Structurally distinct telomere-binding proteins in *Ustilago maydis* execute non-overlapping functions in telomere replication, recombination, and protection

Eun Young Yu¹, Syed S. Zahid¹, Swapna Ganduri¹, Jeanette H. Sutherland¹, Min Hsu¹, William K. Holloman¹ & Neal F. Lue¹,²

Duplex telomere binding proteins exhibit considerable structural and functional diversity in fungi. Herein we interrogate the activities and functions of two Myb-containing, duplex telomere repeat-binding factors in *Ustilago maydis*, a basidiomycete that is evolutionarily distant from the standard fungi. These two telomere-binding proteins, *UmTay1* and *UmTrf2*, despite having distinct domain structures, exhibit comparable affinities and sequence specificity for the canonical telomere repeats. *UmTay1* specializes in promoting telomere replication and an ALT-like pathway, most likely by modulating the helicase activity of Blm. *UmTrf2*, in contrast, is critical for telomere protection; transcriptional repression of *Umtrf2* leads to severe growth defects and profound telomere aberrations. Comparative analysis of *UmTay1* homologs in different phyla reveals broad functional diversity for this protein family and provides a case study for how DNA-binding proteins can acquire and lose functions at various chromosomal locations. Our findings also point to stimulatory effect of telomere protein on ALT in *Ustilago maydis* that may be conserved in other systems.
Eukaryotic chromosome ends, or telomeres, harbor repetitive DNA sequences that mediate the assembly of a special, dynamic nucleoprotein structure that is crucial for genome stability. This dynamic structure stabilizes chromosome ends by suppressing aberrant repair events at the termini, and by promoting the retention and replenishment of telomere DNA through successive rounds of replication. Special mechanisms are required to retain and replenish telomere DNA owing to (i) the propensity of telomere DNA to form barriers that block replication forks; and (ii) the “end-replication” problem that prevents the complete synthesis of lagging strand duplex. Both the protective and DNA maintenance functions of telomeres are mediated by specific proteins within the dynamic assembly, often through direct physical interactions between telomere proteins and functionally relevant downstream targets.

Within the dynamic telomere nucleoprotein assembly, duplex telomere repeat-binding factors are known to execute crucial functions including telomere length regulation, telomere protection and telomere replication. Studies in diverse organisms have revealed considerable structural and functional variabilities in these duplex-binding factors. In mammals, the two major duplex telomere-binding proteins, Telomerase Repeat-Binding Factor 1 and 2 (TRF1 and TRF2), are each part of the shelterin complex that collectively protects chromosome ends and suppresses the DNA damage response. While TRF1 and TRF2 are structurally similar in having an N-terminal TRF homology dimerization domain (TRFH) and a C-terminal Myc domain, each protein serves distinct functions in telomere regulation. TRF2, in particular, plays a key role in telomere protection by inhibiting telomere fusions and the DNA damage response. TRF1, on the other hand, promotes efficient telomere replication by recruiting BLM helicase to unwind replication barriers. (TRF2 also physically interacts with BLM, but the functional significance of this is unclear.) In the fission yeast Schizosaccharomyces pombe, a single TRF1/2-like protein named Tay1 alone performs multiple functions in telomere protection and replication, although it is unclear how the replication function is executed. Interestingly, S. pombe contains two other Myb-containing proteins capable of binding telomere repeat-like sequences. One of these, named Tbf1, resembles mammalian TRF1/2 according to sequence alignment in the TRFH and Myb domains (in fact, more so than does Tay1), but exhibits no compelling telomere functions. The other protein, named SpTeb1, bears two consecutive Myb motifs away from the C-terminus and evidently functions primarily as a transcription factor. Interestingly, the SpTeb1 homolog in the budding yeast Yarrowia lipolytica (named Tay1, a well-studied member of this protein family) is reported to be a key telomere-binding protein and plays a critical role in telomere maintenance and protection. While the evolutionary basis for the structural and functional divergence of duplex telomere-binding proteins is not fully understood, a likely driver is the irregularity of the fission and budding yeast telomere repeats (e.g., TTAC0-1A 0-1C 0-1G1-9 for S. pombe and TG1-3 for Saccharomyces cerevisiae) and its deviation from the canonical GGGTTA repeat found in mammals and all major eukaryotic branches. (For a more detailed discussion of telomere sequence divergence in budding and fission yeasts and the corresponding changes in duplex telomere-binding proteins, see Sepsiova et al. and Steinberg-Neifach and Lu.)

A potentially informative model with which to address mechanistic and evolutionary issues related to duplex telomere-binding proteins is the basidiomycete Ustilago maydis. This model was originally developed by Robin Holliday to investigate recombinational repair. We were initially attracted to this model—in comparison to other fungal models—because (1) Ustilago maydis carries a telomere repeat unit that is identical to the mammalian repeat; (2) the recombination and repair machinery in this organism bears greater resemblances to those in mammals; and (3) repair proteins in this organism (e.g., Rad51, Blm, and Brh2) promote telomere replication and telomere recombination just like their mammalian ortholog counterparts. Accordingly, we have taken advantage of U. maydis to examine the mechanisms and regulation of repair proteins at telomeres. Interestingly, with respect to duplex telomere-binding proteins, our initial database query (as well as an earlier bioinformatics screen) revealed in the U. maydis genome just one potential duplex telomere-binding protein that is structurally similar to SpTeb1 and YfTay1, and that we and others had named UmTrf1. We showed in a published report that this protein has high affinity and sequence specificity for the cognate telomere repeats, and that it interacts directly with Blm helicase, just like mammalian TRF1. To underscore that this protein is structurally distinct from mammalian TRFs, we will hereafter refer to it as UmTay1. While our initial findings suggest that UmTay1 could be the main duplex telomere repeat factor in U. maydis, subsequent analysis of tay1Δ mutants revealed little telomere protection phenotypes (see below), suggesting that the true telomere capping protein remains to be identified. Indeed, Tomaska and colleagues have noted the existence of a TRF/TRF1-like gene in U. maydis with unknown function and potential TRFH and Myb domains, which we will hereafter designate as UmTrf2 (UniProt ID: A0A0D1E1Z3). Thus, U. maydis apparently harbors two structurally distinct, Myb-containing proteins that may mediate distinct or overlapping functions in telomere protection and maintenance. Because Basidiomycetes occupy a more basal position in fungal phylogeny than the ascomycete budding and fission yeasts, studies of duplex telomere-binding proteins in this lineage may offer insights on the origin of these protein families in fungi and the division of labor between them vis-à-vis other organisms.

