

POT1 proteins in green algae and land plants: DNA-binding properties and evidence of co-evolution with telomeric DNA

Eugene V. Shakirov, Xiangyu Song, Jessica A. Joseph and Dorothy E. Shippen*

Department of Biochemistry and Biophysics, Texas A&M University, 2128 TAMU, College Station, TX 77843-2128, USA

Received July 29, 2009; Revised August 27, 2009; Accepted September 5, 2009

ABSTRACT

Telomeric DNA terminates with a single-stranded 3' G-overhang that in vertebrates and fission yeast is bound by POT1 (Protection Of Telomeres). However, no *in vitro* telomeric DNA binding is associated with Arabidopsis POT1 paralogs. To further investigate POT1–DNA interaction in plants, we cloned *POT1* genes from 11 plant species representing major branches of plant kingdom. Telomeric DNA binding was associated with POT1 proteins from the green alga *Ostreococcus lucimarinus* and two flowering plants, maize and Asparagus. Site-directed mutagenesis revealed that several residues critical for telomeric DNA recognition in vertebrates are functionally conserved in plant POT1 proteins. However, the plant proteins varied in their minimal DNA-binding sites and nucleotide recognition properties. Green alga POT1 exhibited a strong preference for the canonical plant telomere repeat sequence TTTAGGG with no detectable binding to hexanucleotide telomere repeat TTAGGG found in vertebrates and some plants, including Asparagus. In contrast, POT1 proteins from maize and Asparagus bound TTAGG G repeats with only slightly reduced affinity relative to the TTTAGGG sequence. We conclude that the nucleic acid binding site in plant POT1 proteins is evolving rapidly, and that the recent acquisition of T TAGGG telomere repeats in Asparagus appears to have co-evolved with changes in POT1 DNA sequence recognition.

INTRODUCTION

Telomeres are the ancient nucleoprotein structures that define the physical ends of eukaryotic chromosomes, protecting them from deleterious activities, such as

recombination and nucleolytic attack, and providing a means to replenish telomeric DNA lost during replication (1). Defects in telomere structure or length maintenance result in cell proliferation and genome maintenance abnormalities, senescence or apoptosis (2). Telomere structure and composition is conserved across different eukaryotic lineages. Telomeric DNA typically consists of tandem arrays of short G-rich repeats that can reach thousands of nucleotides in length. The extreme 3'-ends of the chromosomes terminate in a single-stranded protrusion termed the G-overhang. Several evolutionarily conserved proteins bind directly to the double-stranded region of the telomeric DNA or to the single-strand G-overhang to form the first layer of telomere-associated protein factors. Together with bridging proteins, these DNA-binding factors constitute a telomere-specific protein complex termed shelterin (2,3).

The green plant lineage represents a monophyletic group of photosynthetic organisms that evolved near the base of eukaryotic life and shared the last common ancestor with fungi and animals ~1.5 billion years ago (bya) (4). Despite such long divergence time, many aspects of telomere biology are well conserved between plants and animals. The telomere repeat sequence in the vast majority of plants is TTTAGGG, one nucleotide longer than the 6-base sequence TTAGGG found in vertebrates (5). Interestingly, a few outliers exist in the plant kingdom, that harbor atypical or unknown telomere sequence. For example, while many green algae exhibit the canonical plant TTTAGGG repeats (6) (E. Shakirov and D. Shippen, unpublished results), telomeres in the model fresh water alga *Chlamydomonas reinhardtii* are composed of the eight nucleotide repeat TTTTAGGG (7). Onions (*Allium cepa*) and related species lack simple telomere repeats and instead appear to harbor terminally located satellite DNA (8). Most intriguing is the situation in two phylogenetically unrelated groups of flowering plants (angiosperms), in which canonical TTTAGGG repeats have been replaced by the vertebrate-type hexanucleotide TTAGGG sequence. One of these plant groups is

*To whom correspondence should be addressed. Tel: +1 979 862 2342; Fax: +1 979 845 9274; Email: dshippen@tamu.edu

comprised of a handful of obscure Solanaceae (tomato family) species native to South America (9). The second group consists of a large number of families in the Asparagales order, and represents one of the most successful lineages of extant flowering plants, with 22 000–25 000 currently known species, or nearly 10% of all angiosperms, including Irises, Hyacinths, Agaves and Amaryllis (10). The switch in telomeric DNA sequence from TTTAGGG to TTAGGG is thought to have occurred approximately 90 million years ago (mya) (11) and likely corresponds to a single nucleotide deletion in the template region of the telomerase RNA subunit. This sequence change may have posed a challenge for the plant shelterin complex to maintain chromosome end protection, although the successful diversification of Asparagales argues that these plants accommodated the mutation in telomeric DNA sequence in a short evolutionary time frame. Such compensatory changes likely involved co-evolution of telomere-binding proteins to allow binding to the new telomere repeat sequence, but the molecular mechanisms of these changes are currently unknown.

In vertebrates and fission yeast, POT1 (Protection Of Telomeres) binds single-strand G-rich telomeric DNA with high affinity and plays a pivotal role in mediating telomere length regulation as well as chromosome end protection and cell viability (2,3,12). POT1 proteins are defined by the presence of two N-terminal oligosaccharide/oligonucleotide binding folds (OB-folds), which are responsible for the specific interaction with single-stranded telomeric DNA. OB1 and OB2 contacts with telomeric DNA are primarily mediated by aromatic amino acids, which form stacking interactions with DNA nucleotides (13). Most POT1 proteins studied to date exhibit a minimum binding site (MBS) of 10–12 nucleotides, roughly corresponding to two telomeric repeats, though the most preferred repeat permutation in each case appears to be species-specific (14–16). Interestingly, only a subset of MBS nucleotides is specifically recognized by POT1 proteins, with nucleotides crucial for protein binding scattered throughout the MBS. For many POT1 proteins, the most 3'-terminal MBS nucleotide is buried deep inside the OB-folds (14,17), suggesting a mechanism for how POT1 can protect the G-overhang from nucleases or telomerase action. In addition, while some POT1 proteins clearly prefer 3'-terminal telomeric repeats (12), others can associate with telomeric sequences in the middle or on the 5'-terminus of oligonucleotide substrate (16,18), suggesting that *in vivo* they can also bind to the displaced G-rich strand in the context of the T-loop. While specific interaction with telomeric DNA *in vivo* is an essential feature of all POT1 proteins studied to date, human POT1 is delivered to the telomere via protein–protein interactions with its binding partner, TPP1/ACD. The interaction with TPP1 is achieved through a structurally undefined C-terminal domain (19,20). The POT1-TPP1 heterodimer is required for proper shelterin assembly and to regulate telomerase access and processivity (20–23).

POT1-like proteins have also been identified in plants (24–28). Similar to the situation in most other eukaryotes,

only a single *POT1* gene has been detected in most plants surveyed (Shakirov *et al.*, manuscript in preparation). However, Arabidopsis is an exception as it encodes three highly divergent POT1-like proteins (29; Nelson *et al.*, manuscript in preparation). All three Arabidopsis POT1 proteins are involved in telomere biology, but their functions differ. AtPOT1a is a positive regulator of telomerase activity that physically interacts with the telomerase RNP (30,31), while AtPOT1b and AtPOT1c negatively regulate telomerase activity and participate in chromosome end protection (29; Nelson *et al.*, manuscript in preparation). Strikingly, although AtPOT1 proteins have an architecture similar to yeast and vertebrate POT1 with two N-terminal OB-folds and a C-terminal domain, no *in vitro* telomeric DNA binding has been demonstrated for POT1 proteins from Arabidopsis or two other closely related plants (26,30). Thus, it is unclear whether telomeric DNA binding is a conserved function of POT1 proteins from the plant kingdom.

Plant systematics has witnessed a remarkable influx of new data revealing evolutionary relationship of various lineages of the green plants. We took advantage of this detailed phylogenetic map to clone *POT1* genes from eleven representative members of major plant evolutionary branches. Here we report the initial characterization of the DNA-binding activities of three plant POT1 proteins and provide evidence for significant biochemical differences in POT1 proteins across the plant kingdom. We also demonstrate that POT1 proteins from angiosperms have strong affinity for both TTAGGG and TTTAGGG telomeric repeats, providing a possible explanation for how Asparagales adapted to the recent switch in its telomeric DNA repeat sequence.

MATERIALS AND METHODS

Asparagus telomere length analysis

DNA from Asparagus (*Asparagus officinalis*) shoots was extracted as described (32). Terminal restriction fragment (TRF) analysis was performed as described (33) with DNA digested with either *Tru*I or *A*luI restriction enzymes (Fermentas, Hanover, MD). ³²P 5'-end-labeled (T₃AG₃)₄ and (T₂AG₃)₄TTAG oligonucleotides were used as heptanucleotide and hexanucleotide probes, respectively. Radioactive signals were scanned by a Pharos FX Plus Molecular Imager (Bio-Rad Laboratories, Hercules, CA), and the data were analyzed by Quantity One v.4.6.5 software (Bio-Rad).

In vitro translation and EMSA assays

Expression of plant POT1 proteins in rabbit reticulocyte lysate (RRL) was performed as described (24). EMSA assays were conducted as described (26) with slight modifications. Briefly, each reaction (15 µl total volume) contained equal amounts (4 µl) of RRL-translated plant POT1 protein, 0.5 pmol of the corresponding ³²P-labeled telomeric oligonucleotide, 3 µl of 5 × DNA-binding buffer (100 mM Tris–HCl pH 8.0, 250 mM NaCl, 50 mM MgCl₂, 5 mM EDTA, 5 mM DTT, 25% glycerol) and 1 µl each of non-specific RNA, single-stranded and double-stranded

DNA competitors as described in (24). Reactions were incubated at RT for 15 min. For competition assays, 2.5 pmol of cold competitor oligonucleotide was added and the reactions were incubated for an additional 15 min. The complexes were separated on 5% polyacrylamide gel (acrylamide: bisacrylamide 29:1) for 2 h at 150 volts in $0.8 \times$ TBE at RT, dried and exposed to PhosphorImager screens. Screens were scanned by a Pharos FX Plus Molecular Imager and signal intensity was quantified by Quantity One v.4.6.5 software. Each EMSA result was reproduced several times, but due to variations in protein expression levels in RRL, only one representative gel and the corresponding quantification result are shown for each experiment.

