Evolution of the Telomere-Associated Protein POT1a in *Arabidopsis thaliana* Is Characterized by Positive Selection to Reinforce Protein–Protein Interaction

Mark A. Beilstein,*[,1,2] Kyle B. Renfrew,[2] Xiangyu Song,‡[,2] Eugene V. Shakirov,*[,3,4] Michael J. Zanis,[5] and Dorothy E. Shippen*[,2]

[1]School of Plant Sciences, University of Arizona

[2]Department of Biochemistry and Biophysics, Texas A&M University

[3]Department of Integrative Biology, University of Texas at Austin

[4]Institute of Fundamental Medicine and Biology, Kazan (Volga Region) Federal University, Kazan, Republic of Tatarstan, Russia

[5]Department of Biology, Seattle University

‡Present address: Proctor & Gamble, Tian Zhu, Kong Gang Development Zone B, Shunyi, Beijing, China

***Corresponding author:** E-mail: mbeilstein@email.arizona.edu; eugene-shakirov@utexas.edu; dshippen@tamu.edu.

Associate editor: Michael Purugganan

Abstract

Gene duplication is a major driving force in genome evolution. Here, we explore the nature and origin of the *POT1* gene duplication in *Arabidopsis thaliana*. Protection of Telomeres (POT1) is a conserved multifunctional protein that modulates telomerase activity and its engagement with telomeres. *Arabidopsis thaliana* encodes two divergent POT1 paralogs termed AtPOT1a and AtPOT1b. AtPOT1a positively regulates telomerase activity, whereas AtPOT1b is proposed to negatively regulate telomerase and promote chromosome end protection. Phylogenetic analysis uncovered two independent *POT1* duplication events in the plant kingdom, including one at the base of Brassicaceae. Tests for positive selection implemented in PAML revealed that the Brassicaceae *POT1a* lineage experienced positive selection postduplication and identified three amino acid residues with signatures of positive selection. A sensitive and quantitative genetic complementation assay was developed to assess *POT1a* function in *A. thaliana*. The assay showed that *AtPOT1a* is functionally distinct from single-copy *POT1* genes in other plants. Moreover, for two of the sites with a strong signature of positive selection, substitutions that swap the amino acids in AtPOT1a for residues found in AtPOT1b dramatically compromised AtPOT1a function in vivo. In vitro-binding studies demonstrated that all three sites under positive selection specifically enhance the AtPOT1a interaction with CTC1, a core component of the highly conserved CST (CTC1/STN1/TEN1) telomere protein complex. Our results reveal a molecular mechanism for the role of these positively selected sites in AtPOT1a. The data also provide an important empirical example to refine theories of duplicate gene retention, as the outcome of positive selection here appears to be reinforcement of an ancestral function, rather than neofunctionalization. We propose that this outcome may not be unusual when the duplicated protein is a component of a multisubunit complex whose function is in part specified by other members.

Key words: **telomere, *Arabidopsis*, positive selection, CTC1, Brassicaceae, POT1.**

Introduction

Publication of the *Arabidopsis thaliana* whole-genome sequence more than a decade ago (Arabidopsis Genome Initiative 2000) continues to provide biologists with an important view of the composition and evolution of plant genomes, especially as compared with other eukaryotic lineages. One surprising finding is evidence of widespread gene and genome duplications in *A. thaliana*. We now know that *A. thaliana* is not unique among plants in having a genome characterized by duplication-fueled gene expansion (Cui et al. 2006). In fact, hybridization and other genome duplication events have impacted lineage diversification (Beilstein et al. 2010) and may even have permitted some lineages to survive through mass extinction events (Fawcett et al. 2009).

As our appreciation for the extent of duplications increases, theories to explain the retention of duplicate genes have been proposed. These theories fall into three major categories: Neofunctionalization (NF) (Ohno 1970), subfunctionalization (Force et al. 1999), and maintenance of dosage balance (Birchler and Veitia 2007). Since the outline of these alternatives, evolutionary biologists have sought empirical examples to strengthen theory. At the same time, theories that refine these major classes have emerged, including escape from adaptive conflict (EAC) (Des Marais and Rausher 2008) and positive dosage (Kondrashov et al. 2002; Innan and Kondrashov 2010).

Tests of molecular evolution at the protein level permit the processes that underlie some of these theories to be examined. For example, in the NF model one of the duplicate

Article

copies evolves a new function not performed by its single-copy ancestor. At the molecular level, this change in the protein is described by a signature of positive Darwinian selection in which the nonsynonymous substitution rate (d*N*) outpaces the synonymous substitution rate (d*S*), causing the ratio of the two values, ω, to exceed 1 (i.e., $dN/dS = \omega > 1$) (Zhang et al. 2005). In contrast, subfunctionalization parses the functions of the ancestral single-copy gene between the descendant copies and can be driven by changes in expression through differential degeneration of promoter regions. Such a process does not require changes to the protein coding region, and thus the descendant gene copies may lack evidence of positive selection. Finally, retention by dosage balance describes situations following whole-genome duplication where stoichiometry in biochemical pathways must be maintained to achieve optimal function. Similar to subfunctionalization, changes to protein coding regions are not required, nor are changes to regulatory domains necessary. Rather, the expectation is that other members of a particular pathway will be represented in the genome by multiple copies as well.

The framework for examining gene duplication events has become a powerful tool for understanding the evolution of protein function. Here, we examined the duplication history of the *POT1* (Protection of Telomeres 1) gene. Telomeres are an ancient hallmark feature of linear chromosomes, essential for genome stability and long-term proliferative capacity of cells. The GT-rich sequence of telomeric DNA repeats is conserved across eukaryotes, but composition of telomere-associated proteins varies significantly between distant organisms (Linger and Price 2009; Lue 2010). POT1 is one of the few telomere proteins that can be readily identified from various eukaryotic lineages through its signature oligonucleotide/oligosaccharide-binding (OB) folds (Theobald and Wuttke 2004).

POT1 was first described as a single-strand telomeric DNA-binding protein that protects chromosome ends from nucleolytic degradation and end-to-end chromosome fusion (Baumann and Cech 2001). More recent studies indicate that POT1 also acts in concert with other telomere-associated proteins to either positively (Wang et al. 2007) or negatively (Chen et al. 2012) regulate telomerase, the ribonucleoprotein (RNP) reverse transcriptase responsible for synthesizing and maintaining telomeric DNA. Thus, POT1 is a multifunctional protein that promotes genome integrity by physically protecting chromosome ends and modulating telomerase access.

Although most eukaryotes harbor a single *POT1* gene (Baumann et al. 2002), two or more *POT1* paralogs have been reported in mouse (Hockemeyer et al. 2006; Wu et al. 2006), worms (Raices et al. 2008), and some ciliates (Jacob et al. 2007). *POT1* duplication is also evident in the plant kingdom (Shakirov, McKnight, et al. 2009; Shakirov, Song, et al. 2009) and was first described in *A. thaliana* (Shakirov et al. 2005). AtPOT1a and AtPOT1b exhibit relatively low amino acid sequence identity (49%), and unlike the single-copy *POT1* genes from vertebrates, fission yeast and the moss *Physcomitrella patens* (Shakirov et al. 2010), AtPOT1a and AtPOT1b have low affinity for telomeric DNA in vitro (Shakirov, McKnight, et al. 2009). Instead, AtPOT1a and AtPOT1b bind telomerase RNA and assemble into distinct RNP complexes with telomerase enzyme components. AtPOT1a forms RNP with the canonical telomerase RNA subunit TER1 (Cifuentes-Rojas et al. 2011, 2012), which is responsible for the synthesis of telomeric DNA repeats on chromosome ends. AtPOT1a also stimulates telomerase enzyme activity (Surovtseva et al. 2007; Renfrew et al. 2014). In addition, AtPOT1a physically associates with CST (CTC1/STN1/TEN1), the core telomere-binding complex in *A. thaliana* (Renfrew et al. 2014). POT1 interaction with CST components has also been reported in vertebrates (Chen et al. 2012, 2013; Wu et al. 2012), suggesting that this aspect of POT1 behavior is widely conserved and may represent an ancestral function.

In contrast to AtPOT1a, AtPOT1b assembles into an alternative RNP complex that contains a noncanonical TER, TER2 (Cifuentes-Rojas et al. 2012). The POT1b/TER2 RNP is implicated in the negative regulation of telomerase enzyme activity. Thus, AtPOT1a and AtPOT1b modulate telomerase, but with opposing outcomes. As with AtPOT1a, AtPOT1b may possess some of the functional properties of single-copy POT1 proteins. Overexpression of a dominant negative allele of *POT1b* results in catastrophic telomere shortening and end-to-end chromosome fusions (Shakirov et al. 2005), arguing that the chromosome end protection function of POT1 may be retained in POT1b. Altogether, these findings indicate that *A. thaliana* POT1 proteins contribute to multiple aspects of telomere biology, and postduplication at least some of the functions of AtPOT1a and AtPOT1b diverged.