Herein we present our structural and functional analysis of UmTay1 and UmTrf2. We showed that both proteins, despite their unusual structural features (see Results), recognize the canonical telomere repeat GGGTTA/TAACCC with high affinity and sequence specificity. Trf2 but not Tay1 is critical for telomere protection. Tay1 is evidently essential, and transcriptional repression of this gene results in profound telomere aberrations, accumulation of ssDNA, and elevated levels of extra-chromosomal telomere repeats (ECTR). Tay1, on the other hand, contributes to telomere replication and positively stimulates an ALT-like pathway of telomere recombination, most likely by interacting with Blm and controlling its helicase activity. This functional division of labor is unlike that observed in metazoans or in budding/fission yeasts. Cross comparison of our results with those from other fungal and non-fungal systems highlights the potential of structurally distinct telomere proteins to execute similar functions in different lineages. Based on the accumulated data, we propose that the emergence of the Tay1 family of telomere proteins in fungi may have provided added flexibility to the fungal telomere system to adapt to drastic changes in telomere sequences in budding and fission yeasts.

Results

Tay1 binds GGGTTA/TAACCC repeats through tandem Myb motifs. As noted before, UmTay1 is similar to SpTeb1 and YfTay1 in harboring two consecutive Myb motifs that manifest substantial similarities to the Myb motif of TRF proteins (Supplementary Fig. 1a). However, unlike the other homologs, UmTay1 possesses a much longer C-terminal extension, rendering the protein twice as large as SpTeb1 and YfTay1. To confirm the DNA-binding activity of the Myb domain and investigate the potential role of the C-terminus, we purified an N-terminal 274...
amino acid (a.a.) fragment (spanning the duplicated Myb motifs and named Tay1N274) and compared its DNA-binding properties to those of the full length protein (Fig. 1a). In a titration analysis, Tay1 exhibited ~4-fold higher affinity binding to telomere DNA than Tay1N274; the estimated K_d of these proteins for the TTAGGG4 probe are 27 and 130 nM, respectively (Fig. 1b). We also assessed the sequence-specificity of Tay1 using competitor oligonucleotides harboring different telomere sequences. The YlTEL oligo (derived from the Y. lipolytica telomere sequence), in which the canonical 6-bp repeat unit (GGGTTA) is 4-nt apart, was slightly less active than the cognate 6-bp repeat (TTAGGG4) (Fig. 1c). In contrast, the CgTEL oligo (derived from the Candida glabrata telomere sequence), another G-rich repeat that lacks the GGGTTA sequence, was unable to compete. This binding preference is similar to that reported for Y. lipolytica Tay1, specifically with regard to the strong preference for DNA with canonical repeats11,12, and indicates that UmTay1 has the requisite DNA-binding properties to act specifically at U. maydis telomeres.

**Trf2 is also a high affinity telomere repeat-binding factor.** In addition to UmTay1, Sepsiova et al. noted the potential existence of a TRF/TBF1-like protein (Uniprot ID: A0A0D1E1Z3), which, being 1528 a.a. long, is considerably larger than a typical TRF/Trf1 ortholog such as human TRF1 (Fig. 1a). We named this gene Umtrf2, and performed additional bioinformatic analysis by identifying and characterizing closely related homologs in the Pucciniomycotina and Ustilaginomycotina subphyla (Supplementary Fig. 1b, c). Like UmTrf2, these fungal proteins are ~1.5 to 3 times the sizes of mammalian TRFs. Nevertheless, multiple sequence alignment of these proteins against metazoan TRF orthologs revealed the existence of both a central TRFH domain and a C-terminal Myb domain, precisely as predicted for a true TRF ortholog (Supplementary Fig. 1b, c). Notably, the fungal proteins carry multiple insertions in both their TRFH and Myb domains (vis-à-vis the metazoan proteins), thus accounting in part for their larger sizes. Notwithstanding these insertions, structural considerations suggest that the fungal protein domains may retain the function of TRFH and Myb in dimerization and DNA-binding. For example, even though the Myb motif of UmTrf2 harbors three short insertions (Fig. 1a and Supplementary Fig. 1c), these insertions map to regions of the motif that are distal to the DNA-binding surface as judged by sequence alignment and structure of the human TRF1–DNA complex (Fig. 2a). Accordingly, they are not predicted to disrupt UmTrf2-DNA binding. Notably, whereas insertion 1 is quite variable in length (2–15 a.a.), insertions 2 and 3 are relatively invariant in length.