Site-directed mutagenesis

Mutagenesis reactions were performed with *Pfu* Turbo DNA polymerase (Stratagene) according to the manufacturer's instructions using the following PCR conditions: 94°C 5 min; 18 cycles of 94°C 30 s, 55°C 1 min, 68°C 40 min; followed by 10 min at 68°C. After *DpnI* (Promega) treatment at 37°C for 4 h, the reaction product was transformed into TOP10F' competent cells (Invitrogen). Plasmids were purified and mutations verified by sequencing.

Nucleotide sequence accession numbers

cDNAs encoding the following plant POT1 proteins were deposited into the GenBank: HvPOT1 (EU880295), PtrPOT1 (EU880297), HaPOT1 (EU880298), SmPOT1 (EU880301), ZmPOT1a (EU880303), ZmPOT1b (EU880304), GhPOT1 (EU880305), PtaPOT1 (EU880306), StPOT1 (EU883536) and AoPOT1 (FJ516399). The nucleotide sequence of cloned OIPOT1 cDNA was found to correspond to the previously deposited sequence ABO96101.

RESULTS

Telomeric DNA binding by plant POT1 proteins

As part of a larger study on the molecular evolution and functional divergence of POT1 proteins in plants, we cloned *POT1* cDNAs from eleven plant species representing major branches on the plant evolutionary tree. These include POT1 sequences from the green alga *Ostreococcus lucimarinus*, a spikemoss (*Selaginella moellendorffii*), a pine, three monocotyledonous (barley, maize and Asparagus) and four dicotyledonous (potato, sunflower, poplar and cotton) flowering plants (Supplementary Figure S1). Among the plant species analyzed in our study, three genomes have been sequenced: *O. lucimarinus*, *S. moellendorffii* and *Populus trichocarpa* (poplar). *Ostreococcus POT1* is a single-exon gene. *POT1* genes in both *Selaginella* and poplar, as well as the two previously characterized *Arabidopsis thaliana* *POT1* genes, harbor 10 exons with conserved intron positions (data not shown). This evolutionarily conserved gene structure supports the conclusion that the plant *POT1* genes are indeed orthologs. Among all plant

species analyzed in this study, only maize appears to encode more than one POT1 protein. The two maize *POT1* genes were most likely retained after whole-genome duplication in the ancestor of maize (34). *ZmPOT1a* and *ZmPOT1b* encode 54.5 and 54.6 kDa proteins, respectively, with 75% overall amino acid similarity to each other. Like the POT1 proteins from the Brassicaceae plants (26), all eleven of the new POT1 proteins display significant sequence conservation and are predicted to harbor two N-terminal DNA-binding OB-folds with secondary structures similar to the human and fission yeast POT1 proteins (J. Croy and D. Wuttke, personal communication).

To assess the DNA-binding properties of the new POT1 proteins, we attempted to obtain soluble recombinant proteins using standard expression protocols in *Escherichia coli*. However, as was previously shown for POT1 proteins from *Arabidopsis* and related species (26,30), we were unable to generate a sufficient amount of soluble protein from any clone for DNA-binding studies. Therefore, we turned to an *in vitro* rabbit reticulocyte lysate (RRL) expression system, previously shown to produce soluble vertebrate and plant POT1 proteins (24,26,35). While such approach will not allow us to define DNA-binding constants, we could perform qualitative binding experiments that in previous studies with the mammalian POT1 proteins have provided important comparative insights into POT1 interaction with telomeric DNA (24,35). RRL-expressed plant POT1 proteins were obtained in a soluble form (Supplementary Figure S2) and were subjected to electrophoretic mobility shift assays (EMSA) (Figure 1).

For the yeast and vertebrate POT1 proteins, two telomeric repeats are sufficient for *in vitro*-binding (13).

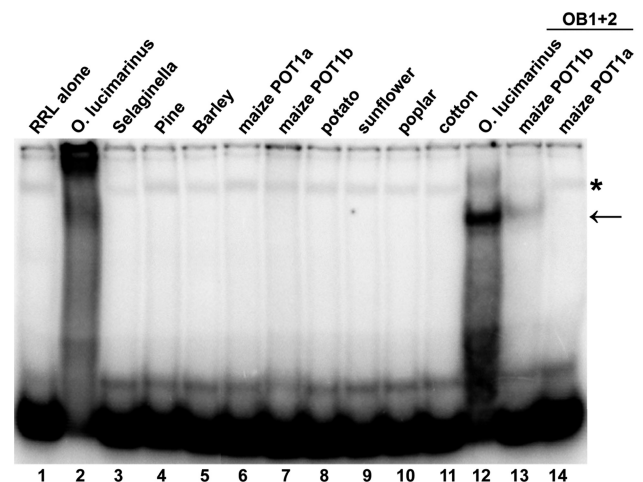


Figure 1. Analysis of DNA-binding capacity of recombinant plant POT1 proteins. EMSA was performed with a cocktail of seven 32 P-labeled 2PLANT oligonucleotides. POT1 proteins from the corresponding plant species are shown above each lane. Binding assays were performed with either full-length POT1 (lanes 2–11) or with truncated proteins bearing only two N-terminal OB-fold domains (lanes 12–14). Asterisk designates a non-specific band often present in the negative RRL-only control (lane 1). Protein–DNA complexes specific to maize and *Ostreococcus lucimarinus* POT1 proteins are indicated by an arrow.

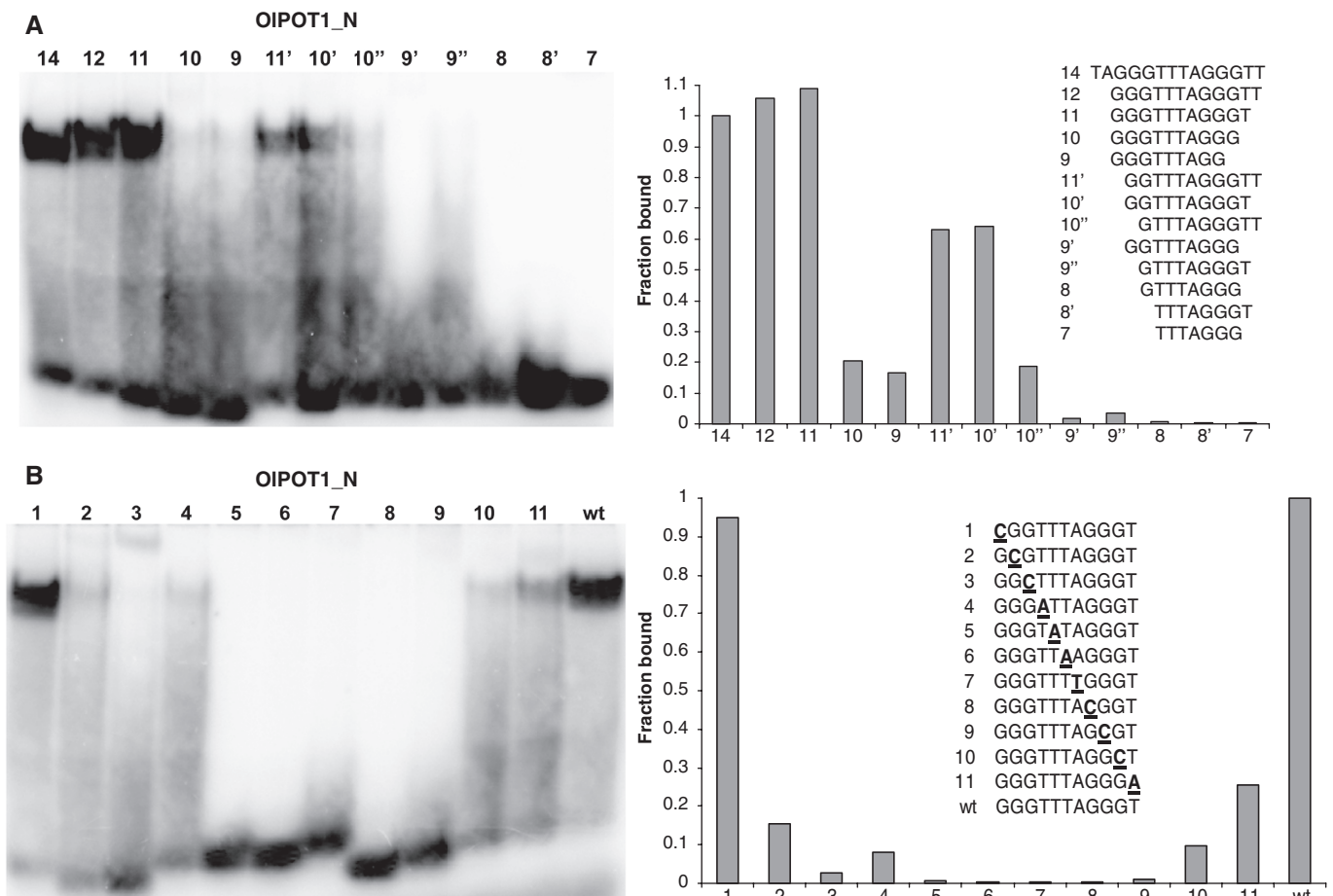


Figure 2. Characterization of DNA-binding activity of recombinant POT1_N protein from *Ostreococcus lucimarinus*. (A) EMSA identifies the minimum binding site of OIPOT1_N. Equal amounts of OIPOT1_N were incubated with the indicated oligonucleotides and the protein-DNA complexes were separated by native PAGE. (B) Identification of nucleotides recognized by OIPOT1_N in GGGTTTAGGGT. Numbers indicate nucleotide positions that were substituted with complementary nucleotides (bold and underlined). Representative EMSA scans are shown in left panels. For each scan, the signal intensity (fraction of protein bound) is plotted on the right with binding to TAGGGTTTAGGGTT (A) or GGGTTTAGGGT (wt) (B) set at 1.0.