Here, we investigate the timing of the duplication event responsible for the two AtPOT1 proteins and characterize the processes underlying retention of the duplicate copies. Using a combination of in vivo genetic complementation and in vitro protein-binding assays, we set out to test whether neo- or subfunctionalization models better describe the evolution of *POT1* genes postduplication and assess whether specific amino acids with a significant signal of positive selection are important for function. Our results reveal surprisingly rapid divergence between the two *POT1* lineages in Brassicaceae, and uncover selection for an enhanced interaction with a core component of the CST telomere complex. The findings also support a new model for protein evolution in the context of multisubunit complexes.

Results

POT1 Phylogeny and Duplication

Several well-documented genome duplication events occurred during the evolution of land plants (Blanc and Wolfe 2004; Jiao et al. 2011). However, our bioinformatic and phylogenetic analyses indicated that duplicate plant *POT1* genes are rare and that most plant genomes, from green algae *Ostreococcus lucimarinus* to the moss *P. patens* to most flowering plants, harbor a single *POT1* gene (fig. 1*A*). Only two *POT1* gene duplication events were detected for the taxa sampled which occurred independently in different lineages of the Angiosperms.

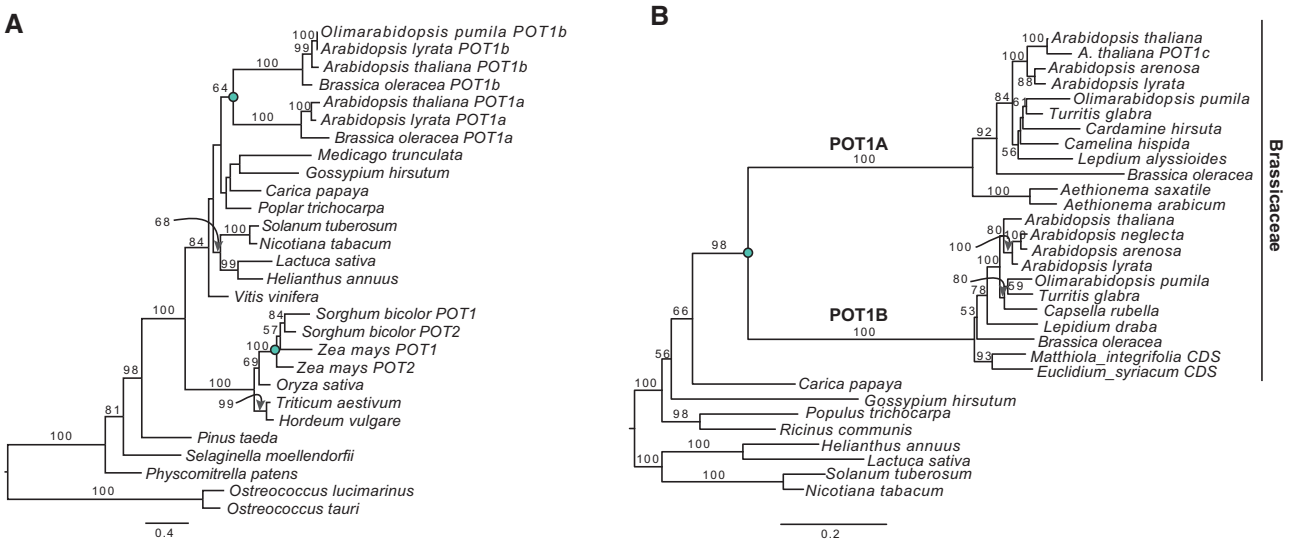

FIG. 1. Phylogenetic tree of plant POT1 proteins. (*A*) POT1 phylogeny for 23 plant species from green algae to flowering plants inferred from an amino acid alignment under the PROTGAMMAWAG option in RAxML v 7.0.4. (*B*) Phylogeny of *POT1a* and *POT1b* from 17 species spanning the family Brassicaceae and other eudicots inferred from a nucleotide alignment of *POT1* under the GTR + Γ model in RAxML. Blue circles at nodes indicate duplication events.

The first example of *POT1* gene duplication is a Panicoideae-specific event in grasses that occurred less than 30 Ma after the divergence of this lineage from the last common ancestor with rice, wheat, and barley (Kellogg 1998; Paterson et al. 2004). The overall amino acid sequence conservation between POT1 proteins within or between the different grass species with one or two *POT1* genes is 70–75%, similar to the situation for the rodent *POT1a* and *POT1b* paralogs (Hockemeyer et al. 2006).

The second duplication occurred in the Brassicaceae (order Brassicales) of eudicots, which includes *A. thaliana*. To explore the origin of this *POT1* duplication in more detail, we analyzed POT1 sequences from 17 species within the family Brassicaceae, as well as *Carica papaya* (papaya, family Caricaceae, order Brassicales), *Gossypium hirsutum* (cotton, order Malvales, sister to Brassicales), and six other species of eudicots. The two *A. thaliana POT1* gene copies are found in regions on chromosomes 2 and 5, which do not belong to the set of canonical duplicate chromosomes that share large stretches of colinearity (e.g., chromosomes 2 and 4) (Vandepoele et al. 2002; Blanc and Wolfe 2004). Given the inferred *POT1* gene tree (fig. 1*B* and supplementary alignment data, Supplementary Material online) and as only a single *POT1* gene was recovered in the whole-genome sequences of papaya and cotton, we conclude that the *POT1* gene duplication giving rise to *POT1a* and *POT1b* lineages occurred near the origin of the Brassicaceae after its divergence from the last common ancestor with papaya approximately 100 Ma (Beilstein et al. 2010).

A Novel Genetic Complementation Assay for AtPOT1a Function

To investigate the functional divergence of Brassicaceae POT1 proteins in vivo, we developed a sensitive and quantitative transgenic rescue assay using *A. thaliana* mutants null for *POT1a* (Surovtseva et al. 2007). Telomeres in the *A. thaliana* Col-0 accession range from 2 to 5 kb (Richards and Ausubel 1988; Shakirov and Shippen 2004). As the telomere tracts of *pot1a* mutants shorten by only about 200–500 bp per plant generation (Surovtseva et al. 2007), partial restoration of *AtPOT1a* function would be difficult to measure quantitatively using a standard telomere detection assay. Therefore, we employed plants doubly deficient in *AtPOT1a* and *Ku70* for *POT1a* complementation (supplementary fig. S1, Supplementary Material online). The Ku70/80 heterodimer is a potent negative regulator of telomere length in *A. thaliana* and in its absence telomeres are extended by up to 10 kb (2-fold) within the first generation (Riha et al. 2002). Telomere expansion is dependent on *AtPOT1a* (Surovtseva et al. 2007) and thus in the presence of a functional *POT1a* transgene the telomeres in $pot1a^{-/-}$ $ku70^{-/-}$ mutants are grossly extended. Furthermore, without a functional copy of *POT1a* telomeres in $pot1a^{-/-}$ $ku70^{-/-}$ mutants shorten dramatically (Surovtseva et al. 2007). Accordingly, this genetic complementation assay provides a means to quantitatively assess *POT1a* function in vivo.

Arabidopsis thaliana mutants homozygous for a T-DNA insertion in *POT1a* and heterozygous for a T-DNA insertion in *Ku70* were transformed with a wild-type *AtPOT1a* construct expressed from its native promoter. An amino terminal FLAG-tag was added to monitor protein expression. Transgene expression for this and other constructs was confirmed by Western blotting or semiquantitative reverse transcription polymerase chain reaction (RT-PCR) (supplementary fig. S2, Supplementary Material online). Telomere length was assessed by Terminal Restriction Fragment (TRF) analysis and mean telomere length was quantified using TeloMetric1.2 (Grant et al. 2001). As expected, telomeres in *pot1a ku70* double mutants were short and spanned