---

**Fig. 1** The DNA-binding activity of *Ustilago maydis* Tay1. **a** Schematic illustrations of the domain structures of the *U. maydis* Tay1 and Trf2 proteins analyzed in this study, as well as that of human TRF1, a prototypical member of the TRF/TBF1 protein family. The TRFH and Myb domains are shown in green and purple, respectively. All the *U. maydis* recombinant proteins examined in this work also contain a C-terminal FLAG tag (as indicated by the designations Tay1-FG and Tay1N274-FG), which is not illustrated. **b** EMSA analysis of the DNA-binding activities of Tay1-FG and Tay1N274-FG (Tay1 and Tay1N274 carrying a C-terminal FLAG tag). The assays were performed using a series of progressive, 3-fold increases in protein concentrations ranging from 4.4 to 360 nM. A representative set of assays is shown on the left and the quantitation (average ± SEM, n = 3 independent experiments) shown on the right. Previous investigations of SpTeb1 revealed K_d that range from 20 nM to >1 µM, possibly owing to the different DNA substrates used in these studies12,50. Our estimated K_d for Tay1 and Tay1N274 (27 and 130 nM, respectively) are broadly in line with these earlier analyses. **c** The DNA-binding specificity of Tay1-FG was examined using three different double-stranded competitor oligonucleotides. The sequences of the competitors are shown on top with the canonical GGGTTA repeat unit underlined—only the G-strand sequences are displayed. CgTEL and YITEL are based on the telomere repeat of *Candida glabrata* and *Yarrowia lipolytica*, respectively. A representative series of assays is shown and the quantified results (average ± SEM, n = 3 independent experiments) plotted on the bottom.
and contain well-conserved residues that may serve important functions other than DNA-binding.

To test the DNA-binding activity of UmTrf2 directly, we attempted to purify the full length protein but were unsuccessful. Because the long N-terminal extension of this protein is probably unstructured and poorly conserved even among fungi, we tested and succeeded in purifying a truncated derivative that is missing the first 415 amino acids (Trf2AN-FG3). Purified Trf2AN-FG3 binds with high affinity to double-stranded TTAGGG4 (Kd of ~20 nM) but not the G-strand alone (Fig. 2b and Supplementary Fig. 2b). We then purified and analyzed the putative TRF region and Myb domain of Trf2 separately (amino acid 416-1349 and 1350-1528, respectively), and found that as predicted, only the Myb domain exhibited low affinity binding to the telomeric probe (Supplementary Fig. 2b).

Further analysis indicates that the two U. maydis telomere-binding proteins manifest quite similar DNA-binding properties. For example, oligo competition experiments indicate that just like Tay1, Trf2 strongly prefers DNAs that contain the canonical 6-bp repeat (Fig. 2c). In addition, the two proteins have nearly indistinguishable minimal DNA target sizes of ~2 to 3 repeats (Supplementary Fig. 3a, b). The TTAGGG2,5 probe, which harbors 2 copies of the GGGTTA repeats (the target of a single Myb motif according to structural analysis), was bound with lower affinity by both Tay1 and Trf2 in comparison to TTAGGG3,5 and TTAGGG4, which have 3 copies of GGGTTA. On the other hand, TTAGGG2,0, with just 1 copy of GGGTTA, did not support any detectable binding (Supplementary Fig. 3). Thus, despite the structural differences between Tay1 and Trf2, the two U. maydis proteins evidently recognize telomere repeat DNA with very similar mechanisms. This notion is also consistent with conservation of putative DNA-binding residues between UmTay1 Myb1, UmTay1 Myb2, UmTrf2 Myb, and metazoan TRF Myb domains (Supplementary Fig. 1a, c).

**Tay1 is required for telomere replication but not protection.** To investigate the telomeric functions of Tay1 in vivo, we analyzed the phenotypes of multiple, independently derived tay1Δ clones. These clones were found to exhibit similar growth rates as the parental strain (FB1), and manifest no evidence of senescence following prolonged passages, suggesting that Tay1 is not essential for either telomere protection or telomerase function. However, telomere restriction fragment (TRF) Southern analysis revealed substantially shortened but stably maintained telomeres in the tay1Δ mutants (Fig. 3a). The shortening was evident at late time points (Fig. 3a), but did not worsen at later time points. Indeed, both wild type and tay1Δ telomeres grew slightly during continuous sub-culturing (Fig. 3a), a phenomenon that has been noted earlier. The telomere phenotypes of tay1Δ clones are reminiscent of DNA repair mutants and point to telomere replication as the main function of this protein. Previous characterization of the telomere replication mutants (e.g., blmΔ and rad51Δ) revealed preferential loss of long telomeres. This feature is also true of tay1Δ telomeres, in STELA analysis of U. maydis bearing telomeres, the longest telomere tracts found in tay1Δ were ~300 bp shorter than those in the wild type strain, whereas the shortest tracts were similar in size (Fig. 3b). (UT4 is a previously characterized subtelomeric element in U. maydis.) We also considered the alternative possibility that Tay1 may promote telomerase function rather than telomere replication. To test this, we engineered conditional mutants of trt1Δ and tay1Δ by placing these two genes under the control of the arabinose-dependent promoter (crg1, a gene that is repressed in YPA but expressed in YPD). After 4 streaks (~100 generations) in YPD, the trt1Δcrg1 tay1Δcrg1 double mutant manifested greater telomere loss than either trt1Δcrg1 or tay1Δcrg1, indicating that the two genes do not act in the same pathway (Supplementary Fig. 4).

In contrast to the telomere length defects observed in Southern and STELA analysis, in-gel hybridization assays did not uncover...
obvious changes in ssDNA at tay1Δ telomeres. In addition, we did not detect any C-circles, which are extrachromosomal circular telomere species that often accumulate when de-protected telomeres become recombinogenic19,25. Thus, the predominant telomere function of UmTay1 is evidently limited to telomere replication. We also assessed the function of tay1 in DNA repair, and found that tay1Δ did not manifest greater susceptibility to various DNA-damaging agents than the parental strain, except for a moderate sensitivity to MMS (Fig. 3c, bottom panel).