Therefore, we used a cocktail of ^{32}P 5'-labeled oligonucleotides corresponding to two repeats of the seven possible permutations of the plant telomere repeat (2PLANT cocktail probe) for EMSA. Under standard gel-shift conditions (35), stable telomeric DNA binding was observed for two full-length plant POT1 proteins: OIPOT1 from the green alga *O. lucimarinus* (Figure 1, lane 2) and AoPOT1 from *A. officinalis* (garden asparagus) (see below). In addition, a band with intensity slightly above background was observed in the well for the maize (*Zea mays*) ZmPOT1b protein (Figure 1, lane 7). The intensity of this band increased when a specific oligonucleotide was used instead of the oligonucleotide cocktail (data not shown).

Telomeric DNA binding by mammalian and yeast POT1 proteins is enhanced with constructs containing only the N-terminal OB-folds (12,35,36), possibly due to modulation of DNA binding by the protein C-terminus. The OB-fold domains of all 11 plant POT1 proteins were expressed in RRL (Supplementary Figure S2 and data not shown) and tested in EMSA. Deletion of the C-terminus

improved the binding of OIPOT1_N (amino acids 1–363) to the plant telomeric DNA cocktail, resulting in the formation of a single well-defined protein-DNA complex (Figure 1, lane 12). In addition, a truncated version of maize POT1b, ZmPOT1b_N (amino acids 1–326) also formed single weak protein-DNA complex (Figure 1, lane 13). However, the second maize POT1 protein, ZmPOT1a, failed to bind telomeric DNA as either a full-length protein or an OB-fold truncation (Figure 1, lanes 6 and 14). With the exception of AoPOT1_N (see below), we detected no telomeric DNA binding by any other OB-fold truncated plant POT1 protein (data not shown). We also asked whether plant POT1 proteins could bind an oligonucleotide corresponding to the C-rich strand of telomeric DNA, as is the case for one of the POT1-like proteins from *Caenorhabditis elegans* (37). No C-strand binding was observed for any of the plant POT1 proteins (data not shown). Nevertheless, our data imply that the ability of POT1 to bind G-rich telomeric DNA has not been completely lost throughout plant kingdom.

DNA-binding properties of OIPOT1_N from green alga

Ostreococcus lucimarinus is a species of Prasinophytes, a clade of green algae that belongs to the oldest diverging (over 1 bya) branch of the photosynthetic eukaryotic lineage, and is a sister clade to all land plants (4). Consequently, analysis of the DNA-binding characteristics of OIPOT1_N may provide insight into the mechanisms of telomeric DNA recognition by the ancestral plant POT1 protein and how these properties have evolved in land plants.

To determine which of the seven permutations of the plant telomeric repeat OIPOT1_N binds best, we performed competition experiments with ³²P-labeled 2PLANT cocktail probe and a 5-fold excess of individual cold 2PLANT oligonucleotides. A representative gel and corresponding quantification are shown in Supplementary Figure S3A. The intensity of the shifted band (fraction bound) in the absence of competitors was measured and set as 1.0, and the remaining signal intensity after the addition of competitors was expressed as a fraction of 1. We found that all individual 2PLANT oligonucleotides competed efficiently for binding (Supplementary Figure S3A).

Next we determined the minimum DNA sequence required for OIPOT1_N binding. EMSA was performed using TAGGGTTAGGGTT and a series of single nucleotide truncations from either the 5'- or 3'-end of this substrate (Figure 2A). Deletion of the first two nucleotides from the 5'-end (oligonucleotide 12) did not decrease binding. However, removal of three nucleotides decreased binding by over 30% (oligonucleotide 11') and deletion of four abolished nearly all binding (oligonucleotide 10''). In both cases, a smear trailing down to free probe was observed, suggesting partial dissociation of the complex during electrophoresis. Only one nucleotide could be removed from the 3'-end without detectable loss of DNA binding (oligonucleotide 11). Deletion of two nucleotides (oligonucleotide 10) decreased binding to only 15% of the full-length oligonucleotide. Therefore, the minimum tight-binding sequence (core MBS) for OIPOT1_N appears to be GGTTTAGGGT (oligonucleotide 10'). Addition of one G residue at the oligonucleotide 5'-end improved binding almost 2-fold (oligonucleotide 11). Thus, as little as 10 nucleotides are necessary for OIPOT1_N binding and may comprise its MBS, while an 11-nt GGGTTTAGGGT oligonucleotide represents the best tight-binding substrate.

Studies with non-plant POT1 proteins reveal that only a subset of nucleotides within the MBS are specifically recognized (13). To identify nucleotides in the OIPOT1_N MBS critical for protein interaction, a series of oligonucleotides with single complementary nucleotide substitutions were tested for OIPOT1_N binding (Figure 2B). Unexpectedly, we discovered that all core MBS nucleotides are crucial for OIPOT1_N binding (Figure 2B, lanes 2–11). The only nucleotide change that did not affect OIPOT1_N binding was the most 5'-terminal G (Figure 2B, lane 1), which is not a part of the core MBS. A nucleotide in this position may be required for improved protein-DNA complex stability,

but may not contribute to specific interactions with OIPOT1_N protein. Overall, we conclude that OIPOT1_N requires at least 10 telomeric nucleotides for efficient binding. However, unlike the vertebrate and yeast POT1 proteins, all nucleotides in the core MBS appear to make specific and crucial contacts with OIPOT1_N.

DNA-binding properties of POT1b_N from maize

Maize is an angiosperm species that harbors canonical plant TTTAGGG telomere repeats (38). We tested ZmPOT1b_N binding to the seven permutations of the plant telomere repeat sequence using competition assays (Supplementary Figure S3B). Although incubation with several 2PLANT competitors leads to slightly decreased signal intensity (Supplementary Figure S3B, lanes 6–8), only the addition of (TTTAGGG)₂ and (GTTTAGG)₂ resulted in significant competition (Supplementary Figure S3B, lanes 2 and 3). These data indicate that ZmPOT1b_N recognizes two different permutations of the plant telomere repeat.

Next we determined the minimum number of nucleotides required for ZmPOT1b_N interaction with telomeric DNA (Figure 3A). In contrast to the situation with OIPOT1_N, removal of as many as seven nucleotides from the 5'-end of the oligonucleotide (a full plant TTTAGGG repeat) did not abolish binding (Figure 3A, oligonucleotide 7). On the other hand, only a single nucleotide could be removed from the 3'-end of (TTTAGGG)₂ oligonucleotide (Figure 3A, compare oligonucleotides 13' and 12'). Since two different probes containing 6 nucleotides each failed to bind ZmPOT1b_N (Figure 3A, lanes 6 and 6'), these data indicate that the MBS necessary for efficient ZmPOT1b_N binding consists of seven nucleotides, one full plant repeat.

To further evaluate the MBS of ZmPOT1b_N, we tested ZmPOT1b_N binding to oligonucleotides representing the seven permutations of the plant telomere repeat (1PLANT). Similar to the permutation analysis of 14-nt 2PLANT probes (Supplementary Figure S3B), ZmPOT1b_N stably associated with two 7-nt 1PLANT probes, TTTAGGG and GTTTAGG (Figure 3B, lanes 6 and 7). Among the two, TTTAGGG appeared to be a slightly better substrate, suggesting that this sequence represents the preferred MBS for ZmPOT1b_N.

We tested the relative importance of each MBS nucleotide for efficient ZmPOT1b_N binding using complementary nucleotide substitutions (Figure 3C). ZmPOT1b_N binding was significantly reduced or abolished with substitutions in positions 2 (TATAGGG), 5 (TTTACGG), 6 (TTTAGCG) and 7 (TTTAGGC) (Figure 3C, lanes 2, 5–7). Nucleotide changes in position 1 (ATTAGGG) and position 3 (TTAAGGG) did not lead to a substantial decrease in ZmPOT1b_N binding, while a change in position 4 (TTTGGG) actually improved binding 3-fold over the wild-type (Figure 3C, compare lane 4 with wt). These data suggest that unlike the situation with OIPOT1_N, not all nucleotides in ZmPOT1b_N MBS make crucial contributions to binding.

