10654, Source: CoGe, Genome ID:27408
AATCCCACATCGAAAAATCAGAGTAATTGAGATTCATTATATTTCGATGCCTGGTTAAAAAAACAATAATTCCAGGGGAACGATG AGAGAGATTAGAGAGAGGAATTGCTACTGGTTAAACCCTAAACCGTACTCTTCATTGAGGATTCAGTTGGGCTTTGTTAACGGGC TGTTTATCCGGCGTTTGAGCTCCCGGGTTGAAAGGCCCAATGAAACGCCGATGCACGCGGGTTTCTCTCCCTAATCCTTGGGAAG AGGCTAGTAGGGGGCAAACTTGTTTGCCGATTTCCCTCGCCCTCCCAAATTCCCCTGTTTTCATAA
***Posidonia oceanica* (TR1) (Posidoniaceae)** Sequence ID: GFJT01022271.1 Coordinates: 1-297 Source: NCBI
TTCCATTTCCGCCCAAGCTGCTCACAGTATACGGGCGTCGGCATCCACCTTCCGGGGATGAGGTGGCTTGGGAGTACACTTGCAC CAAAGCATGCTTATGTGTGCTCAAACCCTAAACTTTCCTCTTAAGTGAGGTTCGAGTTAACCTTACCAAAGAGGTACTTTGTCCAG TATTTAACCTCTGTAATAAAAAAAATACTGATGCCTTGCACTTGTACTCCCAAAGTAGATATGCTACGGGAAGGCTTGAAGGGGG TTGCCTCGGCAACCGATATCCTCGCCTTCCAAGTCCCCATT

***Spirodela polyrhiza* (TR1) (Araceae)**

Sequence ID: UNPA01000014.1 Coordinates: 2999451-2999833 Source: NCBI

CCAAAATCCCACATCGGTAAATTTTTCAAGAAACATATGGGTTATATATTCAGCTGACCATTGGTTTTAATCTTCCAGGGGCCAAG
TGGGGATAGGAGGGCAGCCGCTCAGGCACCCGTTGACCGGGTTGTCCGATAAACCCTAAACCTGACTCTACTGAGGTCGACGCT
CCCAGCTTCACTTGGGTAGTGGAAAGAGGGCGGGGGCTGTTTACCCGGCGTTTAAGCCTCCCTTCTCCACCCGTCTTTGAGAAAT
TCAAACGCCGATGCCCGCGGCTGCCCCTCCCATCGATTATATCGTACCACGGGAGCGCTCGCAGGGGGCGGCGCCGCCAGGCTA
GCCGAGGTCCCTCGCCCTCCCAAGCCCCTGTTTTCTCTCTTCGA

***Zostera capensis* (TR1)**

Sequence ID: PRJNA503110 - assembled *de novo* from raw reads Source: NCBI

GTCCCACACAGCCGTTGAGATGAAAAATAGAAGCATATATAAAACTAAAAAGCAAGACATACAACACCTCCTTGGGGGAATTTCA
CAAGGAGATTATTGCAGCAATCCCATCAGCAACATCGTTGGGATTAGAACCCTAAACCTACTTCTCTGGAAGGTCTTTAACGTCCA
CTGTACTGAGAAAGGACTGAGGAGGTACTTTTTCCGGTGTTTAACCTTCTATAAACAATAAAAACATCGATGCCCCTGCCTCTGTT
CTCCTGTGGGTTGTTGTTGGGTAAGTCTTGAAGGGGTTGTTTATGAAAAAAACACCGATGTCCTCGACTCTTCCCAACCCCAAGTT
CCCCTTTTTTATTTCAA

***Zostera marina* (TR1a)**

Sequence ID: LFYR01000762.1 Coordinates: 70571-71209 Source: NCBI (genome assembly by Olsen et al. 2015)