Tay1 binds to and controls the activity of Blm helicase. We have previously shown that just like mammalian TRF1 and TRF2, UmTay1 physically interacts with the conspecific Blm helicase20. Given this interaction and the similarity between the telomere phenotypes of tay1Δ and blmΔ, we hypothesize that Tay1 promotes telomere replication by controlling Blm activity. To test this possibility biochemically, we analyzed the effects of Tay1 on Blm helicase activity in vitro (Fig. 4a). Two double-stranded DNA substrates with 3' overhangs, one carrying a telomeric sequence and the other a non-telomeric sequence, were used to assess the helicase activity of Blm, a 3' to 5' helicase26. Notably, the untreated telomeric substrate migrated as two distinct species in native gels, most likely owing to G-rich overhang-mediated dimer formation, especially given the propensity of G-rich sequences to form higher order structures such as G-quadruplexes (Fig. 4b, leftmost lane; Supplementary Fig. 5, 6th lane). As expected, the addition of Blm resulted in the unwinding of duplex DNA into ssDNA in an ATP-dependent manner (Supplementary Fig. 5). Interestingly, the effects of Tay1 on unwinding were substrate sequence-dependent: on telomeric DNA, Tay1 inhibited DNA unwinding by Blm by up to 70%, whereas on non-telomeric DNA, the effect was stimulatory by up to 2.5 fold (Fig. 4a, b; Supplementary Fig. 6a). The inhibitory effect of full length Tay1 can be recapitulated by the N-terminal Myb domain (Tay1 N274), suggesting that high affinity DNA-binding by Tay1 may be sufficient to block the unwinding activity of Blm (Supplementary Fig 6b). In contrast, Tay1 N274 is unable to stimulate unwinding on the non-telomeric substrate, implying that the C-terminal regions of Tay1 contribute to this activity, possibly through protein–protein interactions (Supplementary Fig. 6a). It should be noted that Blm was included at 12.5 and 25 µM for the non-telomeric and telomeric assays, respectively, in order to accentuate the stimulatory and inhibitory effects of Tay1.

To examine functional interactions between UmTay1 and Blm in vivo, we generated blmΔ and tay1Δ blmΔ mutants in the same strain background as tay1Δ and compared their telomere phenotypes. All three mutants manifested stably shorted telomeres, consistent with a telomere replication defect (Fig. 4c). Notably, as judged by the loss of long telomeres in both telomere Southern and STELA analysis, the defect is more severe in tay1Δ blmΔ than in either single mutant (Fig. 4c; Supplementary Fig. 7). The exacerbated phenotype of tay1Δ blmΔ vis-à-vis tay1Δ and blmΔ is most notable in the upper range of STELA size distribution; the longest fragments in tay1Δ and blmΔ (~1.3 to 1.5 kb) are mostly absent in tay1Δ blm1Δ (Supplementary Fig. 7). These results suggest that the two genes do
not act exclusively in a single pathway in promoting telomere replication. Indeed, we had previously detected an interaction between *U. maydis* Blm and Pot1, raising the possibility that Blm may be regulated by multiple telomere proteins. **Likewise, it is possible that Tay1 may interact with another factor(s) that contributes to telomere replication. However, with respect to potential interacting partners, we note that Tay1 does not appear to bind Rad51/Brh2, another DNA repair complex implicated in telomere replication**.

Another previously established function of Blm at telomeres is to trigger an ALT-like pathway when telomeres are de-protected. More specifically, when the *ku70*/*80* complex is transcriptionally repressed in *U. maydis* (e.g., in the *ku70*/*80* mutant), telomeres become highly aberrant and manifest a collection of abnormalities that resemble those in ALT cancer cells, including telomere length heterogeneity, ssDNA on the C-strand, and high levels of C-circles. All these phenotypes are completely suppressed in *ku70*/*80* *blmΔ*, implying an essential role for *blm* in triggering the ALT pathway. Given the ability of Tay1 to regulate Blm activity, we proceeded to generate *ku70*/*80* *tay1Δ* mutants and analyzed their telomeres. Notably, all three independently generated *ku70*/*80* *tay1Δ* exhibited milder telomere length abnormalities than *ku70*/*80* when *ku70* expression was repressed. In particular, telomere length distributions became less heterogeneous, and the levels of C-circles were substantially reduced in *ku70*/*80* *tay1Δ* relative to that in *ku70*/*80* (Fig. 4d, e). The levels of C-strand ssDNA in the double mutants were also lower than that in the *ku70*/*80* single mutant (Supplementary Fig. 9a). Consistent with the mild telomere aberrations, the *ku70*/*80* *tay1Δ* double mutants exhibited very little growth defects and did not manifest the enlarged and elongated morphology characteristic of G2/M arrest (Supplementary Fig. 9b, c).

**Trf2 is crucial for telomere protection.** The lack of an obvious telomere de-protection phenotype in *tay1Δ* mutants suggests that *Um*Trf2 may be the key telomere protection factor. To test this idea, we attempted to generate null mutants by gene replacement. Multiple attempts were unsuccessful, suggesting that the gene may be essential. We then constructed transcriptionally conditional mutants of *trf2* (named *trf2*/*crg1*) in which the endogenous *trf2* promoter is replaced by the *crg1* promoter. As noted before,
Trf2 is critical for cell proliferation and telomere protection. A trf2
crg1 clone was passaged on YPD (a medium that represses trf2
transcription), and the growth of the clone at successive streaks
was assessed along with the parental strain. The plates were
incubated at 30 °C for 3 days. B A
trf2
crg1 clone was grown in liquid YPA or YPD, and examined
under the microscope. C DNAs from multiple, independently
derived trf2
crg1 mutants that had been grown in YPA or YPD were
subjected to TRF Southern analysis. D Untreated or Pvul-digested
cromosomal DNAs were fractionated on agarose gel, transferred
to nylon membrane, and then subjected to Southern hybridization
using a TTAGGG82 probe. The short telomere fragments that can be
detected in the absence of Pvul digestion are highlighted as extra-chromosomal
 telomere repeats (ECTR). E Multiple independent trf2
crg1 clones were passaged on YPD plates by repeated re-streaking.
DNAs from liquid culture out-growth of the resulting colonies were subjected to
telomere restriction fragment Southern analysis.