Interestingly, only three out of seven nucleotide positions (1, 4 and 5) differ in the two acceptable

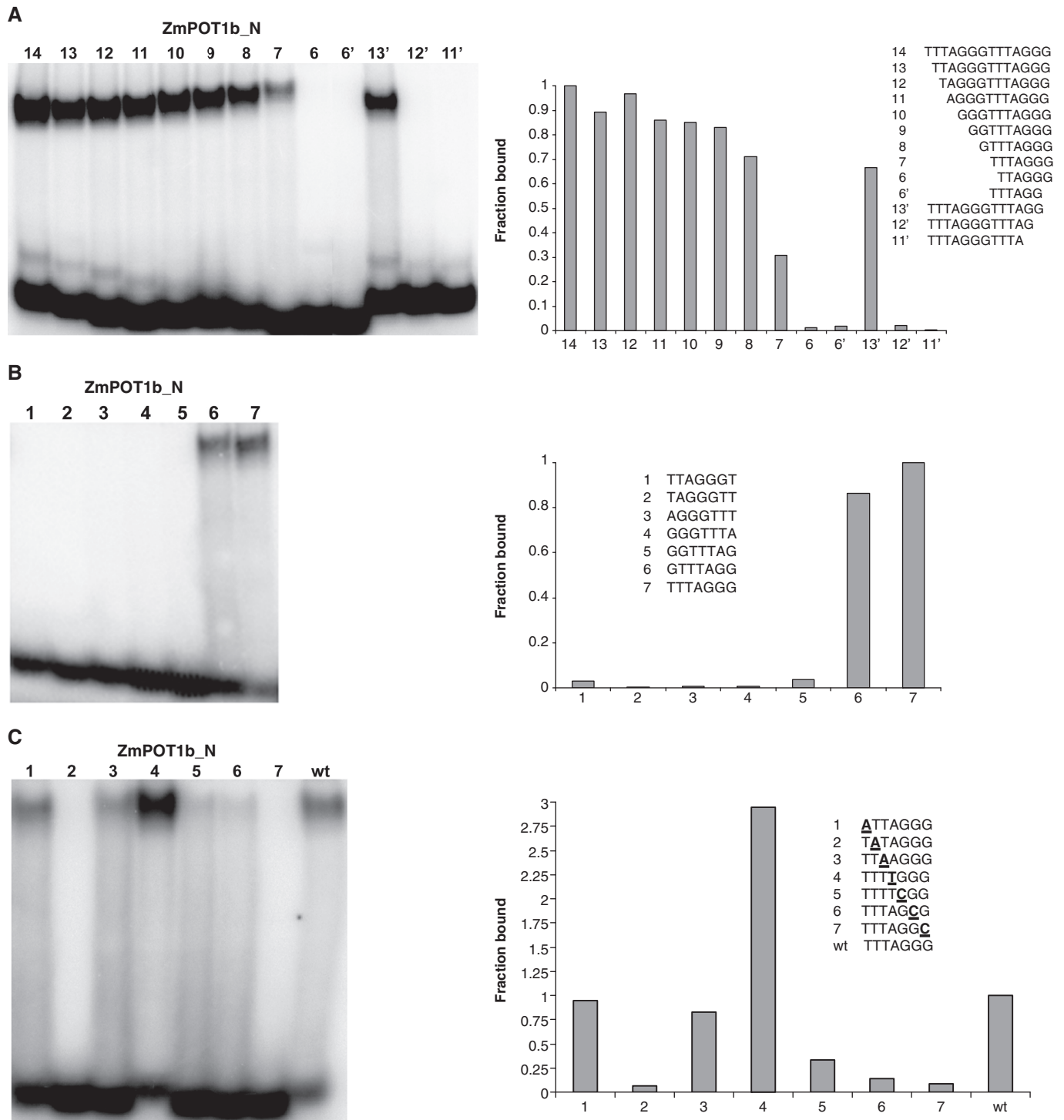


Figure 3. Characterization of ZmPOT1b_N interaction with telomeric DNA. (A) Identification of the MBS for ZmPOT1b_N. Equal amounts of ZmPOT1b_N were incubated with the indicated oligonucleotides and protein-DNA complexes were separated by native PAGE. (B) Analysis of ZmPOT1b_N binding to different permutations of a single plant telomere repeat. Numbers indicate seven possible repeat permutations. (C) Analysis of nucleotides specifically recognized by ZmPOT1b_N in TTTAGGG. Numbers indicate nucleotide positions that were substituted with complementary nucleotides (bold and underlined). For all panels, EMSA scans are shown on the left, and radioactive signal intensity is plotted on the right with ZmPOT1b_N binding to TTTAGGGTTTAGGG (A) or TTTAGGG (B, C) set at 1.0.

ZmPOT1b_N permutations, $\underline{T}_1T_2T_3\underline{A}_4G_5G_6G_7$ and $\underline{G}_1T_2T_3\underline{T}_4A_5G_6G_7$. Since changes in positions 1 and 4 do not lead to decreased protein binding, the increased ZmPOT1b_N binding to TTTAGGG versus GTTTAGG

may reflect a difference at position 5, with G being preferred over A. Overall, although TTTAGGG is the natural telomere repeat sequence in both *O. lucimarinus* and maize, the MBS and the relative importance of

individual nucleotides within it vary dramatically between OIPOT1_N and ZmPOT1b_N.

DNA-binding properties of Asparagus POT1 protein

The plant lineages leading to maize and Asparagus diverged only 100–110 mya (11), a relatively recent event in the evolutionary history of land plants. To study Asparagus POT1, we first confirmed that *A. officinalis* possesses hexanucleotide TTAGGG repeats using the TRF assay. As expected, a consensus plant telomere probe (T₃AG₃)₄ hybridized to the control Arabidopsis telomeric DNA (Supplementary Figure S4A, lanes 1 and 3), but not to Asparagus DNA (lanes 2 and 4), while a TTAGGG-specific telomere probe recognized Asparagus DNA, but not Arabidopsis (Supplemental Figure 4A, lanes 5–8). Strikingly, Asparagus telomere tracts appear to be at least an order of magnitude longer than in *Arabidopsis* (Supplemental Figure 4A, compare lanes 1 and 3 with 6 and 8).

To evaluate Asparagus POT1 binding to the hexanucleotide telomere repeat sequences, we examined the affinity of full-length AoPOT1 for an oligonucleotide cocktail containing equal amounts of all six possible permutations of two TTAGGG telomere repeats (2HEXA cocktail probe). A single shifted band was formed (Supplementary Figure 3C, lane 1). To determine which permutation(s) of the human telomere repeat were recognized, competition assays were performed with a 5-fold excess of individual cold 2HEXA oligonucleotides (Supplementary Figure S3C, lanes 2–7). Among all 2HEXA oligonucleotides, (TTAGGG)₂ was the most efficient competitor (Supplementary Figure S3C, lane 2). The specificity of AoPOT1 for (TTAGGG)₂ was confirmed in a direct binding assay. AoPOT1 efficiently bound (TTAGGG)₂, while the mouse mPOT1a_N protein was unable to bind this repeat permutation (Supplementary Figure S4B, compare lanes 1 and 3). Thus, although the telomere repeat sequence in Asparagus and vertebrates is the same, AoPOT1 prefers the permutation terminating in TTAGGG, while mPOT1a_N and other vertebrate POT1 proteins prefer the permutation terminating in GGTTAG (14,16,18,35).

AoPOT1 binding was assessed with a series of single-nucleotide truncations of the (TTAGGG)₂ oligonucleotide to define the MBS. Removal of even a single nucleotide from the 3'-end of the oligonucleotide was not tolerated (Figure 4A, oligonucleotides 11'-7'). On the other hand, removal of as many as 4 nucleotides from the 5'-end did not completely abolish binding, though it led to a substantial reduction in signal intensity (Figure 4A, oligonucleotide 8). While these data suggest that the MBS is the 8-nt GGTTAGGG, we noticed that AoPOT1 binding to a 9-nt oligonucleotide GGGTTAGGG is improved 2-fold (oligonucleotide 9), suggesting that GGGTTAGGG represents the best tight-binding substrate.

We next examined the relative contribution of each MBS nucleotide to AoPOT1 binding using a series of complementary nucleotide substitutions in GGGTTAGGG (Figure 4B). Although mutations in the first three 5'-terminal positions decreased signal intensity to only

34–45% of the original oligonucleotide (Figure 4B, oligonucleotides 1–3), mutations in all other positions abolished binding almost completely (Figure 4B, oligonucleotides 4–9). Thus, nucleotides 4–9 (a complete TTAGGG telomere repeat) make the most important contributions to AoPOT1 binding, while guanines 1–3 may be important for the stability of AoPOT1-DNA interaction. This situation is different from OIPOT1_N, where all the nucleotides in the MBS are important for binding, and ZmPOT1b_N, where several nucleotide positions in the MBS exhibit more relaxed specificity.

POT1 proteins from Asparagus and maize bind both hexa- and heptanucleotide telomere repeats

Our TRF and EMSA data indicate that AoPOT1 may bind TTAGGG repeats *in vivo*, however it was unclear whether this protein retained the ability to interact with the ancestral TTTAGGG sequence. We found that AoPOT1 binding to two heptanucleotide repeats is 2-fold better than to two hexanucleotide repeats (Figure 5, lanes 1 and 2), suggesting that the ancestral TTTAGGG sequence is still a preferred substrate for AoPOT1. We next tested whether the ability to bind both types of telomere repeats is a conserved feature of POT1 proteins. As with AoPOT1, ZmPOT1b_N bound both types of repeats (Figure 5, lanes 7 and 8), displaying 2.5-fold better binding to (TTTAGGG)₂ than to (TTAGGG)₂. In striking contrast, we could not detect binding by *O. lucimarinus* POT1_N to any permutation of the hexanucleotide repeat in a direct binding assay (Figure 5, lanes 4 and 10). Similarly, mouse POT1a_N failed to bind any permutation of the plant telomere repeat (Figure 5, lanes 6 and 11). Thus, recognition of both hexanucleotide and heptanucleotide telomere repeats is an evolutionarily conserved feature of POT1 proteins from maize and Asparagus, which dates back to at least 100–110 mya. The inability of OIPOT1_N to bind hexanucleotide repeats suggests that the algal POT1 protein either lost the ability to bind such repeats after the divergence of land plants and green algae, or this property evolved independently in angiosperms.