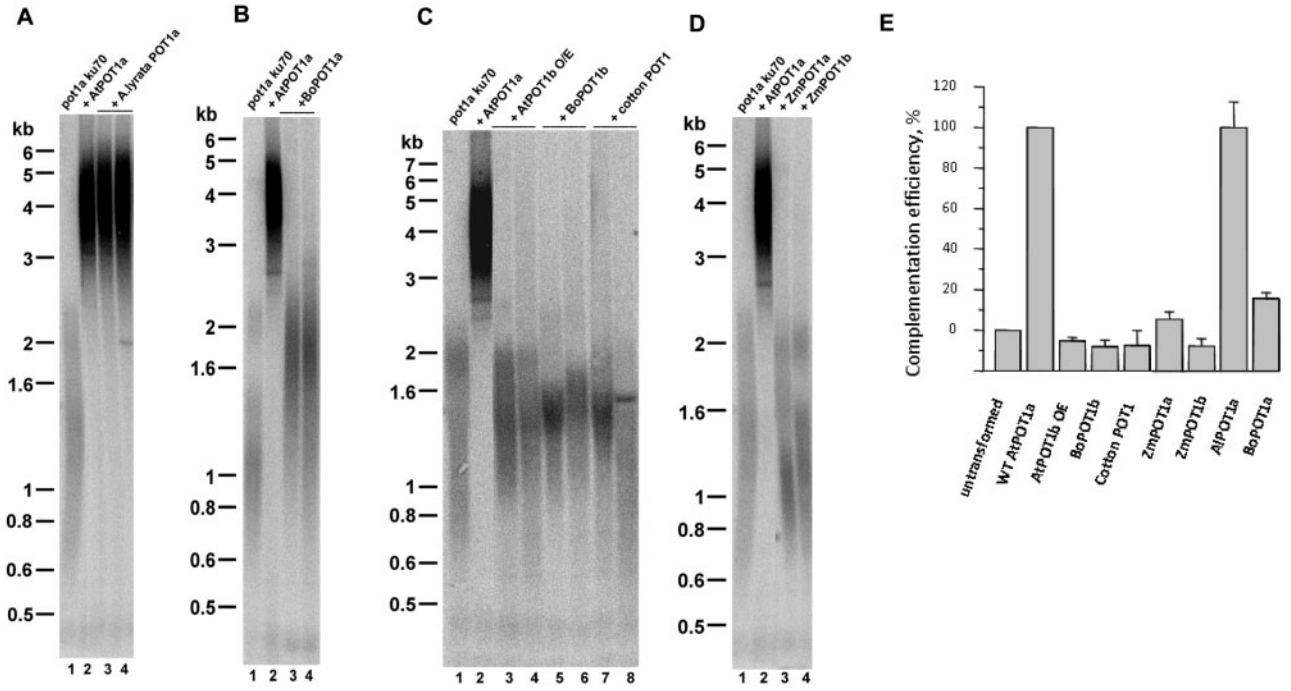

FIG. 2. Cross-species complementation analysis of *AtPOT1a* deficiency. TRF analysis of untransformed *pot1a ku70* mutants and transformants expressing POT1 proteins from *Arabidopsis thaliana* or other plant species. (*A*) Lane 1, untransformed *pot1a ku70* plant; lanes 2–4, *pot1a ku70* plants expressing WT AtPOT1a (lane 2) and *Arabidopsis lyrata* POT1a (lanes 3 and 4). (*B*) Lane 1, untransformed *pot1a ku70*; lanes 2–4, *pot1a ku70* mutants expressing WT AtPOT1a (lane 2) or *Brassica oleracea* (Bo) POT1a (lanes 3 and 4). (*C*) Lane 1, untransformed *pot1a ku70* plant; lanes 2–8, *pot1a ku70* plants expressing WT AtPOT1a (lane 2), overexpressing AtPOT1b (lanes 3 and 4), overexpressing *Brassica oleracea* BoPOT1b (lanes 5 and 6), and *Gossypium hirsutum* GhPOT1 (lanes 7 and 8). (*D*) Lane 1, untransformed *pot1a ku70* plant; lanes 2–4, *pot1a ku70* plants expressing WT AtPOT1a (lane 2), *Zea mays* ZmPOT1a (lane 3), and *Z. mays* ZmPOT1b (lane 4). (*E*) Quantification of telomere length in (*A*)–(*D*). Error bars represent standard error from three independently transformed lines of each genotype.

0.5–2.3 kb (fig. 2*A*, lane 1; supplementary fig. S3*A*, lane 1 and supplementary fig. S3*B*, lane 1, Supplementary Material online). Transformation with an empty vector did not alter telomere length (supplementary fig. S3*A*, lane 2 and supplementary fig. S3*B*, lane 2, Supplementary Material online). Conversely, transgenic plants expressing wild-type AtPOT1a either from the strong 35S promoter or from its native promoter had elongated telomeres that ranged from 2.5 to 6 kb in length (fig. 2*A*, lane 2; supplementary fig. S3*A*, lanes 3, 4 and supplementary fig. S3*B*, lane 3, Supplementary Material online), indicating that the AtPOT1a deficiency was fully complemented by both wild-type AtPOT1a constructs.

To further validate the assay, we analyzed the effects of several AtPOT1a mutations that, based on data for POT1 proteins from yeast and humans, were likely to result in a loss of POT1a function. In *Schizosaccharomyces pombe* POT1, deletion of as few as five amino acids from the highly conserved C-terminus completely abolishes POT1 function (Bunch et al. 2005). The extreme C-terminus of AtPOT1a is highly similar to the corresponding region in *S. pombe* POT1 (supplementary fig. S3*C*, Supplementary Material online). Telomeres in $pot1a^{-/-}$ $ku70^{-/-}$ plants expressing AtPOT1a-C10Δ, which lacks the ten C-terminal amino acids, were not elongated (supplementary fig. S3*B*, lane 4 and supplementary fig. S3*D*, Supplementary Material online). This finding supports the conclusion that the C-terminus of AtPOT1a bears a critical and likely conserved functional motif.

In mammalian POT1, a highly conserved phenylalanine residue (F65 in AtPOT1a) within the first OB-fold domain (OB1) is essential for telomeric DNA binding (Lei et al. 2004; He et al. 2006). In addition, mutation of the corresponding residue in the *Asparagus officinale* POT1 and the *Zea mays* POT1a proteins to alanine completely abolishes DNA binding in vitro (Shakirov, Song, et al. 2009). Consistent with these findings, telomeres in *A. thaliana* $pot1a^{-/-}$ $ku70^{-/-}$ mutants expressing AtPOT1a–F65A were elongated to only 12% of the level observed with wild-type AtPOT1a (supplementary fig. S3*B*, lane 5 and supplementary fig. S3*D*, Supplementary Material online). This result implies that nucleic acid binding is an important component of AtPOT1a protein function in vivo. Overall, we conclude that key functional elements within POT1 proteins are conserved in AtPOT1a, and further that our genetic complementation assay provides a reliable gauge of AtPOT1a function in vivo.

Evolution of Brassicaceae POT1a Proteins

We used the genetic complementation assay to examine conservation of POT1a function across different members of Brassicaceae. *Arabidopsis lyrata* shared the last common ancestor with *A. thaliana* approximately 13 Ma (Beilstein et al. 2010), and hence is one of the closest extant relatives of

A. thaliana. *Arabidopsis lyrata* POT1a protein exhibits 94% amino acid similarity to AtPOT1a overall. Cross-species complementation using AlPOT1a fully rescued the *AtPOT1a* function (fig. 2*A*, lanes 3 and 4). *Brassica oleracea* (cauliflower) diverged from *A. thaliana* approximately 43 Ma (Beilstein et al. 2010), and BoPOT1a protein exhibits only 74% amino acid similarity to AtPOT1a overall. In contrast to AlPOT1a, *A. thaliana* mutants expressing BoPOT1a exhibited only 15% complementation efficiency relative to wild-type AtPOT1a (fig. 2*B*, lanes 3 and 4).

We next asked whether there is a complete separation-of-function between the Brassicaceae *POT1a* and *POT1b* lineages by testing whether overexpression of *AtPOT1b* or *BoPOT1b* from the strong constitutive 35S CaMV promoter can rescue *AtPOT1a* deficiency. No complementation was observed in either case (fig. 2*C*, lanes 3–6), consistent with previous functional analysis of *AtPOT1a* and *AtPOT1b* (Shakirov et al. 2005).

The NF model predicts that if the *POT1a* genes from Brassicaceae species acquired a novel function, a single-copy *POT1* gene from a non-Brassicaceae species would fail to complement *AtPOT1a* deficiency. As predicted, the single-copy *POT1* gene from cotton (*G. hirsutum*), which shared the last common ancestor with *A. thaliana* over 100 Ma (Beilstein et al. 2010), failed to complement *AtPOT1a* deficiency (fig. 2*C*, lanes 7 and 8). Finally, we asked whether the duplication of *POT1* genes in maize led to the evolution of functions similar to those associated with *AtPOT1a*. Overexpression of neither ZmPOT1a nor ZmPOT1b rescued the *pot1a* mutant (fig. 2*D*, lanes 3 and 4). Taken together, our phylogenetic data support the independence of the duplication events within grasses versus Brassicaceae, whereas genetic analyses indicate emergence of unique functions among all the POT1 lineages. These results also show that the ability of *POT1a* genes from *A. thaliana* relatives to rescue *POT1a* deficiency is correlated with phylogenetic distance.