The role of Trf2 in telomere protection is also supported by
analysis of single-stranded telomere DNA through in-gel
hybridization. In particular, low levels of G-strand ssDNA
(single-strand DNA) were detected in trf2-expressing clones,
consistent with normal telomeres harboring G-strand, 3′-
overhangs (Fig. 6a). In contrast, in trf2-repressed clones,
G-strand ssDNA was reduced whereas C-strand ssDNA became
greatly elevated (Fig. 6a). Moreover, the amounts of C-circles,
another marker of abnormal telomere recombination, were
also increased in these clones (Fig. 6b). Notably, these ssDNA
and C-circle aberrations are also found in ku70- or ku80-
repressed mutants, suggesting that the abnormal telomere
processing reactions in the trf2 and ku-deficient mutants may
share some mechanistic resemblance. However, there may be
differences as well, given that the levels of C-circles in the
ku70nar1 are more than 10-fold higher than those in trf2
crg1 (cf. Figs. 4e, 6b, note that the two sets of assays were performed

![Diagram showing telomere repeats and fragments](image_url)
concurrently, analyzed on the same blot, and quantified using the same scale).

To further examine the protective function of trf2, we developed a PCR-based assay to assess telomere–telomere (T–T) fusions between UT4/5 and UT6-bearing telomeres (Supplementary Fig. 10a). Notably, the vast majority of PCR products generated in these assays hybridized to both a UT4/5 probe and a UT6 probe, as predicted for products derived from authentic fusions (Supplementary Fig. 10a). Cloning and sequencing of a 1.1 kb product further confirmed the presence of UT4/5 and UT6 sequences on the two ends of this fragment (Supplementary Fig. 10b). Applying the fusion assay to the U. maydis trt1Δtrf2Δ strains grown in YPA or YPD concurrently, analyzed on the same blot, and quantified using the same scale).

Fig. 6 Ustilago maydis Trf2 suppresses multiple telomere aberrations. a DNAs from multiple, independently constructed trf2Δtrg1 strains grown in YPA or YPD were assayed for the levels of unpaired G-strand and C-strand telomere DNA by in-gel hybridization. b The indicated chromosomal DNA preparations were subjected to duplicate C-circle assays. The scan of the blot is shown on the left and the quantification (means; n = 2 experimental replicates; normalized against the average value of ku70Δ samples in Fig. 4e) plotted on the right. Note that these assays were performed concurrently with those in Fig. 4e and quantified from the same blot. c The DNAs from the indicated strains grown in the given media were subjected to telomere fusion analysis.

Discussion
Fungi have been a rich source of discovery for telomere biology. Nevertheless, except for a few standard and atypical organisms in the ascomycete phylum, the telomere machinery in different fungal lineages remains poorly understood. We have sought to characterize telomere regulation in Ustilago maydis as a starting basis for understanding telomeres in Basidiomycetes. Because of the resemblance of the repair machinery in U. maydis to that in metazoans, this fungus could also serve as a useful model for the interplay between the telomere and repair machinery. In this study, we focused on the structural, functional, and mechanistic features of two double-strand telomere-binding proteins in U. maydis, named UmTay1 and UmTrf2. Notably, we showed that both proteins, despite their unusual and quite different structural features, recognize the canonical telomere repeat GGGTTA/TAACCC with very similar affinity and sequence specificity. Furthermore, the two proteins mediate quite distinct telomere regulatory functions, with Tay1 contributing primarily to telomere replication and telomere recombination, and Trf2 being the chief telomere protection factor. These observations have interesting mechanistic and evolutionary ramifications, as discussed below.

The DNA-binding mechanisms of TRF/TBF1 protein family have been studied extensively. High resolution structures of the Myb domain–DNA complexes revealed crucial contacts between the two Myb protomers (head-to-tail arrangement) and consecutive GGG/CCC triplets in the telomere DNA targets. In addition, single particle EM analysis of full length TRF1 suggests that the TRFH dimerization domain helps to position the Myb protomer relative to the DNA target.

In contrast, small angle x-ray scattering (SAXS) envelope of TRF2 dimer revealed an extended conformation that leaves the two Myb protomers far apart. It seems likely that, given its similar structural organization and similar DNA-binding specificity, UmTrf2 would recognize telomere repeats in the same fashion as mammalian TRF1 and TRF2, i.e., through specific interactions between Myb protomers and neighboring GGG/CCC triplets. One notable feature of UmTrf2 is the large linker between its TRFH and Myb domains (~600 a.a. compared to ~100 a.a. in human TRF1 and ~200 a.a. in human TRF2), a feature that could provide additional flexibility to UmTrf2 in positioning its two Myb protomers relative to the DNA target. This idea is consistent with the ability of UmTrf2 to bind U. maydis and Y. lipolytica telomere sequences with similar affinities, despite the differences in gaps between the consecutive GGG/CCC triplets (3 bp for the U. maydis repeat and 7 bp for the Y. lipolytica repeat).
Less is known about the DNA-binding mechanism of the Tay1 protein family, which are characterized by two tandem Myb motifs separated by ~20 a.a. (or much longer in some homologs such as those from *Coccidioides posadasii* or *Uncinocarpus reesei*)\(^{11,20}\). While no high resolution structural information is available for this protein family, the similarities in the DNA target size and sequence preference exhibited by *Un*Tay1 and *Um*Tay2 suggest that Tay1-like proteins may position their two Myb motifs on telomere DNA in the same way as TRF/TBF1 proteins. Notably, structural considerations suggest that the linker between the two Myb motifs in Tay1 (minimally 20 a.a.) is adequate to span the distance between the C-terminus of one Myb protomer and the N-terminus of the other protomer as judged by the crystal structure of the TRF1–DNA complex (30–40 Å depending on the combination). The long linker between the Myb motifs may also account for the ability of *Um* Tay1 to bind *Y. lipolytica* telomeres, in which the GGGTTA core repeat is separated from each other by four nucleotides. Even though the two Myb motifs are expected to be further apart along the DNA axis, they would lie on the same side of the DNA double helix, thus reducing the spatial separation that needs to be spanned by the protein linker. Therefore, we propose that a Tay1 monomer might be sufficient for high affinity and sequence-specific binding to telomere repeats. Indeed, our preliminary co-expression/pull down experiments did not reveal detectable interactions between differently tagged *Um* Tay1 (Supplementary Fig. 11), suggesting that the protein does not form a stable dimer. However, we cannot exclude the possibility of DNA-induced oligomerization, which was suggested by EM analysis of *Y. lipolytica* Tay1–DNA complex\(^{11}\).