GTCCCACACCGCCAAAGAAATAAATTTTAAAGCATATATATAATAAAAAAACCTATTTATAAAATAACTTATTGGAGGGTTTGTGT
TTCCACCAGTAGTTTATTGCAGTTTGTCTATTCTAACCCTAAATACTTCTTTCTTCTAGAAAAGAGTTAATGTATGCTAAAAAGTAG
GTATTTTCTCCGAGATTTCTACCTGAAATGAATATGAAAAATCTCGATGCTTCTGCCTCTGTACTTCTGTGATTGATTGTTGGGGAA
ATCTTTGTAGGGAGTATTGTTTGTTCATAACTGATATCCTCGGCTTTCCTACAACAATCAAAATTCCCTCTGTTTTCATTTAATTT

***Zostera marina* (TR1b)**

Sequence ID: SRR10664318 – assembled *de novo* from raw transcriptomic reads Source: This study

GTTTGTGTTTCCACCAGTAGTTTATTGCAGTTTGTCTATTCTAACCCTAAATACTTCTTTCTTCTAGAAAAGAGTTAATGTATGCTAA
AAAGTAGGTATTTTCTCCGAGATTTCTACCTGAAATGAATATGAAAAATCTCGATGCTTCTGCCTCTGTACTTCTGTGATTGATTGT
TGGGGAAATCTTTGTAGGGAGTATTGTTTGTTCATAACTGATATCCTCGGCTTTCCTACAACAATCAAAATTCCCTCA

***Zostera marina* (TR2)**

Sequence ID: LFYR01001054.1 Coordinates: 856312 - 856652 Source: NCBI (genome assembly by Olsen et al. 2015)

CATGATATACATCCTTACCGGAGGGTTTGTGTATTGATCATAAGTACGTAGTTTATTACAGTTTGTTTCCACCAGTCGGTATAGGT
GGTATACTAACCTTAAATTCTCCTTTCTTTTAGAAAAGAGTTGATTTATGTAAAAATCAGATAGTTTCTTTGAGATTTATGTCTGAA
ATAAATATGAAAAATCTCGATGCTTCTGCCTCTGTAATCGATCTTGTCATTGATTCTTGGGGAAAGTTTAATTTGTAAGGAGTGAT
GCTTGTTCATCACTGATATATATCCTCGGCTTTCCTACAACAATCAAATTCCCTCTGTTTTTGATGTCAGAATATACGATTA

***Zostera noltii* (TR1)**

Sequence ID: SRR2409708 – assembled from raw reads Source: NCBI

CTTGGGGGAATTTACAAGGAGATTATTGCAGCAATCCCATCAGCAACATCGTTGGGATTAGAACCCTAAACCTACTTCTCTGGAA
GGTCTTTAATGTCCANNTGTATGAGAGAGGACTGAGGAGGTACTTTTTCCGGTGTTTAACCTTCTATAAACAATAAAAACATCGA
TGCCCCTGCCTCTGTTCTCCTGTGGGTTGTTG

***Zostera muelleri* (TR1)**

Sequence ID: Zmu_v1_s5374_62084 Coordinates: 11965-12335 Source: ABG

AGTCCCACACAGCCGTTGGGAAGAAAAATAGAAGCATATATAAAACAAATAAGCAAAATTTACATCACTTTCTTGGGGGGAATTT
CACTAGGAGACCAATGCAGCTTTCCCACCAGCAACATCGGTGGGATTAAACCCTAAACCAACTTCTTCGGAAGGTCTTTGATCAT
GTCCATTGTACTGGAGAGGACTGAGGAGGTACTTTTACCGGTGTTTAACCTTCTATAAACAATAAAAACATCGATGCCCCTGCGTC
TGTGTTCTCCAGTGGGTTGTTGGTTGGGTTAAGTCTTGAAGGGGGTGTTTATGAAAAAACACCATTGTCCTCGACTCTTCCCAACC