Positive Selection along the POT1a Lineage

Given the functional differences between the duplicated *POT1* paralogs in *A. thaliana*, we hypothesized that either *POT1a* or *POT1b* had undergone adaptive evolution and that the corresponding amino acid substitutions would be correlated with functional diversification. To test these hypotheses we asked whether sites in the Brassicaceae *POT1a* or *POT1b* lineages experienced positive selection. Specifically, we examined the rate ratio ($\omega = dN/dS$) of nonsynonymous (dN) to synonymous (dS) changes along the branches leading to either *POT1a* or *POT1b*. We used the branch-sites test in PAML (Zhang et al. 2005) with the foreground branch represented by either the *POT1a* lineage or the *POT1b* lineage. Background branches consist of site classes whose ω values are not permitted to exceed 1. In contrast, the foreground branch contains an additional site classes ($\omega > 1$). This analysis resulted in a significant difference between the null model (background and foreground branches are evolving under the same rates) and the alternative model (some sites in the foreground branch have $\omega > 1$), $P = 0.00014$, when the POT1a branch, but not the POT1b branch was tested (table 1). Bayes Empirical Bayes (BEB) was used to calculate the posterior probability of sites coming from the site class with $\omega > 1$. Three sites with posterior probability > 0.90, E35 (BEB = 0.915), S212 (BEB = 0.921), and E293 (BEB = 0.902), were treated as having potentially adaptive roles in the function of POT1a and chosen for functional analysis (table 2). Each of these sites is located in a distinct protein domain: E35 in OB1, S212 in OB2, and E293 in the C-terminal region of POT1a protein (supplementary fig. S4, Supplementary Material online) (Lei et al. 2004; Theobald and Wuttke 2004; Trujillo et al. 2005; Croy et al. 2006). A fourth residue, L132 in OB1 domain is not predicted to be under positive selection and served as a negative control.

Sites of Positive Selection Are Required for AtPOT1a Function In Vivo

If the ancestor of *AtPOT1a* was indeed subjected to an extensive evolutionary sweep, substituting the identified positively selected amino acids with residues found in the *AtPOT1b* copy is expected to decrease the ability of AtPOT1a to complement the *pot1a* null mutant. We tested this prediction with AtPOT1 constructs bearing either the E35F, S212A, E293F or L123D mutations. Plants expressing the control L132D POT1a variant displayed over 90% complementation efficiency, confirming that L132 is not essential for POT1a function in this context (fig. 3). Conversely, the complementation efficiency dropped to 57% of wild-type level in transgenic plants expressing the E35F POT1a variant. S212L complementation efficiency was similarly reduced to below 50% (fig. 3). Notably, plants expressing the E293F POT1a variant exhibited 88% complementation efficiency (fig. 3), suggesting that either this residue has little effect on POT1a function under these conditions or both Glu and Phe function similarly in this context. Overall, our quantitative transgenic rescue experiments demonstrate that two of the three sites (E35 and S212) predicted to be under positive selection are indeed important for AtPOT1a function in vivo, whereas the third site E293 has a minimal effect on POT1a function. Moreover, our data further demonstrate that in silico predictions of positive selection sites require functional confirmation in vivo.

Sites of Positive Selection Affect the Binding Affinity of POT1a for CTC1 In Vitro

Recent work revealed that AtPOT1a physically interacts with CTC1 and STN1, core components of CST (CTC1/STN1/TEN1) complex (Renfrew et al. 2014). This interaction is specific, because AtPOT1a does not interact with TEN1. The CST complex is a heterotrimeric protein complex that plays a crucial role in modulating telomeric DNA replication at telomeres and likely other sites that bear repetitive DNA sequences (Casteel et al. 2009; Gu et al. 2012; Stewart et al. 2012). Notably, the single-copy human POT1 and the mouse POT1 paralog POT1b also interact with the corresponding CTC1 and STN1 proteins (Chen et al. 2012, 2013; Wu et al. 2012), suggesting that this interaction may be evolutionarily conserved. To investigate whether positive

Table 1. Likelihood Scores and Likelihood Ratio Test Results for Brassicaceae *POT1a* and *POT1b*[a].

Lineage	Log-Likelihood (ln *L*) Null Model	Log-Likelihood (ln *L*) Alternative Model	Likelihood Ratio Test ($2 \times \lvert \text{Ln l null} - \text{Ln l alternative} \rvert$), X^2 *P* Value
POT1a	−16,958.41	−16,951.01	14.8, $P = 0.00014$
POT1b	−16,959.68	−16,958.81	1.7, $P = 0.187$

[a]Likelihood scores and likelihood ratio test comparing the null model (no site classes with $\omega > 1$) with the alternative model (an additional site class where $\omega > 1$) of the branch-sites test implemented in the codeml package of PAML for the branch leading to either Brassicaceae *POT1a* or *POT1b*.

Table 2. Sites Predicted to be under Positive Selection Post *POT1* Duplication at the Base of Brassicaceae[a].

Amino Acid Position	POT1a Amino Acid	BEB Value
212	S	0.921
35	E	0.915
293	E	0.902
213	S	0.869
14	P	0.867
84	N	0.866
403	Y	0.804
131	C	0.801
405	W	0.794
299	M	0.767
186	S	0.763
26	L	0.753
393	E	0.739
328	T	0.734
290	Q	0.704
73	S	0.702
244	S	0.682
48	N	0.663
219	H	0.651
206	R	0.649
13	S	0.599
391	T	0.591
278	K	0.582
141	E	0.551
201	S	0.543
217	S	0.539
349	C	0.535
231	S	0.522

[a]Numbers in the left column are amino acid alignment positions and numbers in the far right column are BEB values generated in PAML. Letters in the center column are amino acids encoded in *POT1a* at the indicated alignment position. The three sites with BEB values exceeding the 0.90 threshold were chosen for further analysis.

selection within the POT1a lineage in Brassicaceae influences AtPOT1a interaction with CTC1 or STN1, we employed an in vitro biochemical approach. Recombinant POT1a, CTC1, and STN1 were expressed in rabbit reticulocyte lysate. Due to insolubility of full length CTC1, we used an amino terminal truncation construct CTC1ΔN previously shown to support POT1a–CTC1 protein interaction (Renfrew et al. 2014). Immunoprecipitation assays were conducted with recombinant RRL-expressed proteins using T7-tagged wild-type or mutant POT1a and [S^{35}] methionine-labeled CTC1ΔN. The ability of wild-type AtPOT1a or the positive selection variants to coprecipitate CTC1 was visualized and quantified from the amount of [^{35}S] methionine-labeled CTC1 that was pulled down (B, bound fraction) in the reaction (fig. 4). As expected, wild-type T7-tagged POT1a protein bound CTC1 efficiently. In contrast, all three positive selection site mutants had reduced affinity for CTC1 in vitro (fig. 4*A*), even E293F, which was largely able to rescue the *pot1a* null mutant in our in vivo complementation assay. Importantly, the L132D mutation did not perturb the POT1a–CTC1 interaction (fig. 4*B*), consistent with its ability to fully complement telomere length (fig. 3). We also examined AtPOT1b binding to CTC1ΔN. The affinity of POT1b for CTC1ΔN was significantly lower than for wild-type AtPOT1a, and similar to binding exhibited by the positive selection site mutants (fig. 4*C*). Thus, POT1a and not POT1b appears to be evolving an enhanced association with CTC1.

Finally, we asked whether the positive selection site mutations affected the interaction of POT1a with STN1. In marked contrast to the results obtained for CTC1ΔN, none of the amino acid substitutions significantly perturbed POT1a–STN1 binding in vitro (fig. 5). This observation provides additional confirmation that positive selection site mutations do not cause global protein misfolding or other types of aberrant conformational changes within AtPOT1a. Instead, these residues specifically affect the affinity of POT1a association with CTC1. Taken together, our data support the hypothesis that selective pressure on AtPOT1a specifically increases its binding affinity for CTC1, leading to a more robust telomere extension by telomerase.

Discussion

In this study, we investigated the role of POT1 in the maintenance of plant telomeres through a combination of evolutionary, genetic, and biochemical approaches. Our inquiries focused on the relatively recent duplication of *POT1* within the Brassicaceae and capitalized on advanced understanding of phylogenetic relationships within the plant kingdom and the wealth of functional data available for the telomerase enzyme and its regulation. By exploiting a sensitive genetic complementation assay, we addressed both the biological function and evolution of POT1. Our results not only reveal new insight into the molecular basis underlying positive selection at specific residues within *A. thaliana* POT1a, but also shed new light on the evolutionary fate of duplicated genes in the context of essential macromolecular complexes.

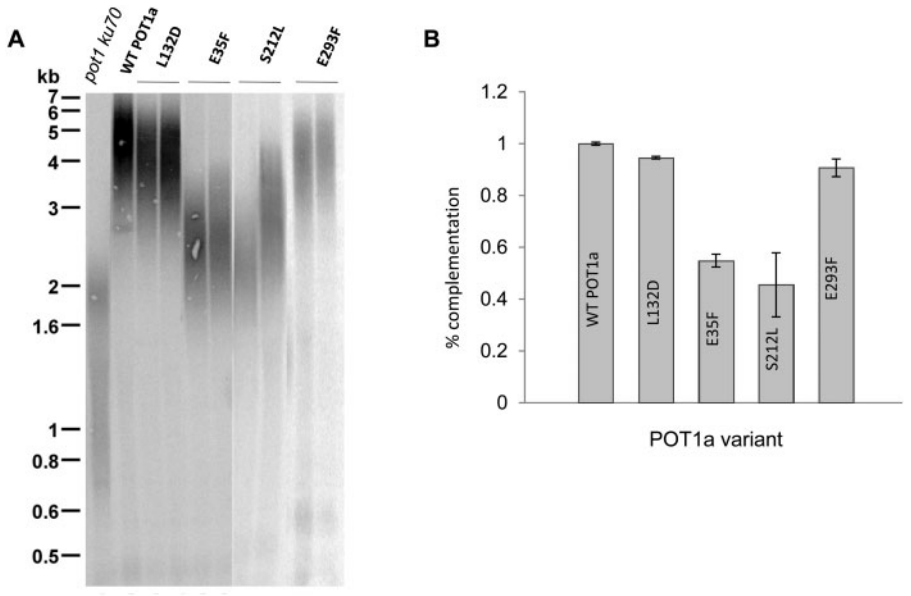

Fig. 3. Transformants expressing AtPOT1a variants with positively selected sites substituted by amino acids found in AtPOT1b show reduced ability to complement the *pot1a* null mutation. (*A*) TRF analysis showing telomere length for *pot1a ku70* mutant, *pot1a ku70* mutants expressing WT AtPOT1a, the L132D mutant (not under positive selection), and the three mutants of positively selected sites: E35F, S212L, E293F. (*B*) Quantification of telomere length in (*A*). Error bars represent standard error from three independently transformed lines of each genotype.