We noted that the first six 3'-terminal positions in the MBS of POT1 proteins from angiosperms (GGGTTAGGG for AoPOT1 and TTTAGGG for ZmPOT1b_N) are identical and constitute one full hexanucleotide repeat. The remaining 5'-terminal nucleotides in each MBS do not appear to make crucial contacts with the corresponding POT1 proteins (Figures 4B, lanes 1–3 and 3C, lane 1). Thus, the ability of AoPOT1 and ZmPOT1b_N to bind both hexa- and heptanucleotide telomere repeats can be explained if both proteins fail to discriminate between different nucleotides in the 5'-terminal positions of their respective MBS. In support of this model, we found that AoPOT1 and ZmPOT1b_N can bind to all variations of GGNTTAGGG and NNTTAGGG, respectively (Supplementary Figure S5).

End-binding specificity of plant POT1 proteins

Vertebrate and yeast POT1 proteins differ in their preference for telomeric repeats at the 3'-end of the DNA

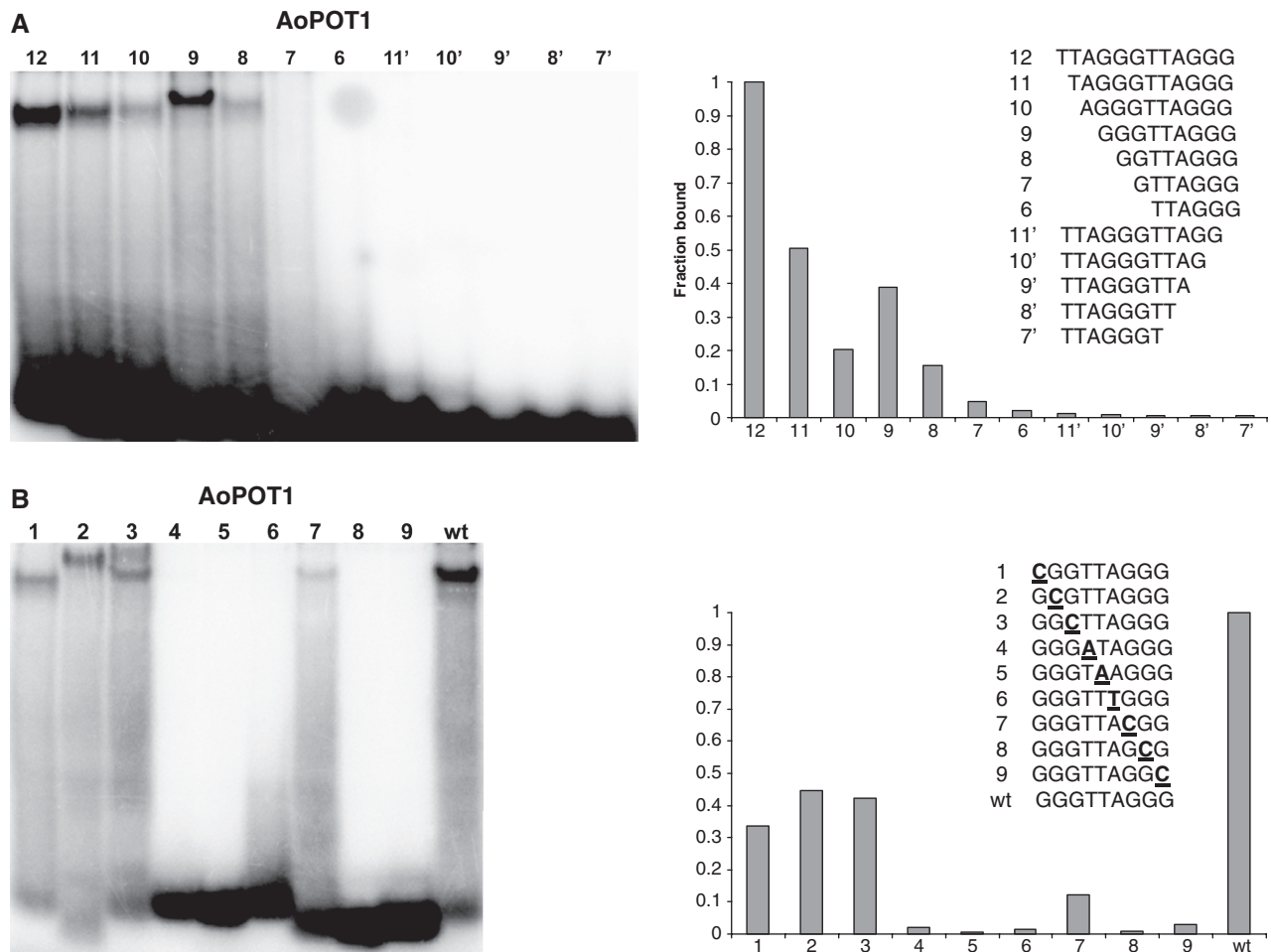


Figure 4. Analysis of Asparagus POT1 interaction with telomeric DNA. (A) Identification of the MBS for AoPOT1. Equal amounts of AoPOT1 were incubated with the indicated oligonucleotides and protein-DNA complexes were separated by native PAGE. (B) Analysis of nucleotides specifically recognized by AoPOT1 in GGGTTAGGG. Numbers indicate nucleotide positions that were substituted with complementary nucleotides (bold and underlined). For all panels, EMSA scans are shown on the left, and radioactive signal intensity is plotted on the right with AoPOT1 binding to TTAGGGTTAGGG (A) or GGGTTAGGG (B) set at 1.0.

substrate (13). Therefore, we asked if plant POT1 proteins also exhibit such preference by performing competition assays with oligonucleotides containing two telomere repeats at the 5'- or 3'-end or in the middle of a longer DNA oligonucleotide (Figure 6). Interestingly, the plant POT1 proteins behaved differently in these competition assays. Complex formation of OIPOT1_N with labeled (TTAGGGT)₂ was significantly reduced or abolished with all competitors, suggesting that this protein binds telomeric repeats regardless of their position in the substrate (Figure 6A). On the other hand, increasing competition was observed for ZmPOT1b_N with oligonucleotides carrying the telomere repeats on the 5'-end, middle and on the 3'-end (Figure 6B), respectively. Finally, AoPOT1 binding to (TTAGGG)₂ could only be competed with a substrate harboring two hexanucleotide telomere repeats on the 3'-end of the oligonucleotide (Figure 6C). We conclude that POT1 preference for the position of telomeric repeats on the DNA substrate is species-specific and not evolutionarily conserved.

Mutational analysis of plant POT1 proteins

Structurally characterized POT1 proteins from humans and *Schizosaccharomyces pombe* share a number of conserved primary sequence and secondary structure elements, which are crucial for specific interaction with telomeric DNA (14,39). The availability of Asparagus POT1 provides an opportunity to analyze the importance of these amino acids and protein regions for TTAGGG repeat recognition in the context of a full-length plant POT1 protein. As expected from studies in yeast and humans (12,35,36), the C-terminal region of AoPOT1 (amino acids 322–504) was dispensable for DNA binding (Figure 7A, lane 2). Moreover, full-length AoPOT1 and AoPOT1_N have similar binding properties, although the latter showed slightly reduced binding to the 8-nt GGTTA GGG substrate (Supplementary Figure S6). Further truncation of AoPOT1 to eliminate the second OB-fold (amino acids 1–167) completely abolished DNA-binding activity (Figure 7A, lane 3). Thus, two N-terminal OB-folds are sufficient for telomeric DNA binding by AoPOT1.

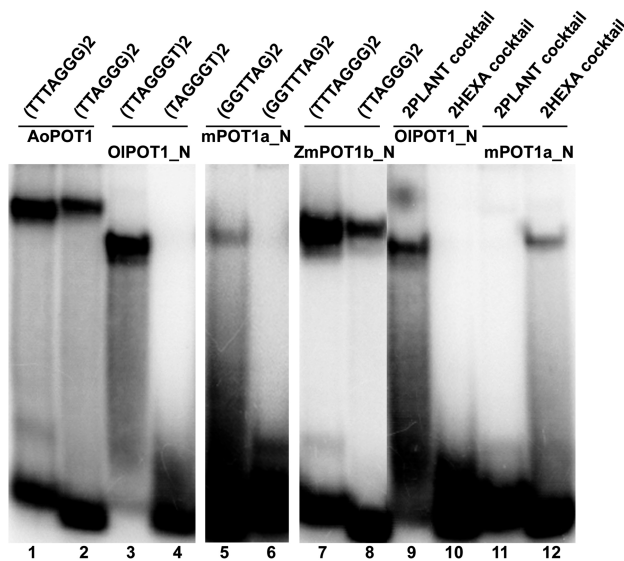


Figure 5. AoPOT1 and ZmPOT1b_N bind both hepta- and hexanucleotide telomere repeats. EMSA results are shown for AoPOT1, OIPOT1_N, mPOT1a_N and ZmPOT1b_N. POT1 proteins were incubated with the indicated oligonucleotides consisting of two full hepta- or hexanucleotide telomere repeats, or with an oligonucleotide cocktail containing all possible permutations of the heptanucleotide or hexanucleotide telomere repeat.