The Tay1 protein family, which encompasses *Um* Tay1, provides an interesting case study in the plasticity of the telomere machinery (Fig. 7). Tay1 homologs appear to be confined to fungi, and database analysis suggests that they emerged prior to the divergence of the Dikarya clade, since family members can be identified in both Basidiomycete and Ascomycetes subphyla (e.g., homologs in *Aspergillus* (Pezizomycotina), *Schizosaccharomyces* (Taphrinomycotina), *Puccinia* (Pucciniomycotina), and *Yarrowia* (Saccharomycotina)), but not in more basal branches\(^{11,12}\). Interestingly, this gene was evidently lost in late branching lineages of Saccharomycotina yeasts (e.g., *Saccharomyces* and *Candida*), the complete absence of the GGGTTA sequence renders telomeres refractory to binding by either Tay1 or TRF/TBF1, thus necessitating the utilization of another Myb-containing protein (Rap1) with a more relaxed sequence specificity to bind and protect telomeres\(^{13}\). While TBF1 in these organisms might have been retained because of its roles in transcription and subtelomere regulation\(^{32–34}\), Tay1 was probably lost due to the lack of any critical function. Altogether, the emergence of Tay1 in fungi and its subsequent loss in specific lineages provide interesting illustrations of functional acquisition and loss in evolution by a new protein family. The ability of this

**Fig. 7 Functional plasticity of the Tay1 protein family.** A simplified fungal phylogeny is shown along with the distribution of the TRF/TBF1 and Tay1 protein family in the phylogeny. The tree is adapted from a previous study\(^{51}\). The phyla in which the telomere repeats deviate from the canonical GGGTTA sequence are indicated in red. The existence and functions of TRF/TBF1 and Tay1 homologs in selected organisms are also illustrated.
protein family members to evolve different telomeric and non-telomeric functions may have provided added flexibility to fungi to meet the challenges posed by telomere sequence alterations. Indeed, beyond TRF/TvB1 and Tay1, the plasticity of duplex telomere-binding proteins is further illustrated by Rap1 and Tar1, two other Myb/Sant-containing proteins that emerged later to serve as the major TBPs in budding and fission yeasts, respectively.22,23

Another key finding in this study is the demonstration of a stimulatory function for UmmTay1 in the ALT-like pathway in U. maydis. This is to our knowledge the first example of a telomere protein exerting a positive regulatory function in ALT-related pathways and could have relevance for similar pathways in mammalian systems. Both mammalian TRF1 and TRF2 have been shown to bind and possibly regulate the cognate BLM, and protein exerting a positive regulatory function in ALT-related pathways with distinct factor requirements. In this regard, the rather limited telomere function of Tay1 in U. maydis allowed us to address the issue more easily through the use of tay1Δ null mutations.

The current work also points to potential complications in studies that investigated ALT mechanisms through the use of TRF1-Fok1, a fusion protein that induces telomere-specific breaks.37,38 The DNA synthesis/repair triggered by the induction of TRF1-Fok1 has been proposed to mimic the natural ALT pathway. However, if TRF1-Fok1 retains the ability to recruit or activate BLM, the breaks generated by this fusion protein might be preferentially processed by BLM, and preferentially channeled toward specific repair pathways rather than those utilized by endogenous ALT substrates. Given that there are multiple ALT pathways with distinct factor requirements,39-41 the roles of telomere proteins in regulating these pathways and in modulating the outcomes of BIR-like DNA synthesis at telomeres would appear to warrant further investigation.

Methods

Ustilago maydis strains and growth conditions. Standard protocols were employed for the genetic manipulation of U. maydis. All U. maydis strains used in this study were haploid and were derived from the UCM350 or FB1 background and are listed in Supplementary Table S1. The kat7Δmutants harbored a conditional kat7Δ under the control of the nitrate-dependent promoter from nar1.11 Accordingly, the kat7Δ gene in this strain is expressed in the MM medium and repressed in YPD. The tay1Δ mutants were generated by replacing the entire tay1 open reading frame in FB1 with a cassette expressing resistance to hygromycin (HygP) through homologous recombination.45,46 Briefly, the plasmid harboring the tay1 disruption cassette with the HygP marker (named pUT1-Tay1-Hyg) was constructed through Gibson Assembly (NEB) of four overlapping PCR fragments, which were generated using appropriate primers (Supplementary Table S2) and either genomic DNA or plasmid pUMar261 as templates. The four PCR fragments linked by Gibson assembly include: (i) 5′ UTR of tay1 connected to the 5′ end of the hygromycin ORF, (ii) the complete hygromycin ORF flanked by short segments from the 5′ and 3′UTR of tay1, (iii) 3′ end of the hygromycin ORF linked to 3′ UTR of tay1, and (iv) the vector sequence from pUMar261 flanked by short segments from the 5′ and 3′UTR of tay1. The disruption cassette was released from the vector by cleaving pUUTay1-Hyg with EcoRI and BamHI, and used for transformation. The kat7Δkat7Δ tay1Δ mutants were constructed from the tay1Δ mutant using the pUK70mut1 plasmid.17