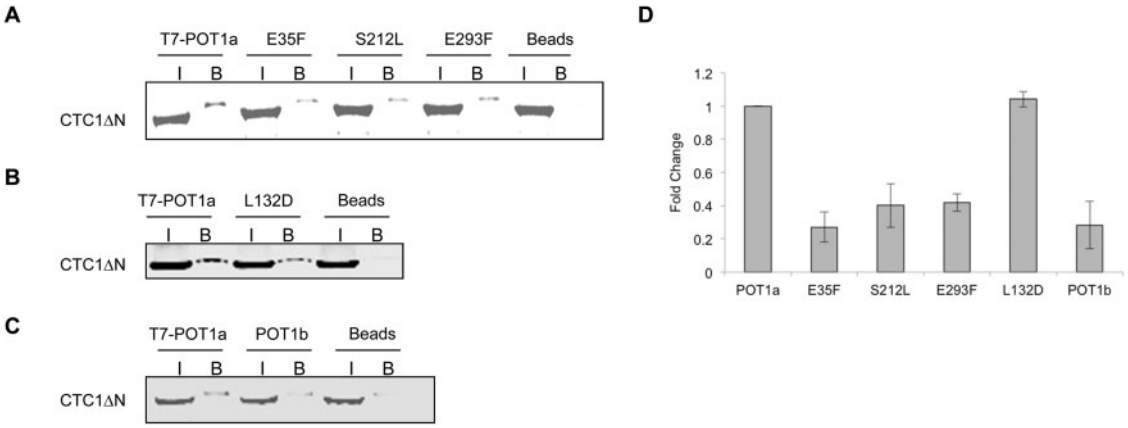

Fig. 4. Positive selection site mutations reduce AtPOT1a affinity for CTC1. (*A*) In vitro co-IP assays with recombinant RRL-expressed proteins were performed using T7-tagged AtPOT1a, positive selection site AtPOT1a mutants, AtPOT1b and [S^{35}] methionine-labeled AtCTC1ΔN. Control reactions for nonspecific binding of the [35S]Met-labeled CTC1ΔN protein to the beads were performed in the absence of T7-tagged proteins (beads control). (I) denotes protein input, (B) indicates bound protein. (*B*) Quantification of binding relative to wild-type POT1a protein. Binding was calculated as a fraction of B/B + I followed by subtraction of the background signal (beads) and normalized to wild-type T7-POT1a signal. Error bars represent standard error of the mean from three independent pull-down experiments.

Retention of POT1 Duplicates Is Rare among Land Plants

Despite multiple ancient whole-genome polyploidization events in vertebrates, yeast, and plants (Kellis et al. 2004; Dehal and Boore 2005; Cui et al. 2006), *POT1* remains single-copy in most eukaryotic genomes, including most plants. Among the 21 representatives of the green plant lineage we surveyed, only two instances of independent *POT1* gene duplication were detected, one in Brassicaceae and one in the Panicoideae subfamily of grasses (Poaceae). Given the typically large number of gene family members in plants (Guo 2013), including the well-documented increase in genes encoding putative double-strand telomere-binding proteins in *A. thaliana* (Karamysheva et al. 2004; Shakirov and Shippen 2012; Nelson et al. 2014), our data suggest that *POT1* duplication is not well tolerated. This finding argues that additional copies of *POT1* are most often lost following duplication, a pattern observed for components of the CST complex (Shakirov and Shippen 2012; Nelson et al. 2014). Thus, the duplication in *POT1* that characterizes genomes in the Brassicaceae appears to be an unusual event in *POT1* evolution.

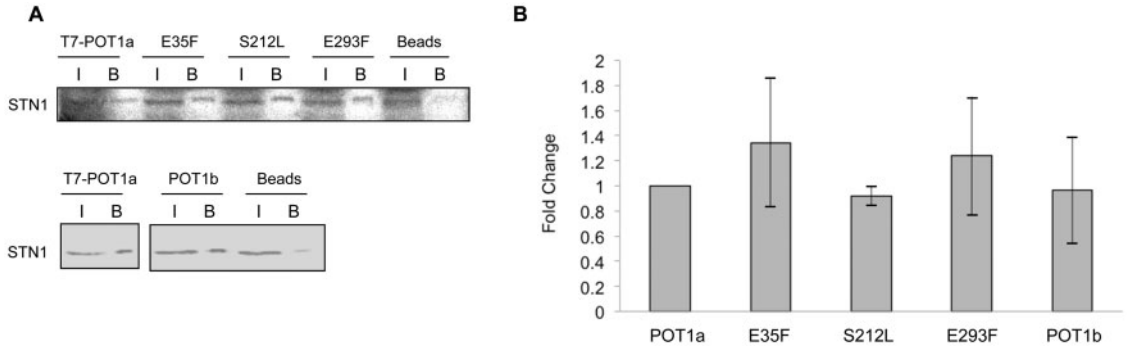

Fig. 5. Positive selection site mutations do not alter AtPOT1a affinity for STN1. (*A*) In vitro co-IP assays with recombinant RRL-expressed proteins were performed using T7-tagged AtPOT1a, positive selection site AtPOT1a mutants, AtPOT1b and [S^{35}] methionine-labeled AtSTN1. The beads control contained no T7-tagged POT1 proteins. (I) denotes protein input, (B) indicates bound protein. (*B*) Quantification of binding relative to wild-type POT1a protein. Binding was calculated as a fraction of B/B + I followed by subtraction of the background signal (beads) and normalized to wild-type POT1a signal. Error bars represent standard error of the mean from three independent pull-down reactions.

POT1a Divergence Following POT1 Duplication

The sensitive genetic complementation system we developed allowed us to obtain quantitative data on the degree of functional conservation of POT1 proteins among different plant species. Considering the dramatic functional divergence in the two *A. thaliana* POT1 paralogs, we first investigated the extent to which these functions are conserved within a single-copy *POT1* gene by performing cross-species complementation experiments. The single-copy *POT1* ortholog from *G. hirsutum* (cotton) did not complement the *pot1a* null mutation in *A. thaliana*. One possible explanation is that POT1 proteins from non-Brassicaceae species cannot interact with a key binding partner of AtPOT1a. A similar model has been proposed for the mouse POT1 paralogs (Hockemeyer et al. 2007). Expression of a single-copy human *POT1* fails to complement mouse *POT1a* or *POT1b* deficiencies, but coexpression of human POT1 with its interacting partner TPP1 fully restores mouse *POT1a* function. Mouse *POT1b* deficiency is not rescued with these same constructs, indicating that mouse POT1b evolved a novel function or interacts with a different binding partner (Hockemeyer et al. 2006).

To better understand divergence within Brassicaceae *POT1a* sequences postduplication, we examined *POT1a* orthologs from relatives of *A. thaliana* at varying phylogenetic distances. *Arabidopsis lyrata* POT1a fully complemented the telomere length defect of *the A. thaliana pot1a* mutant, whereas the *B. oleracea* (cauliflower) POT1a showed only 15% complementation efficiency. Hence, Brassicaceae POT1a proteins have continued to diverge following the duplication event, and that similar to our observations in cotton, the ability of BoPOT1a to bind other partners in the complex may be altered. Regardless of mechanism, the capacity of *POT1a* orthologs to complement the *A. thaliana* null mutation is correlated with phylogenetic distance.

As expected based on previous biochemical and genetic analysis of AtPOT1b, overexpression of this gene did not rescue the *pot1a* mutant. This result is consistent with the rapid and concerted sequence divergence observed in our analyses of molecular evolution. Notably, neither of the two POT1 proteins from *Z. mays* could rescue the *A. thaliana pot1a* mutant, suggesting that evolution of duplicated *POT1* genes may have proceeded differently in different plant lineages.