The crystal structure of human POT1 protein identified conserved F62 and Y89 residues in OB1, which are critical for POT1 interaction with telomeric DNA (Figure 7B) (14). Most plant POT1 proteins have a nearly invariant phenylalanine in the first position (F67 in AoPOT1) and a large amino acid with a bulky side chain (mostly Y, H or F) in the second position (Y94 in AoPOT1). Consistent with previous reports for mammalian POT1 proteins (14,35), alanine substitutions of the corresponding Asparagus POT1 amino acids F67 and Y94 completely abolished DNA binding (Figure 7A, lanes 4 and 5). A similar result was obtained in ZmPOT1b_N with a F89A mutation (corresponding to F62 in human POT1) (Figure 7A, compare lanes 7 and 8). As a control, a F60A mutation in AoPOT1, which affects a non-conserved amino acid, had no effect on AoPOT1 binding (Figure 7A, lane 6). Altogether, these results argue that evolutionarily conserved aromatic amino acids in OB1 are important for telomeric DNA binding across kingdoms, and imply that the overall architecture of OB1 in plant POT1 proteins is similar to that of its mammalian and yeast counterparts.

DISCUSSION

The POT1 protein family represents an evolutionarily conserved group of telomeric DNA-binding factors with essential functions in chromosome end protection and telomere length regulation. Although the major POT1 functions appear to be conserved in most branches of eukaryotic life, previous data in plants indicated that Arabidopsis POT1 proteins evolved unusual functions in regulating telomerase, a property not dependent on

physical contact with telomeric DNA (30). This study addresses this phenomenon further and provides a reconciling view that in some plant species POT1 may indeed have functions similar to those described for non-plant POT1 proteins.

DNA-binding properties of plant POT1 proteins

A well-established phylogenetic hierarchy for the major groups within the green plant lineage allows us an opportunity to examine changes in telomere-related genes in an evolutionary context. Here we compare the DNA-binding properties of three plant POT1 proteins: OIPOT1_N, from the earliest branching lineage analyzed in our study, and AoPOT1 and ZmPOT1b_N from angiosperms. Our biochemical analysis reveals fundamental differences in the nucleic acid binding activity of POT1 proteins across the plant kingdom (Supplemental Table S1). Several of the DNA-binding properties of OIPOT1 are reminiscent of non-plant POT1 proteins and contrast sharply with AoPOT1 and ZmPOT1b. Specifically, the minimum tight-binding sequence of OIPOT1_N consists of 10 nucleotides, a number similar to MBS of POT1 proteins from yeast and animals (13). In contrast, POT1 proteins from angiosperms require fewer nucleotides for efficient binding, with Asparagus POT1 (8–9 nucleotides) being on the lower end of the spectrum and ZmPOT1b_N displaying a short MBS of only 7 nucleotides. Interestingly, two polypeptides harboring a single OB-fold, *S. pombe* Pot1pN and *C. elegans* CeOB2, recognize a shorter, 6-nucleotide MBS (37,40). Thus, ZmPOT1b_N has the smallest MBS among all currently characterized POT1 proteins bearing at least two OB-folds.

Second, plant POT1 proteins show significant variation in the way they interact with cognate DNA. For other POT1 proteins, only a subset of MBS nucleotides is specifically recognized and makes important contributions to binding (13). A similar situation is observed for ZmPOT1b_N, where four out of seven nucleotides in the MBS are required for binding. Asparagus POT1 follows the same trend, with six of the nine MBS nucleotides needed for binding, although all these crucial AoPOT1 MBS nucleotides are localized at the 3'-end of the oligonucleotide. In contrast, all 10 MBS nucleotides are required for OIPOT1 binding, a phenomenon not previously observed for other POT1 proteins. These data suggest that the mechanism responsible for specific recognition of single-stranded telomeric DNA may significantly differ between POT1 proteins from green algae and land plants. Overall, we conclude that the telomeric DNA-binding properties of plant POT1 proteins are evolving rapidly.

Evidence for co-evolution of telomeric DNA and POT1 proteins in plants

The biochemical similarities and differences between ZmPOT1b_N and AoPOT1 may provide clues to the apparent co-evolution of Asparagus POT1 and the telomere repeat sequence in this species. ZmPOT1b_N requires only 7 nt (one full TTAGGG repeat) for efficient binding, while AoPOT1 requires a very similar, but longer

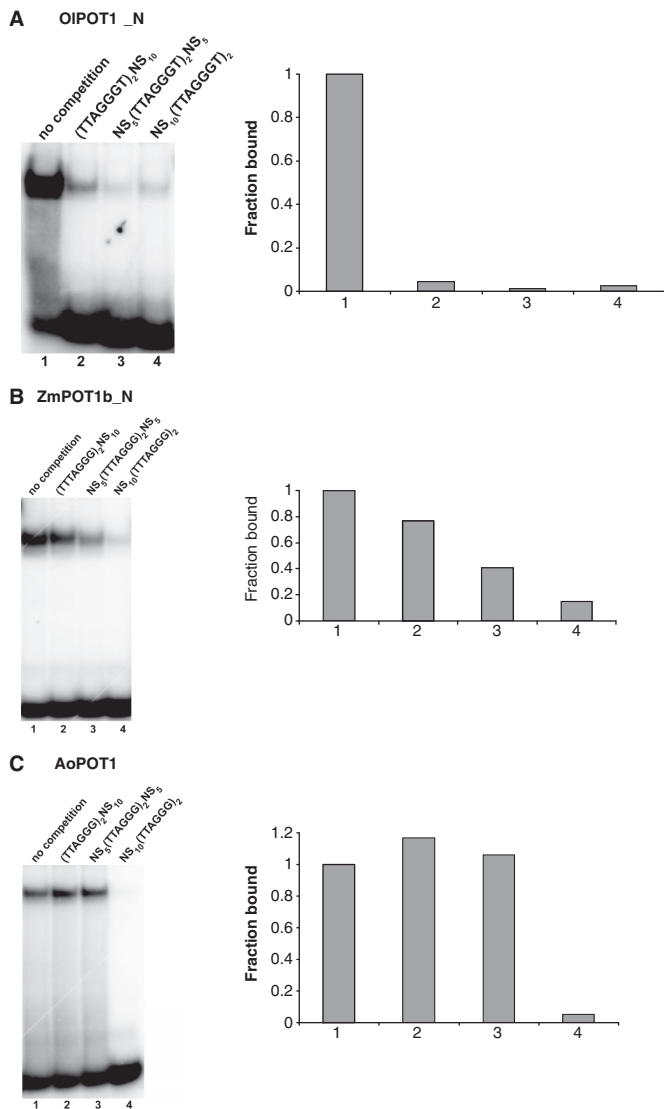


Figure 6. End-binding preference of plant POT1 proteins. (A) EMSA results are shown for reactions with OIPOT1_N using radioactively labeled (TTAGGGT)₂ oligonucleotide as a probe in the absence (lane 1) or presence of 5X excess cold competitors containing the same telomeric sequence located either 5'-terminally (lane 2) or 3'-terminally (lane 4) to the 10-nt non-telomeric sequence NS₁₀ (CTCTACCAAA), or flanked by 5-nt non-telomeric NS₅ sequences (CTCTA and CCAAA) on both ends (lane 3). The fraction of complex bound to the labeled oligonucleotide is plotted on the right with binding in the absence of competitor set at 1.0. (B) and (C) Competition assays for ZmPOT1b_N bound to radioactively labeled (TTTAGGG)₂ oligonucleotide (B) and AoPOT1 bound to radioactively labeled (TTAGGG)₂ oligonucleotide (C). Lane designation and quantification as in (A).

nine-nucleotide substrate GGGTTAGGG. Since the three 5'-terminal Gs do not appear to make significant contributions to AoPOT1 binding, our data suggest that these additional nucleotides in Asparagus POT1 MBS may stabilize this protein's interaction with shorter human-type telomere repeats. Notably, the remaining six nucleotides in AoPOT1 MBS, TTAGGG, are identical to the 3'-end nucleotides present in ZmPOT1b_N MBS. Complementary substitutions in most of these nucleotides completely

abolish protein binding, suggesting that TTAGGG sequence is crucial for specific interaction with POT1 proteins. These data may also explain the requirement of AoPOT1 and, to a lesser extent, ZmPOT1b_N for the presence of telomeric repeats on the 3'-end of DNA oligonucleotides. Such 3'-end positioning is likely necessary to improve or stabilize Asparagus POT1 interaction with the G-overhang, which, in turn, may result in better regulation of G-overhang length or interaction with telomerase.

What is the molecular basis for the stable association of POT1 proteins from angiosperms with hexanucleotide telomere repeats? Since both ZmPOT1b_N and AoPOT1 are capable of specifically binding to TTAGGG repeats, this biochemical feature must have evolved in the common ancestor of maize and Asparagus. We note that AoPOT1 and ZmPOT1b have a similar tolerance to nucleotide substitutions in certain MBS positions. Specifically, two different nucleotides, T (the extra T nucleotide present only in the plant TTTAGGG repeat, but not in the hexanucleotide TTAGGG sequence) and G (which replaces this T in the context of AoPOT1 MBS GGGTTAGGG) are equally tolerated by ZmPOT1b and AoPOT1 in the seventh MBS position (counting from the 3'-end of each oligonucleotide). This property may have originally evolved as a response to the known ability of many plant telomerases to naturally generate mutant telomere repeats containing one less or one more T (25,41,42). These so-called T-slippage events represent the most commonly detected type of telomerase error in plants *in vitro* and *in vivo*. The ability to tolerate T-slippage could have potentially allowed the ancestral POT1 protein to remain bound to the mutant human-type TTAGGG sequences. Another possibility is that decreased affinity to the cognate telomere repeats may help to dislodge POT1 from the G-overhang by other DNA-binding proteins, such as RPA or the CST complex during telomere replication (43).