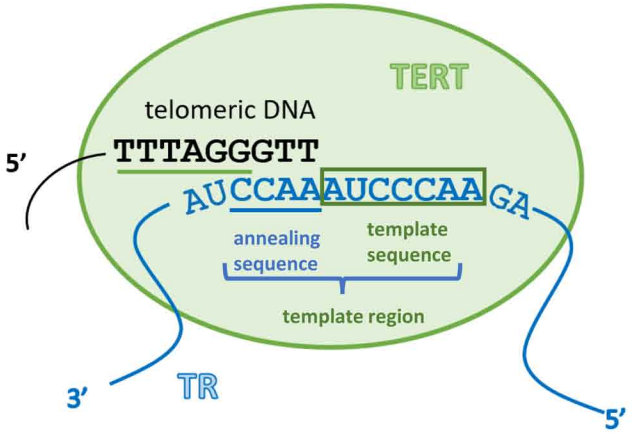
CCAAATTCCCCTCTTTTAATTTAATTTAA
***Zostera muelleri* (TR2)** Sequence ID: Zmu_v1_s7592_29784_9366_29784 Coordinates: 8581-8221 Source: ABG
GTCCCACACAGCCGTTGGGATAAAAAATAGAAGCATATATAAAACTCAAAAGCAAGACTACGACACCTTCTTGGGGGAATTTCAC AAGGAGATTATTGCAGCAATCCCATCAGCAACATCTTTGGGATTAGAACCCTAAACCTACTTCTTTGGAAGGTCTATAATGTCCAC CGTACTGAGAGAGGACTGAGGAGGTACTTTTTCCGGTGTTTAACCTTCTATAAACAATAAAAACATCGATGCTCCTGCCTCTGTTC TCCTGTGGGTTGTTGTTGGGTAAGTCTTGAAGGGGTTGTTTATGAAAAAACACCGATGTCCTCGACTCTTCCCAACCCCAAGTTCC CCTCTTTTAATTTAATTT