AtPOT1a Shares Some Functional Domains with Yeast and Vertebrate POT1 and Retains Some of the Ancestral Functions

The function of POT1 in plants is known only from *A. thaliana* and *P. patens*, and at first glance, the difference between these proteins appears to be dramatic. *Arabidopsis thaliana* POT1 paralogs assemble into telomerase RNP complexes (Cifuentes-Rojas et al. 2011, 2012), whereas the moss POT1 protects chromosome ends from nucleolytic attack and inappropriate recombination (Shakirov et al. 2010). Despite these remarkable distinctions, our data also reveal substantial conservation of POT1 function. First, mutation of F65, a residue in human and yeast POT1 predicted to be critical for DNA binding based on the crystal structure of OB1 (Lei et al. 2003, 2004), diminishes telomere elongation in *A. thaliana pot1a* mutants to only 12% of the wild-type level. This observation is intriguing given that *A. thaliana* POT1 proteins bind telomerase RNA instead of single-stranded telomeric DNA (Cifuentes-Rojas et al. 2011, 2012) and implies that the nucleic acid-binding interface is conserved. Second, we uncovered an essential role for the extreme C-terminus of AtPOT1a. This motif is critical for *S. pombe* POT1 function (Bunch et al. 2005) and for human POT1 interaction with its binding partner, TPP1 (Liu et al. 2004; Ye et al. 2004). Altogether, our genetic analysis indicates that a subset of core ancestral functions within *A. thaliana* POT1a are retained despite the evolutionary distance separating fungal, plant, and animal lineages.

Positive Selection within the POT1a Lineage Enhances the AtPOT1a Interaction with CTC1

The identification of a large number of plant *POT1* sequences also provided an opportunity to search for evidence of selective pressure on the duplicated POT1 proteins of Brassicaceae. BEB implemented in the branch-sites test of PAML identified a number of positively selected sites in *POT1a*, which were

prioritized on the basis of BEB-assigned statistical values. Three sites with the highest ω values were subjected to functional genetic and biochemical analysis. Remarkably, we discovered that two of these sites are critical for POT1a function in vivo, and all three specifically increase the POT1a interaction with CTC1 in vitro.

E35 is found in OB1, and its importance is illustrated by a 2-fold reduction in complementation efficiency in plants expressing the E35F POT1a variant. This mutation also reduced POT1a binding to CTC1 in vitro by approximately 70%. Structural modeling of AtPOT1a suggests that E35 is unlikely to play a direct role in DNA binding, but it does lie in the vicinity of DNA-binding pocket. Thus, E35 may be involved in protein–protein interactions and higher order complex assembly. A second site with a strong signature of positive selection, S212, is found in the OB2 domain, which is immediately adjacent to OB1 and is proposed to reinforce nucleic acid binding and protein interaction activities of POT1 proteins (Trujillo et al. 2005; Croy and Wuttke 2006). As with E35, our data indicate that S212 is important for AtPOT1a function in vivo and for CTC1 interaction in vitro.

The third site E293 lies within the C-terminal domain, a region of POT1 that has not been structurally characterized. Although the E293F mutation does not significantly reduce AtPOT1a function in vivo, it does diminish AtPOT1a affinity for CTC1 to the same extent as the other two positive selection site mutations. Why E293F mutation does not affect telomere length maintenance in vivo is unclear, but it is possible that in the context of the entire TER1/POT1a RNP, the diminished POT1a–CTC1 interaction in the E293F mutant is masked or compensated by other interacting RNP subunits. Importantly, mutations of all three positively selected sites reduce POT1a–CTC1 interaction down to the POT1b level, and yet do not affect AtPOT1a interaction with other CST members such as STN1. Thus, our findings support the conclusion that the selective evolutionary pressure on AtPOT1a acts specifically to enhance its binding to CTC1. A model depicting the interaction of AtPOT1a in *A. thaliana* is shown in figure 6.

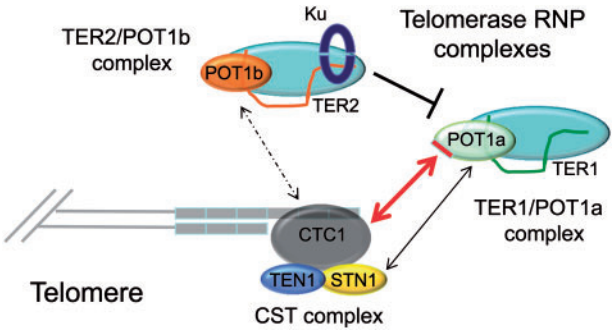

Fig. 6. Positive selection enhances the POT1a-binding interface with CTC1 in *Arabidopsis thaliana*. *Arabidopsis thaliana* harbors two alternative telomerase RNP complexes containing the RNA subunits TER1 or TER2. TER1 is associated with the POT1a protein, one of two POT1 paralogs, whereas TER2 is associated with Ku and POT1b. The TER1/POT1a complex is responsible for telomere maintenance. It engages the telomere through POT1a interaction with two components of the telomeric CST complex, CTC1 and STN1. In this study, we show that the CTC1-binding interface on POT1a is under positive selection. Although POT1b exhibits a basal level of affinity toward CTC1, selective pressure enhances binding of POT1a to CTC1. The increased affinity of POT1a for CTC1 may favor TER1/POT1a RNP recruitment to telomeres over the TER2/POT1b complex, thereby promoting telomere synthesis on chromosome ends.

Functional Evolution of POT1 Genes in A. thaliana

Essential components of the *A. thaliana* telomere complex have been shaped by gene duplication events, and have resulted in the emergence of at least two distinct telomerase RNP particles with partially opposing functions (fig. 6). Each RNP is defined by the presence of a unique TER subunit (AtTER1 or AtTER2). The *A. thaliana* TER genes, like *POT1*, are the products of gene duplication. However, in this case the event was very recent as all other Brassicaceae have a single locus of orthology with AtTER1 and AtTER2 (Beilstein et al. 2012). Because the AtTER1 RNP associates with POT1a whereas the AtTER2 RNP assembles with POT1b, the duplication of *POT1* in Brassicaceae produced the necessary components for two alternative telomerase RNPs much earlier in evolutionary history than is suggested by the later duplication in TER. Thus, it is plausible that post *POT1* duplication, alternative RNPs were specified and subsequently evolved opposing functions.

Whether retention of *POT1a* and *POT1b* paralogs in Brassicaceae is more consistent with a model of NF or EAC depends on the function of POT1 in species with a single copy. In the NF model for the retention of duplicated genes, one gene copy acquires a novel, beneficial function, whereas the other copy retains most of the ancestral gene functions (Ohno 1970; Lynch and Conery 2000). At the molecular level, the gene copy that gained a novel function is expected to undergo a period of positive Darwinian evolution following the duplication event that marks its origin (Lynch and Conery 2000). As *POT1a* experienced positive selection postduplication, we might expect that the AtPOT1a association with telomerase (Surovtseva et al. 2007; Cifuentes-Rojas et al. 2011), its role in telomere length homeostasis (Shakirov et al. 2005), and its interaction with CTC1 (Renfrew et al. 2014) are functions that were acquired postduplication.

However, several of our observations lie in conflict with the NF model. For example, the ability of both AtPOT1a and AtPOT1b to assemble into telomerase RNP complexes suggests that the ancestral single-copy POT1 protein played a similar role. Moreover, AtPOT1a is a member of the canonical telomerase RNP complex required for enzyme function in vivo (Surovtseva et al. 2007), arguing that this POT1–CTC1 interaction may be ancestral given that AtPOT1a, AtPOT1b, and POT1 from vertebrates associate with CTC1 (Chen et al. 2012, 2013; Wu et al. 2012; Renfrew et al. 2014; this study). Therefore, we postulate that the CTC1–POT1 interaction is conserved between plants and animals, and that positive selection in *POT1a* resulted in an enhancement of CTC1 binding in comparison to POT1b (fig. 6).

If we reject the NF model, why are *POT1a* and *POT1b* retained postduplication? In EAC, conflicting functions of a single-copy gene are resolved upon duplication as each descendant copy is free to specialize. Positive selection is expected in both copies as they evolve toward an optimum no longer constrained by the necessity to perform the conflicting role (Des Marais and Rausher 2008). Here, we find no evidence for positive selection in POT1b, raising doubt that EAC adequately describes retention. However, unlike the proposed theoretical models, POT1a and POT1b are members of multisubunit RNP complexes with different constituents (fig. 6). AtPOT1a uniquely associates with the TER1 RNP to negatively influence telomerase enzyme activity. In contrast, AtPOT1b uniquely associates with TER2 and Ku, which are both implicated in the negative regulation of telomerase and telomere length, respectively (Riha et al. 2002; Cifuentes-Rojas et al. 2012). It is therefore plausible that an ancestral single-copy POT1 assembled into two alternative telomerase RNPs with opposing functions that are defined by different accessory proteins. In this scenario, both the stimulation of telomerase activity through CTC1 interaction and the negative regulation of telomerase by association with Ku could have been ancestral functions of a single-copy POT1 protein. Retention was favored following duplication to reduce functional conflict. Positive selection need only affect one copy in order to reinforce one of the two functions, in this case enhancing the interaction of AtPOT1a–CTC1. Furthermore, if AtPOT1b did not convey the properties of negative regulation, it would not be subjected to selective pressure.