Recognition of non-cognate telomere repeats is not unique to higher plant POT1 proteins. *Oxytricha nova* telomere end-binding protein (OnTEBP) stably associates with non-cognate telomeric sequences by facilitating significant conformational changes in DNA oligonucleotides via a phenomenon termed nucleotide shuffling, during which DNA sequence register shifts and entire nucleotides are excluded from the protein-DNA complex (44). Chicken and human POT1 proteins are also capable of interacting with non-cognate DNA oligonucleotides in competition assays (12,16). Similarly, in direct EMSA assays, *S. pombe* POT1 specifically binds DNA sequences resembling telomere repeats present in *Tetrahymena thermophyla*, *O. nova* and even *Saccharomyces cerevisiae* (45). Taken together, these observations suggest that relaxed DNA sequence specificity may be a common characteristic of POT1 proteins. This property could be especially beneficial in organisms such as Paramecium, where telomerase synthesizes an unusually high number of mutant telomere repeats (46). Likewise, in Asparagus plant lineage relaxed POT1 telomeric DNA sequence specificity would be beneficial in an evolutionary context and may have contributed to the survival of this entire plant order.

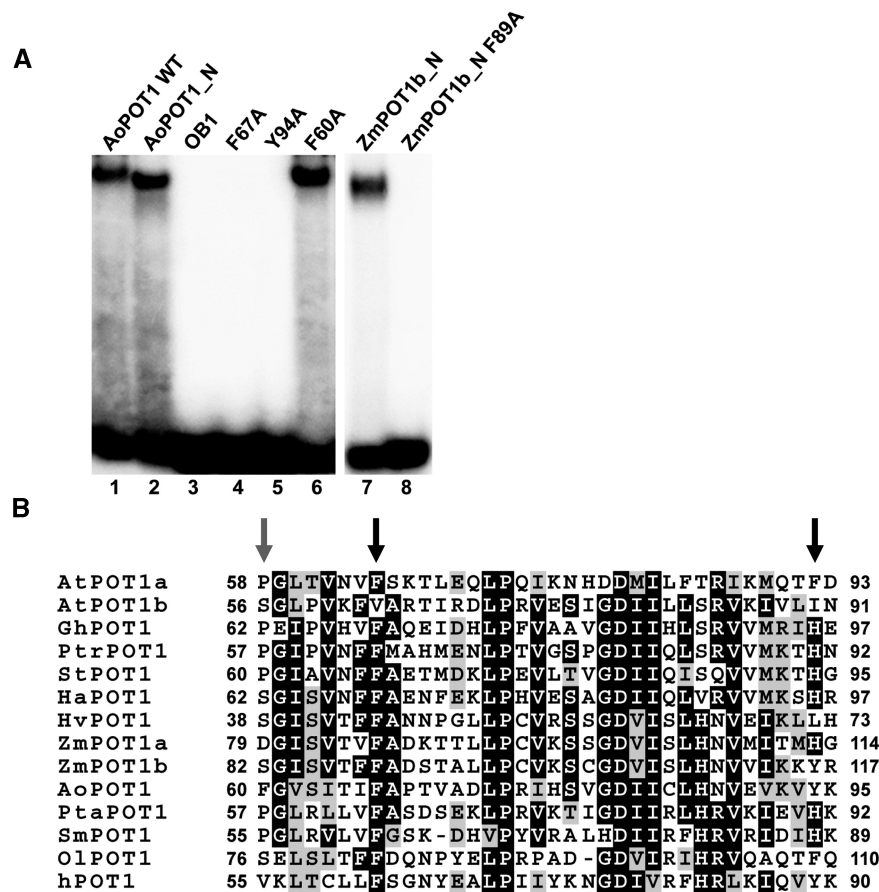


Figure 7. Mutational analysis of plant POT1 proteins. (A) EMSA assays of AoPOT1 and ZmPOT1b_N truncation and point mutants. DNA-binding reactions were performed with wild-type AoPOT1 (lane 1), AoPOT1 truncation constructs (lanes 2 and 3) and point mutants (lanes 4–6) as well as with wild-type ZmPOT1b_N (lane 7) and its point mutant F89A (lane 8). The labeled probes are GGGTTAGGG for AoPOT1 and TTTAGGG for ZmPOT1b_N. (B) Partial alignment of plant POT1 proteins with human POT1. An OBI region with a high degree of inter-kingdom amino acid similarity is shown. Black arrows indicate the positions of two catalytically important human POT1 residues and the corresponding aromatic amino acids F67 and Y94 in AoPOT1. Grey arrow designates the location of F60 in AoPOT1, which has no effect on telomeric DNA binding. Numbers indicate amino acid positions relative to the start codon. Abbreviations: At, *Arabidopsis thaliana*; Gh, *Gossypium hirsutum*; Ptr, *Populus trichocarpa*; St, *Solanum tuberosum*; Ha, *Helianthus argophyllus*; Hv, *Hordeum vulgare*; Zm, *Zea mays*; Ao, *Asparagus officinalis*; Pta, *Pinus taeda*; Sm, *Selaginella moellendorffii* and Ol, *Ostreococcus lucimarinus*. Alignment was generated with MEGA 3 software (52) and visualized in the BOXSHADE format.

Evolutionary changes in plant POT1 functions

Interestingly, we failed to detect *in vitro* telomeric DNA binding for eight POT1 proteins from the evolutionarily diverse group of plant organisms analyzed in this study and for six previously characterized POT1 proteins from the Brassicaceae family of plants, which includes *Arabidopsis* (26). We can not rule out the possibility that some RRL-expressed plant POT1 proteins lack proper post-translational modifications or other requisites for efficient binding to telomeric DNA *in vitro*. However, we note that RRL-expressed POT1 proteins from yeast (12), mammals (36) and three different plants (this study) can efficiently bind telomeric DNA under the same conditions. Moreover, several lines of evidence suggest that telomeric DNA binding may not be the major *in vivo* function of POT1 proteins in *Arabidopsis* and, perhaps, in other plants. In striking contrast to yeast and mammalian POT1 proteins, *Arabidopsis* POT1a acts as a positive regulator of telomerase activity and is only

enriched at the telomeres in S-phase, when telomerase is thought to act (30). Moreover, *Arabidopsis* POT1b appears to be a negative regulator of telomerase activity (E. Shakirov, A. Nelson and D. Shippen, in preparation). Recent data indicate that AtPOT1a and AtPOT1b associate directly with *Arabidopsis* telomerase RNA in regions outside the telomere template domain (C. Cifuentes-Rojas *et al.*, manuscript in preparation). While RNA binding is associated with other OB-fold containing proteins, for example, translation factors (47), POT1 proteins have not been previously reported to bind RNA. Thus, the interaction of *Arabidopsis* POT1a and POT1b with telomerase RNA appears to represent a major evolutionary shift in plant POT1 functions from DNA to RNA binding.

Arabidopsis and maize are currently the only plants known to harbor more than one *POT1* orthologue. In both cases, *POT1* genes were likely duplicated around 30mya, when the lineages leading to *Arabidopsis*

and maize experienced independent whole-genome duplication events (34,48,49). Despite a similar evolutionary timeframe, the fate of these duplicated *POT1* genes appears to be distinct. First, like the two mouse *POT1* proteins, which evolved partially non-overlapping functions (36,50,51), the maize *POT1* paralogs share ~75% amino acid similarity. In contrast, Arabidopsis *POT1a* and *POT1b* display much lower sequence conservation, retaining only ~50% amino acid similarity overall. Second, while both Arabidopsis *POT1* proteins bind telomerase RNA instead of telomeric DNA, only one of the maize *POT1* proteins, *ZmPOT1a*, lost the ability to bind telomeric DNA, raising the interesting possibility that *ZmPOT1a* evolved to bind the maize telomerase RNA. Although further analysis of the telomere complex in maize will be required to test this model, the comparative biochemical analysis of plant *POT1* proteins described here underscores the remarkably rapid evolution of the OB-fold nucleic acid binding interface.

ACCESSION NUMBERS

EU880295, EU880297, EU880298, EU880301, EU880303–EU880306, EU883536, FJ516399.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Sandy Chang for providing the mouse *POT1a_N* construct. They are also grateful to Thomas D. McKnight and to members of the Shippen laboratory for insightful discussions and critical reading of the manuscript.

FUNDING

National Institutes of Health [GM065383 to D.E.S.]. Funding for open access charge: National Institutes of Health [GM065383 to D.E.S.].

Conflict of interest statement. None declared.