465

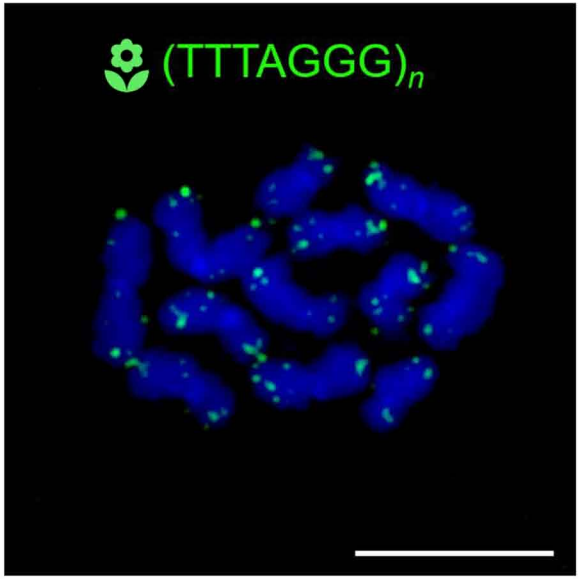
A

Zostera noltii

B



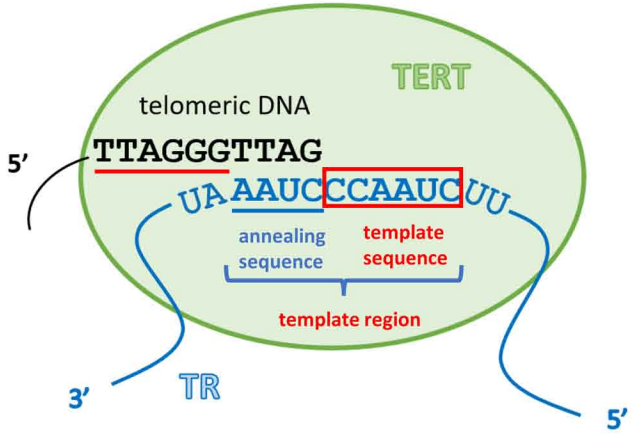
C



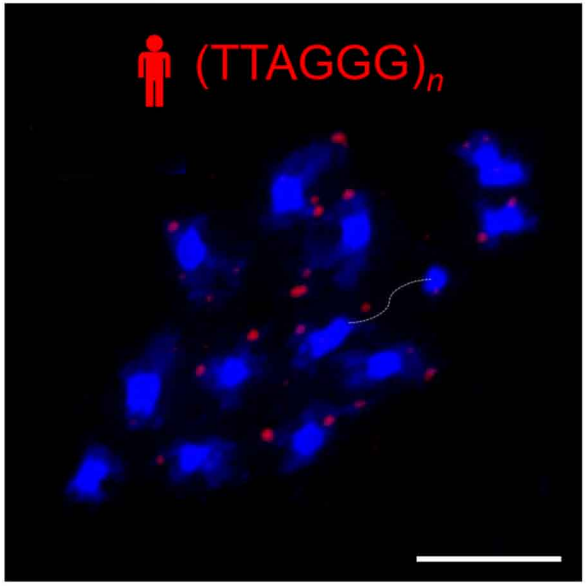
D

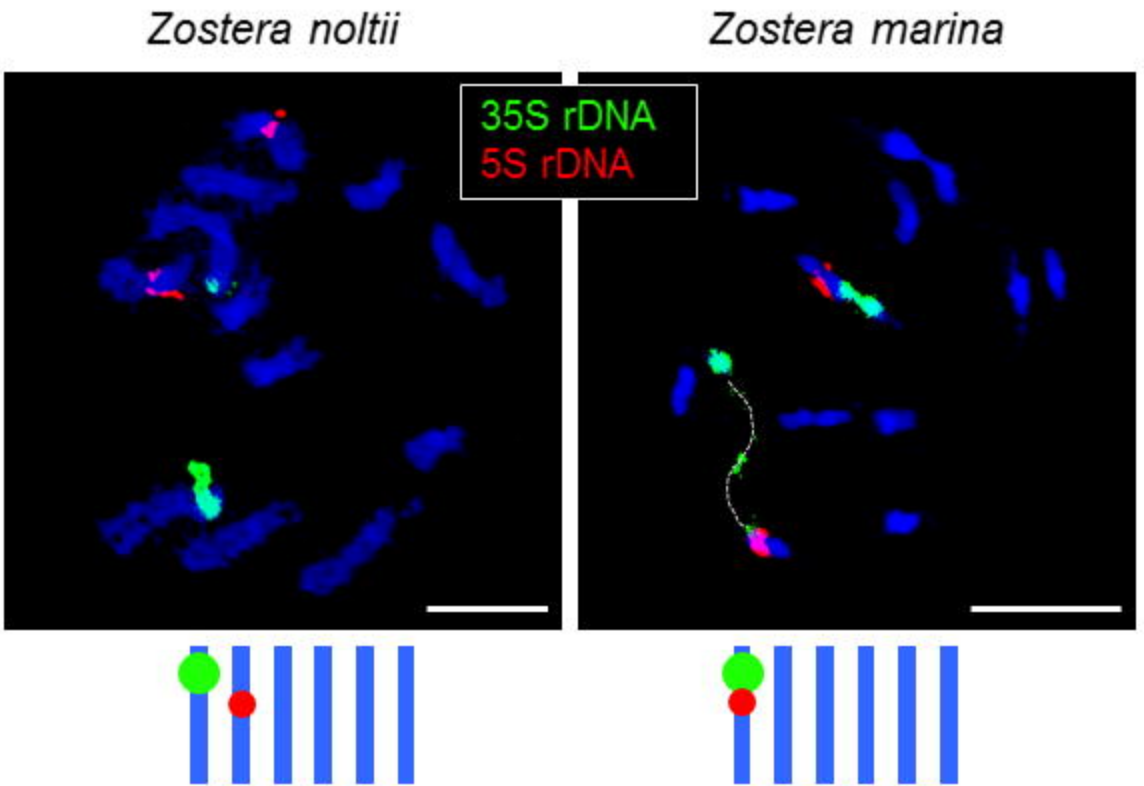
Zostera marina

E



F




Zostera noltii
Zostera marina
35S rDNA
5S rDNA