Whether retention of *POT1a* and *POT1b* is best explained by NF or EAC requires additional functional studies in species with a single *POT1* gene, such as the genetically amenable moss *P. patens*. In the absence of such data, we favor EAC for the reasons discussed. Taken together, our data indicate that signatures of selection resulting from tests of molecular evolution may not always be consistent with expectations from theory. The observations highlight the power of a combined interdisciplinary approach to extend in silico predictions of molecular evolution toward understanding molecular and biochemical mechanisms in vivo.

Materials and Methods

Recovery of POT1 Orthologs from Sampled Species

POT1 BLAST searches of the plant genomes were performed using the BLASTp or tBLASTn options available at the corresponding genome portals (http://www.phytozome.net/; last accessed October 15, 2014) with *A. thaliana* POT1 proteins as a query. BLAST searches with human or *S. pombe* POT1 proteins were also attempted, but did not improve the outcome. For species in Brassicaceae lacking whole-genome sequence, we designed degenerate primers to amplify either the *POT1a* or *POT1b* paralog using hiTAIL-PCR (Liu and Chen 2007). Sequences retrieved by BLAST searches were annotated as *POT1* homologs, whereas *POT1* sequences resulting from amplification were deposited in GenBank. Accession numbers are provided in supplementary table S1, Supplementary Material online.

Phylogenetic and Positive Selection Analyses

Translated amino acid sequences were aligned using ClustalW (Larkin et al. 2007). The alignment was then used to correct the nucleotide alignment using MacClade Vers. 4.08 (Maddison DR and Maddison WP 2005), and subsequent alignment corrections were made by eye. Phylogenetic trees were inferred using RAxML v7.0.4 (Stamatakis 2006). The *POT1* land plant tree was generated from an amino acid alignment under the PROTGAMMAWAG model of protein evolution, which permits Γ distributed rate heterogeneity among sites and uses the WAG (Whelan and Goldman) amino acid transition matrix (Whelan and Goldman 2001). The Eurosid *POT1* nucleotide alignment was used to infer phylogeny under the GTRGAMMA model of sequence evolution, which employs the general time reversible model of sequence evolution with Γ distributed rate heterogeneity.

Using PAML4.0 (Yang 1997), we implemented the branch-site model A test with the foreground branch represented by either the *POT1a* lineage or the *POT1b* lineage. In the null model of the branch-site test, both the background and foreground branches consist of sites where $0 < \omega < 1$ or $\omega = 1$ (Zhang et al. 2005). In the alternative model, the designated foreground branch is permitted additional site classes in which $\omega > 1$ (Zhang et al. 2005).

Plant Growth and Transformation Procedures

Arabidopsis thaliana seeds were cold treated overnight at 4 °C, and then placed in an environmental growth chamber and grown under a 16-h light/8-h dark photoperiod at 23 °C. *pot1a-1*, *ku70*, and *pot1a-1*$^{-/-}$ *ku70*$^{-/-}$ mutants were described previously (Riha et al. 2002; Surovtseva et al. 2007). For genetic complementation experiments, *POT1* cDNAs were subcloned into the pCBK05 binary vector carrying the *bar* gene as a selectable marker (Riha et al. 2002) under the control of *AtPOT1a* native promoter (a 1.5-kb region immediately upstream of the start codon) or the CaMV35S promoter for overexpression. Complementation constructs were introduced into the *Agrobacterium tumefaciens* GV3101 strain, which was used to transform *pot1a-1*$^{-/-}$ *ku70*$^{+/-}$ plants by the modified in planta method (Bechtold and Pelletier 1998). T1 primary transformants were selected on 0.5 Murashige and Skoog basal medium supplemented with 25 mg/l of phosphinothricine (BASTA) (Crescent Chemical, Islandia, NY) and genotyped by PCR to identify homozygous *pot1a-1 ku70* knock-out plants harboring the transgene. PCR genotyping was also used to identify their siblings without the transgene. RT-PCR and Western blotting with FLAG antibody (Pierce) were used to verify transgene expression in the transformants.

Telomere Length Analysis and Quantification

DNA from individual whole plants was extracted as described (Cocciolone and Cone 1993). TRF analysis was performed with DNA digested with *Tru*1I (Fermentas, Hanover, MD) restriction enzyme. ^{32}P 5′-end-labeled $(T_3AG_3)_4$ oligonucleotide was used as a probe (Fitzgerald et al. 1999). Radioactive signals were scanned by a Storm PhosphorImager (Molecular

Dynamics, Sunnyvale, CA), and the data were analyzed by IMAGEQUANT software (Molecular Dynamics). The average telomere length (L) was measured using Telometric-1.2 program (Grant et al. 2001). The average telomere lengths of untransformed *pot1a ku70* mutants, transformants expressing wild-type *AtPOT1a*, and other *POT1* constructs were designated as L_0, L_1, and L_x, respectively. We set the complementation level of wild-type *AtPOT1a* transformants (positive control) as one, and that of untransformed *pot1a ku70* mutants (negative control) as zero. The complementation efficiency (E) of each *POT1* construct was calculated as: $E = (L_x - L_0)/(L_1 - L_0)$*100%. At least three individual transformant lines for each construct were analyzed for statistical support.

POT1a Protein Interactions

Proteins were expressed from Rabbit Reticulocyte Lysate according to the manufacturer's instructions (Promega). Proteins were expressed with a T7 tag from the pET28a vector (Novagen) or without a tag from the pCITE4a vector (Novagen). Proteins without a tag were translated with [^{35}S] methionine to detect the amount of coprecipitated protein. Expression of T7-tagged proteins was verified by labeling an aliquot of the expression master mix with [^{35}S] methionine. The coimmunoprecipitation procedure was carried out as described (Renfrew et al. 2014). Binding quantification was performed using Quantity One software (Bio-Rad). The bound signal was represented as a fraction of the bound plus unbound total signal followed by subtracting the amount of background noise (beads control). Samples were then normalized to wild-type POT1a binding to CTC1 or STN1. Error bars represent the standard of the mean from three independent replicates.

Acknowledgments

The authors are grateful to members of the Shippen lab for insightful discussions. This work was supported by National Institute of Health (grant number R01 GM065383 to D.E.S. and National Research Service Award number F32GM09363 to M.A.B.); start up funds from Purdue to M.J.Z. and funds from the Russian Government Program of Competitive Growth of Kazan Federal University to E.V.S.

Supplementary Material

Supplementary tables S1 and figures S1–S4 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

References

Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815.

Baumann P, Cech TR. 2001. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* 292:1171–1175.

Baumann P, Podell E, Cech TR. 2002. Human Pot1 (protection of telomeres) protein: cytolocalization, gene structure, and alternative splicing. *Mol Cell Biol.* 22:8079–8087.

Bechtold N, Pelletier G. 1998. In planta Agrobacterium-mediated transformation of adult *Arabidopsis thaliana* plants by vacuum infiltration. *Methods Mol Biol.* 82:259–266.

Beilstein MA, Brinegar AE, Shippen DE. 2012. Evolution of the *Arabidopsis* telomerase RNA. *Front Genet.* 3:188.

Beilstein MA, Nagalingum NS, Clements MD, Manchester SR, Mathews S. 2010. Dated molecular phylogenies indicate a Miocene origin for *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A.* 107:18724–18728.

Birchler JA, Veitia RA. 2007. The gene balance hypothesis: from classical genetics to modern genomics. *Plant Cell* 19:395–402.

Blanc G, Wolfe KH. 2004. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell* 16:1667–1678.

Bunch JT, Bae NS, Leonardi J, Baumann P. 2005. Distinct requirements for Pot1 in limiting telomere length and maintaining chromosome stability. *Mol Cell Biol.* 25:5567–5578.

Casteel DE, Zhuang S, Zeng Y, Perrino FW, Boss GR, Goulian M, Pilz RB. 2009. A DNA polymerase-alpha primase cofactor with homology to replication protein A-32 regulates DNA replication in mammalian cells. *J Biol Chem.* 284:5807–5818.

Chen LY, Majerska J, Lingner J. 2013. Molecular basis of telomere syndrome caused by CTC1 mutations. *Genes Dev.* 27:2099–2108.

Chen LY, Redon S, Lingner J. 2012. The human CST complex is a terminator of telomerase activity. *Nature* 488:540–544.

Cifuentes-Rojas C, Kannan K, Tseng L, Shippen DE. 2011. Two RNA subunits and POT1a are components of *Arabidopsis* telomerase. *Proc Natl Acad Sci U S A.* 108:73–78.

Cifuentes-Rojas C, Nelson AD, Boltz KA, Kannan K, She X, Shippen DE. 2012. An alternative telomerase RNA in *Arabidopsis* modulates enzyme activity in response to DNA damage. *Genes Dev.* 26: 2512–2523.

Cocciolone SM, Cone KC. 1993. Pl-Bh, an anthocyanin regulatory gene of maize that leads to variegated pigmentation. *Genetics* 135: 575–588.