REFERENCES

- Gilson,E. and Geli,V. (2007) How telomeres are replicated. *Nat. Rev. Mol. Cell Biol.*, **8**, 825–838.
- Xin,H., Liu,D. and Songyang,Z. (2008) The telosome/shelterin complex and its functions. *Genome Biol.*, **9**, 232.
- Palm,W. and de Lange,T. (2008) How shelterin protects mammalian telomeres. *Annu. Rev. Genet.*, **42**, 301–334.
- Yoon,H.S., Hackett,J.D., Ciniglia,C., Pinto,G. and Bhattacharya,D. (2004) A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.*, **21**, 809–818.
- McKnight,T.D. and Shippen,D.E. (2004) Plant telomere biology. *Plant Cell*, **16**, 794–803.
- Higashiyama,T., Maki,S. and Yamada,T. (1995) Molecular organization of *Chlorella vulgaris* chromosome I: presence of telomeric repeats that are conserved in higher plants. *Mol. Gen. Genet.*, **246**, 29–36.
- Petracek,M.E., Lefebvre,P.A., Silflow,C.D. and Berman,J. (1990) Chlamydomonas telomere sequences are A+T-rich but contain three consecutive G-C base pairs. *Proc. Natl Acad. Sci. USA*, **87**, 8222–8226.
- Pich,U., Fuchs,J. and Schubert,I. (1996) How do Alliaceae stabilize their chromosome ends in the absence of TTTAGGG sequences? *Chromosome Res.*, **4**, 207–213.
- Sykorova,E., Lim,K.Y., Chase,M.W., Knapp,S., Leitch,I.J., Leitch,A.R. and Fajkus,J. (2003) The absence of Arabidopsis-type telomeres in Cestrum and closely related genera Vestia and Sessea (Solanaceae): first evidence from eudicots. *Plant J.*, **34**, 283–291.
- Fajkus,J., Sykorova,E. and Leitch,A.R. (2005) Telomeres in evolution and evolution of telomeres. *Chromosome Res.*, **13**, 469–479.
- Wikstrom,N., Savolainen,V. and Chase,M.W. (2001) Evolution of the angiosperms: calibrating the family tree. *Proc. Biol. Sci.*, **268**, 2211–2220.
- Baumann,P. and Cech,T.R. (2001) Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science*, **292**, 1171–1175.
- Croy,J.E. and Wuttke,D.S. (2006) Themes in ssDNA recognition by telomere-end protection proteins. *Trends Biochem. Sci.*, **31**, 516–525.
- Lei,M., Podell,E.R. and Cech,T.R. (2004) Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat. Struct. Mol. Biol.*, **11**, 1223–1229.
- Croy,J.E., Podell,E.R. and Wuttke,D.S. (2006) A new model for *Schizosaccharomyces pombe* telomere recognition: the telomeric single-stranded DNA-binding activity of Pot1-389. *J. Mol. Biol.*, **361**, 80–93.
- Wei,C. and Price,C.M. (2004) Cell cycle localization, dimerization, and binding domain architecture of the telomere protein cPot1. *Mol. Cell Biol.*, **24**, 2091–2102.
- Horvath,M.P., Schweiker,V.L., Bevilacqua,J.M., Ruggles,J.A. and Schultz,S.C. (1998) Crystal structure of the *Oxytricha nova* telomere end binding protein complexed with single strand DNA. *Cell*, **95**, 963–974.
- Loayza,D., Parsons,H., Donigian,J., Hoke,K. and de Lange,T. (2004) DNA binding features of human POT1: A nonamer 5'-TAGGGTTAG-3' minimal binding site, sequence specificity, and internal binding to multimeric sites. *J. Biol. Chem.*, **279**, 13241–13248.
- Ye,J.Z., Hockemeyer,D., Krutchinsky,A.N., Loayza,D., Hooper,S.M., Chait,B.T. and de Lange,T. (2004) POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Genes Dev.*, **18**, 1649–1654.
- Liu,D., Safari,A., O'Connor,M.S., Chan,D.W., Laegerel,A., Qin,J. and Songyang,Z. (2004) PTPN22 interacts with POT1 and regulates its localization to telomeres. *Nat. Cell Biol.*, **6**, 673–680.
- Loayza,D. and de Lange,T. (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature*, **424**, 1013–1018.
- Kelleher,C., Kurth,I. and Lingner,J. (2005) Human protection of telomeres 1 (POT1) is a negative regulator of telomerase activity *in vitro*. *Mol. Cell Biol.*, **25**, 808–818.
- Wang,F., Podell,E.R., Zaug,A.J., Yang,Y., Baciu,P., Cech,T.R. and Lei,M. (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature*, **445**, 506–510.
- Baumann,P., Podell,E. and Cech,T.R. (2002) Human Pot1 (protection of telomeres) protein: cytolocalization, gene structure, and alternative splicing. *Mol. Cell Biol.*, **22**, 8079–8087.
- Shakirov,E.V., Salzberg,S.L., Alam,M. and Shippen,D.E. (2008) Analysis of *Carica papaya* telomeres and telomere-associated proteins: insights into the evolution of telomere maintenance in Brassicales. *Tropical Plant Biol.*, **1**, 202–215.
- Shakirov,E.V., McKnight,T.D. and Shippen,D.E. (2009) POT1-independent single-strand telomeric DNA binding activities in Brassicaceae. *Plant J.*, **58**, 1004–1015.
- Kuchar,M. and Fajkus,J. (2004) Interactions of putative telomere-binding proteins in *Arabidopsis thaliana*: identification of functional TRF2 homolog in plants. *FEBS Lett.*, **578**, 311–315.
- Tani,A. and Murata,M. (2005) Alternative splicing of Pot1 (Protection of telomere)-like genes in *Arabidopsis thaliana*. *Genes Genet. Syst.*, **80**, 41–48.
- Shakirov,E.V., Surovtseva,Y.V., Osburn,N. and Shippen,D.E. (2005) The Arabidopsis Pot1 and Pot2 proteins function in telomere length homeostasis and chromosome end protection. *Mol. Cell Biol.*, **25**, 7725–7733.

30. Surovtseva, Y.V., Shakirov, E.V., Vespa, L., Osbun, N., Song, X. and Shippen, D.E. (2007) Arabidopsis POT1 associates with the telomerase RNP and is required for telomere maintenance. *EMBO J.*, **26**, 3653–3661.
31. Rossignol, P., Collier, S., Bush, M., Shaw, P. and Doonan, J.H. (2007) Arabidopsis POT1A interacts with TERT-V(18), an N-terminal splicing variant of telomerase. *J. Cell Sci.*, **120**, 3678–3687.
32. Cocciolone, S.M. and Cone, K.C. (1993) Pl-Bh, an anthocyanin regulatory gene of maize that leads to variegated pigmentation. *Genetics*, **135**, 575–588.
33. Fitzgerald, M.S., Riha, K., Gao, F., Ren, S., McKnight, T.D. and Shippen, D.E. (1999) Disruption of the telomerase catalytic subunit gene from Arabidopsis inactivates telomerase and leads to a slow loss of telomeric DNA. *Proc. Natl Acad. Sci. USA*, **96**, 14813–14818.
34. Paterson, A.H., Bowers, J.E. and Chapman, B.A. (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl Acad. Sci. USA*, **101**, 9903–9908.
35. He, H., Multani, A.S., Cosme-Blanco, W., Tahara, H., Ma, J., Pathak, S., Deng, Y. and Chang, S. (2006) POT1b protects telomeres from end-to-end chromosomal fusions and aberrant homologous recombination. *EMBO J.*, **25**, 5180–5190.
36. Wu, L., Multani, A.S., He, H., Cosme-Blanco, W., Deng, Y., Deng, J.M., Bachilo, O., Pathak, S., Tahara, H., Bailey, S.M. *et al.* (2006) Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres. *Cell*, **126**, 49–62.
37. Raices, M., Verdun, R.E., Compton, S.A., Haggblom, C.I., Griffith, J.D., Dillin, A. and Karlseder, J. (2008) *C. elegans* telomeres contain G-strand and C-strand overhangs that are bound by distinct proteins. *Cell*, **132**, 745–757.
38. Burr, B., Burr, F.A., Matz, E.C. and Romero-Severson, J. (1992) Pinning down loose ends: mapping telomeres and factors affecting their length. *Plant Cell*, **4**, 953–960.
39. Lei, M., Podell, E.R., Baumann, P. and Cech, T.R. (2003) DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature*, **426**, 198–203.
40. Lei, M., Baumann, P. and Cech, T.R. (2002) Cooperative binding of single-stranded telomeric DNA by the Pot1 protein of *Schizosaccharomyces pombe*. *Biochemistry*, **41**, 14560–14568.
41. Fitzgerald, M.S., Shakirov, E.V., Hood, E.E., McKnight, T.D. and Shippen, D.E. (2001) Different modes of *de novo* telomere formation by plant telomerases. *Plant J.*, **26**, 77–87.
42. Mizuno, H., Wu, J., Katayose, Y., Kanamori, H., Sasaki, T. and Matsumoto, T. (2008) Chromosome-specific distribution of nucleotide substitutions in telomeric repeats of rice (*Oryza sativa* L.). *Mol. Biol. Evol.*, **25**, 62–68.
43. Gao, H., Cervantes, R.B., Mandell, E.K., Otero, J.H. and Lundblad, V. (2007) RPA-like proteins mediate yeast telomere function. *Nat. Struct. Mol. Biol.*, **14**, 208–214.
44. Theobald, D.L. and Schultz, S.C. (2003) Nucleotide shuffling and ssDNA recognition in *Oxytricha nova* telomere end-binding protein complexes. *EMBO J.*, **22**, 4314–4324.
45. Trujillo, K.M., Bunch, J.T. and Baumann, P. (2005) Extended DNA binding site in Pot1 broadens sequence specificity to allow recognition of heterogeneous fission yeast telomeres. *J. Biol. Chem.*, **280**, 9119–9128.
46. McCormick-Graham, M., Haynes, W.J. and Romero, D.P. (1997) Variable telomeric repeat synthesis in *Paramecium tetraurelia* is consistent with misincorporation by telomerase. *EMBO J.*, **16**, 3233–3242.
47. Theobald, D.L., Mitton-Fry, R.M. and Wuttke, D.S. (2003) Nucleic acid recognition by OB-fold proteins. *Annu. Rev. Biophys. Biomol. Struct.*, **32**, 115–133.
48. Kellogg, E.A. (1998) Relationships of cereal crops and other grasses. *Proc. Natl Acad. Sci. USA*, **95**, 2005–2010.
49. Schranz, M.E. and Mitchell-Olds, T. (2006) Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. *Plant Cell*, **18**, 1152–1165.
50. Hockemeyer, D., Daniels, J.P., Takai, H. and de Lange, T. (2006) Recent expansion of the telomeric complex in rodents: two distinct POT1 proteins protect mouse telomeres. *Cell*, **126**, 63–77.
51. Palm, W., Hockemeyer, D., Kibe, T. and de Lange, T. (2009) Functional dissection of human and mouse POT1 proteins. *Mol. Cell Biol.*, **29**, 471–482.
52. Kumar, S., Tamura, K. and Nei, M. (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform.*, **5**, 150–163.