The tay1ΔΔ strain was constructed by integrating a linearized pRU12-trf2 plasmid containing fragments that match (1) 700 bp of the trf2 promoter and (2) the first 700 bp of the trf2 ORF into the trf2 genomic locus in UCM350. Briefly, the two trf2 fragments were generated by PCR using appropriate oligos with flanking restriction enzyme sites (Supplementary Table S3), and then cloned together in between the NdeI and EcoRI sites of plasmid pRU12tri.27 Correct recombination of Xhol treated pRU12-trf2 into the trf2 locus results in the insertion of the CBX8 marker and the crgl promoter precisely adjacent to and upstream of the trf2 ATG initiation codon. Following transformation, the correct mutant strains were selected on YPA supplemented with both carboxin (600 μg/mL) and conduritol B-β-glycoside (50 μM), and confirmed by Southern blot analysis. The tay1ΔΔ strain was made using the same approach, i.e., by constructing the equivalent pRU12-tay1 plasmid and integrating Xhol-linearized pRU12-Tay1 into the tay1 genomic locus. The trf2Δ1 strains were confirmed by PCR reactions except that the trf1 fragments (corresponding to an upstream region and a 5′ portion of ORF) were cloned in between the NdeI and BamHI sites of pRU11-hyg, and the transformants were selected on YPA supplemented with hygromycin (150 μg/mL).

Another key finding in this study is the demonstration of a stimulatory function for UmmTay1 in the ALT-like pathway in U. maydis. This is to our knowledge the first example of a telomere protein exerting a positive regulatory function in ALT-related pathways and could have relevance for similar pathways in mammalian systems. Both mammalian TRF1 and TRF2 have been shown to bind and possibly regulate the cognate BLM, and protein exerting a positive regulatory function in ALT-related pathways with distinct factor requirements. In this regard, the rather limited telomere function of Tay1 in U. maydis allowed us to address the issue more easily through the use of tay1Δ null mutations.

The current work also points to potential complications in studies that investigated ALT mechanisms through the use of TRF1-Fok1, a fusion protein that induces telomere-specific breaks.37,38 The DNA synthesis/repair triggered by the induction of TRF1-Fok1 has been proposed to mimic the natural ALT pathway. However, if TRF1-Fok1 retains the ability to recruit or activate BLM, the breaks generated by this fusion protein might be preferentially processed by BLM, and preferentially channeled toward specific repair pathways rather than those utilized by endogenous ALT substrates. Given that there are multiple ALT pathways with distinct factor requirements,39-41 the roles of telomere proteins in regulating these pathways and in modulating the outcomes of BIR-like DNA synthesis at telomeres would appear to warrant further investigation.
Statistical and reproducibility. Student’s t tests were performed using GraphPad Prism 6, and the p-values were calculated and indicated as follows: *p < 0.05; **p < 0.01; ***p < 0.001. Each experiment was repeated two or three times to ensure reproducibility. All comparisons between different mutants were conducted using two or three independently generated mutants of the same genotypes.

Data availability
The data generated and/or analyzed during the current study are included within the paper and Supplementary Information or available from the corresponding author upon reasonable request. The original, uncropped gel and blot images for Figs. 1-6 are included in the Supplementary Information File as Supplementary Fig. 13. The source data for all the plots in the paper are included in the Supplementary Data 1 file.

Code availability
The source code for TeSLA can be downloaded at the following link: https://www.nature.com/commsbio.