Croy JE, Podell ER, Wuttke DS. 2006. A new model for *Schizosaccharomyces pombe* telomere recognition: the telomeric single-stranded DNA-binding activity of Pot11-389. *J Mol Biol.* 361: 80–93.

Croy JE, Wuttke DS. 2006. Themes in ssDNA recognition by telomere-end protection proteins. *Trends Biochem Sci.* 31: 516–525.

Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE, Doyle JJ, Soltis PS, Carlson JE, Arumuganathan K, Barakat A, et al. 2006. Widespread genome duplications throughout the history of flowering plants. *Genome Res.* 16:738–749.

Dehal P, Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol.* 3:e314.

Des Marais DL, Rausher MD. 2008. Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature* 454: 762–765.

Fawcett JA, Maere S, Van de Peer Y. 2009. Plants with double genomes might have had a better chance to survive the Cretaceous-Tertiary extinction event. *Proc Natl Acad Sci U S A.* 106:5737–5742.

Fitzgerald MS, Riha K, Gao F, Ren S, McKnight TD, Shippen DE. 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 U S A.* 96:14813–14818.

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151:1531–1545.

Grant JD, Broccoli D, Muquit M, Manion FJ, Tisdall J, Ochs MF. 2001. Telometric: a tool providing simplified, reproducible measurements of telomeric DNA from constant field agarose gels. *Biotechniques* 31: 1314–1316, 1318.

Gu P, Min JN, Wang Y, Huang C, Peng T, Chai W, Chang S. 2012. CTC1 deletion results in defective telomere replication, leading to

catastrophic telomere loss and stem cell exhaustion. *EMBO J.* 31: 2309–2321.
Guo YL. 2013. Gene family evolution in green plants with emphasis on the origination and evolution of *Arabidopsis thaliana* genes. *Plant J.* 73:941–951.
He H, Multani AS, Cosme-Blanco W, Tahara H, Ma J, Pathak S, Deng Y, Chang S. 2006. POT1b protects telomeres from end-to-end chromosomal fusions and aberrant homologous recombination. *EMBO J.* 25:5180–5190.
Hockemeyer D, Daniels JP, Takai H, de Lange T. 2006. Recent expansion of the telomeric complex in rodents: two distinct POT1 proteins protect mouse telomeres. *Cell* 126:63–77.
Hockemeyer D, Palm W, Else T, Daniels JP, Takai KK, Ye JZ, Keegan CE, de Lange T, Hammer GD. 2007. Telomere protection by mammalian Pot1 requires interaction with Tpp1. *Nat Struct Mol Biol.* 14:754–761.
Innan H, Kondrashov F. 2010. The evolution of gene duplications: classifying and distinguishing between models. *Nat Rev Genet.* 11: 97–108.
Jacob NK, Lescasse R, Linger BR, Price CM. 2007. Tetrahymena POT1a regulates telomere length and prevents activation of a cell cycle checkpoint. *Mol Cell Biol.* 27:1592–1601.
Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, et al. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* 473:97–100.
Karamysheva ZN, Surovtseva YV, Vespa L, Shakirov EV, Shippen DE. 2004. A C-terminal Myb extension domain defines a novel family of double-strand telomeric DNA-binding proteins in *Arabidopsis*. *J Biol Chem.* 279:47799–47807.
Kellis M, Birren BW, Lander ES. 2004. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* 428:617–624.
Kellogg EA. 1998. Relationships of cereal crops and other grasses. *Proc Natl Acad Sci U S A.* 95:2005–2010.
Kondrashov FA, Rogozin IB, Wolf YI, Koonin EV. 2002. Selection in the evolution of gene duplications. *Genome Biol.* 3:RESEARCH0008.
Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.
Lei M, Podell ER, Baumann P, Cech TR. 2003. DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. *Nature* 426:198–203.
Lei M, Podell ER, Cech TR. 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.
Linger BR, Price CM. 2009. Conservation of telomere protein complexes: shuffling through evolution. *Crit Rev Biochem Mol Biol.* 44:434–446.
Liu D, Safari A, O'Connor MS, Chan DW, Laegeler A, Qin J, Songyang Z. 2004. PTOP interacts with POT1 and regulates its localization to telomeres. *Nat Cell Biol.* 6:673–680.
Liu YG, Chen Y. 2007. High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *Biotechniques* 43:649–656.
Lue NF. 2010. Plasticity of telomere maintenance mechanisms in yeast. *Trends Biochem Sci.* 35:8–17.
Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. *Science* 290:1151–1155.
Maddison DR, Maddison WP. 2005. MacClade: analysis of phylogeny and character evolution Ver. 4.08. Sunderland (MA): Sinauer Associates.
Nelson AD, Forsythe ES, Gan X, Tsiantis M, Beilstein MA. 2014. Extending the model of *Arabidopsis* telomere length and composition across Brassicaceae. *Chromosome Res.* 22:153–166.
Ohno S. 1970. Evolution by gene duplication. New York: Springer-Verlag 160p.
Paterson AH, Bowers JE, Chapman BA. 2004. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc Natl Acad Sci U S A.* 101: 9903–9908.
Raices M, Verdun RE, Compton SA, Haggblom CI, Griffith JD, Dillin A, Karlseder J. 2008. *C. elegans* telomeres contain G-strand and C-strand overhangs that are bound by distinct proteins. *Cell* 132: 745–757.
Renfrew KB, Song X, Lee JR, Arora A, Shippen DE. 2014. POT1a and components of CST engage telomerase and regulate its activity in *Arabidopsis*. *PLoS Genet.* 10:e1004738.
Richards EJ, Ausubel FM. 1988. Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* 53:127–136.
Riha K, Watson JM, Parkey J, Shippen DE. 2002. Telomere length deregulation and enhanced sensitivity to genotoxic stress in *Arabidopsis* mutants deficient in Ku70. *EMBO J.* 21:2819–2826.
Shakirov EV, McKnight TD, Shippen DE. 2009. POT1-independent single-strand telomeric DNA binding activities in Brassicaceae. *Plant J.* 58:1004–1015.
Shakirov EV, Perroud PF, Nelson AD, Cannell ME, Quatrano RS, Shippen DE. 2010. Protection of Telomeres 1 is required for telomere integrity in the moss *Physcomitrella patens*. *Plant Cell* 22:1838–1848.
Shakirov EV, Shippen DE. 2004. Length regulation and dynamics of individual telomere tracts in wild-type *Arabidopsis*. *Plant Cell* 16: 1959–1967.
Shakirov EV, Shippen DE. 2012. *Selaginella moellendorffii* telomeres: conserved and unique features in an ancient land plant lineage. *Front Plant Sci.* 3:161.
Shakirov EV, Song X, Joseph JA, Shippen DE. 2009. POT1 proteins in green algae and land plants: DNA-binding properties and evidence of co-evolution with telomeric DNA. *Nucleic Acids Res.* 37: 7455–7467.
Shakirov EV, Surovtseva YV, Osbun N, Shippen DE. 2005. The *Arabidopsis* Pot1 and Pot2 proteins function in telomere length homeostasis and chromosome end protection. *Mol Cell Biol.* 25: 7725–7733.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
Stewart JA, Wang F, Chaiken MF, Kasbek C, Chastain PD 2nd, Wright WE, Price CM. 2012. Human CST promotes telomere duplex replication and general replication restart after fork stalling. *EMBO J.* 31: 3537–3549.
Surovtseva YV, Shakirov EV, Vespa L, Osbun N, Song X, Shippen DE. 2007. *Arabidopsis* POT1 associates with the telomerase RNP and is required for telomere maintenance. *EMBO J.* 26: 3653–3661.
Theobald DL, Wuttke DS. 2004. Prediction of multiple tandem OB-fold domains in telomere end-binding proteins Pot1 and Cdc13. *Structure* 12:1877–1879.
Trujillo KM, Bunch JT, 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.
Vandepoele K, Saeys Y, Simillion C, Raes J, Van De Peer Y. 2002. The automatic detection of homologous regions (ADHoRe) and its application to microcolinearity between *Arabidopsis* and rice. *Genome Res.* 12:1792–1801.
Wang F, Podell ER, Zaug AJ, Yang Y, Baciu P, Cech TR, Lei M. 2007. The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* 445:506–510.
Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol.* 18: 691–699.
Wu L, Multani AS, He H, Cosme-Blanco W, Deng Y, Deng JM, Bachilo O, Pathak S, Tahara H, Bailey SM, et al. 2006. Pot1 deficiency initiates

DNA damage checkpoint activation and aberrant homologous recombination at telomeres. *Cell* 126:49–62.

Wu P, Takai H, de Lange T. 2012. Telomeric 3′ overhangs derive from resection by Exo1 and Apollo and fill-in by POT1b-associated CST. *Cell* 150:39–52.

Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci.* 13:555–556.

Ye JZ, Hockemeyer D, Krutchinsky AN, Loayza D, Hooper SM, Chait BT, 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.

Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol.* 22:2472–2479.