References
1. Jain, D. & Cooper, J. P. Telomeric strategies: means to an end. Annu. Rev. Genet. 44, 243–269 (2011).
2. de Lange, T. Shelterin-mediated telomere protection. Annu. Rev. Genet. 52, 223–247 (2018).
3. Mierova, R. & Blasco, M. A. Replicating through telomeres: a means to an end. Trends Biochem. Sci. 40, 504–515 (2015).
4. Harley, C., Futcher, A. & Greider, C. Telomeres shorten during ageing of human fibroblasts. Nature 345, 548–540 (1990).
5. Zimmermann, M., Kibe, T., Kabir, S. & de Lange, T. TRF1 negotiates TTAGGG repeat-associated replication problems by recruiting the BLM helicase and the TPPI/POT1 repressor of ATR signaling. Genes Dev. 28, 2477–2491 (2014).
6. Lillard-Wetherell, K. et al. Association and regulation of the BLM helicase by the telomeres proteins TRF1 and TRF2. Hum. Mol. Genet. 13, 1919–1932 (2004).
7. Cooper, J. P., Nimmo, E. R., Alloshire, R. C. & Cech, T. R. Regulation of telomere length and function by a Myb-domain protein in fission yeast. Nature 385, 744–747 (1997).
8. Miller, K. M., Rog, O. & Cooper, J. P. Semi-conservative DNA replication through telomeres requires Taz1. Nature 440, 824–828 (2006).
9. Valente, L. P. et al. Myb-domain protein Tbz1 controls histone levels and centromere organisation in fission yeast. EMBO J. 32, 450–460 (2013).
10. Kramara, J. et al. Evolution of telomeres in Schizosaccharomyces pombe and its possible relationship to the diversification of telomere binding proteins. PLoS One 11, e0154225 (2016).
11. Steinberg-Neifach, O. & Lue, N. F. Telomere DNA recognition in Saccharomyces cerevisiae: potential lessons for the co-evolution of SS DNA and dsDNA-binding proteins and their target sites. Front. Genet. 6, 162 (2015).
12. Holoman, W. K., Schirawski, J. & Holliday, R. The homologous recombination system of Ustilago maydis. Fungal Genet. Biol. 45, S31–39 (2008).
13. Guzman, P. A. & Sanchez, J. G. Characterization of telomeric regions from Ustilago maydis. Microbiology 140 (Pt 3), 551–557 (1994).
14. Holoman, W. K. Unraveling the mechanism of BRC2A in homologous recombination. Nat. Struct. Mol. Biol. 18, 748–754 (2011).
15. de Sena-Tomas, C., Yu, E. Y., Calzada, A., Holoman, W. K., Lue, N. F. & Perez-Martin, J. Fungal Ku prevents permanent cell cycle arrest by suppressing DNA damage signaling at telomeres. Nucleic Acids Res. 43, 2138–2151 (2015).
16. Yu, E. Y., Kostrub, C. F., Buchman, A. R. & Holloman, W. K. BRCA2 homolog required for proflavin-induced telomere maintenance in Ustilago maydis. Nucleic Acids Res. 47, 1948–1963 (2018).
17. Sanchez-Alonso, P. & Guzman, P. Predicted elements of telomere organization and function in Ustilago maydis. Fungal Genet. Biol. 45, 554–62 (2008).
18. Mezaros, B., Erdos, G. & Dossztanyi, Z. UPIPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. Nucleic Acids Res. 46, W329–W337 (2018).
19. Court, R., Chapman, L., Fairall, L. & Rhodes, D. How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: a view from high-resolution crystal structures. EMBO Rep. 6, 39–45 (2005).
20. Swapan, G., Yu, E. Y. & Lue, N. F. Single telomere length analysis in Ustilago maydis, a high-resolution tool for examining fugal telomere length distribution and C-strand 5’-end processing. Micro. Cell. 5, 393–403 (2018).
21. Henson, J. D. et al. DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. Nat. Biotechnol. 27, 1181–1185 (2009).
22. Croteau, D. L., Popuri, V., Opreko, P. L. & Bohr, V. A. Human RecQ helicases in DNA repair, recombination, and replication. Annu. Rev. Biochem. 83, 519–532 (2014).
23. Böttin, A., Kamper, J. & Kahmann, R. Isolation of a carbon source-regulated gene from Ustilago maydis. Mol. Gen. Genet. 253, 342–352 (1996).
24. Dillé, R. L., Verma, P., Cho, N. W., Wondisford, A. R. & Greenberg, R. A. RAD52 and RTEL1 compensate telomere loss in the absence of telomerase. Nucleic Acids Res. 46, 2432–2445 (2018).
25. Gauthier, G. et al. A higher-order entity formed by the flexible assembly of RAP1 with TRF2. Nucleic Acids Res. 44, 1962–1976 (2016).
26. Hugues, H. et al. Transcription factor substitution during the evolution of fungal ribosome regulation. Mol. Cell 29, 552–562 (2008).
27. Grohmann, M., W337 (2018).
28. Stavropoulos, D. J. et al. The Bloom syndrome helicase BLM interacts with TRF2 in ALT cells and promotes telomeric DNA synthesis. Hum. Mol. Genet. 11, 3135–3144 (2002).
29. Sobinoff, A. P. et al. BLM and SLX4 play opposing roles in recombination-dependent replication at human telomeres. EMBO J. 36, 2907–2917 (2017).
30. Dillery, R. L., Verma, P., Cho, N. W., Winters, H. D., Wondisford, A. R. & Greenberg, R. A. Break-induced telomere synthesis underlies alternative telomere maintenance. Nature 539, 54–58 (2016).
31. Kojic, M. & Schirawski, J. & Holliday, R. The homologous recombination system of Ustilago maydis. Fungal Genet. Biol. 45, S31–39 (2008).
32. Holoman, W. K., Schirawski, J. & Holliday, R. The homologous recombination system of Ustilago maydis. Fungal Genet. Biol. 45, S31–39 (2008).
33. Kojic, M., Kostrub, C. F., Buchman, A. R. & Holloman, W. K. BRC2A homolog required for proflavin-induced telomere maintenance in Ustilago maydis, a new model system of DNA repair proteins at telomeres. DNA Repair (Amst.) 12, 472–479 (2013).
34. Yu, E. Y., Perez-Martin, J., Holoman, W. K. & Lue, N. F. Mre11 and Blm-dependent formation of ALT-like telomeres in K-deficient Ustilago maydis. PLoS Genet. 11, e1005570 (2015).
35. Yu, E. Y., Hsu, M., Holoman, W. K. & Lue, N. F. Contributions of recombination and repair proteins to telomere maintenance in telomerase-positive and negative Ustilago maydis. Mol. Microbiol. 107, 81–93 (2018).
Acknowledgements

We thank José Pérez-Martín (CSIC, Spain) for plasmids, Qingwen Zhou for purified Rad51 and Brh2/Dss1, Lubomir Tomaska for pointing out the existence of UmTbf1/Trf2, and members of our laboratories for discussion. This work was supported by a pilot project grant from the Sandra and Edward Meyer Cancer Center (N.F.L. and W.K.H.), NSF MCB-1817331, and NIH GM 107287 (N.F.L.).

Author contributions

E.Y.Y., W.K.H., and N.F.L. conceived the study. E.Y.Y., S.Z., and N.F.L. designed and executed the experiments with contributions from S.G., J.S., and M.H.; E.Y.Y. and N.F.L. interpreted the data and wrote the manuscript with advice from S.Z. and W.K.H.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s42003-020-01505-z.

Correspondence and requests for materials should be addressed to N.F.L.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020