SUPPORTING INFORMATION

Biogenesis of telomerase RNA from a protein-coding mRNA precursor

Dhenugen Logeswaran, Yang Li, Khadiza Akhter, Joshua D. Podlevsky, Tamara L. Olson, Katherine Forsberg and Julian J.-L. Chen

SI Materials and Methods

**Fungal strains and plasmids.** The UCM350 haploid strain (*nar1-6 pan1-1 alb1*) of *U. maydis* was a kind gift from Drs. Eun Young Yu and Neal Lue (1). All other strains generated in this study are derivatives of UCM350. Plasmid pPotef-LN22*-3xGNLS-cbx was a gift from Dr. Michael Feldbrügge (Addgene plasmid # 86780)

**Oligonucleotides.** Sequences of oligonucleotides used in this study are listed in Table S4.

**U. maydis cell lysis and western blot analysis.** Log phase *U. maydis* cells were harvested by centrifugation at 5000 xg for 10 mins and the cells were resuspended in lysis buffer (50 mM HEPES pH 7.5, 1.2 mM MgCl₂, 10% glycerol, 0.1 mM EGTA, 0.1 mM EDTA, 1% Tween 20, 5 mM BME, 1X Roche protease inhibitor cocktail, and 1 mM PMSF). Equal volume of 0.5 mm glass beads (Biospec) was added to the lysate and vortexed at the maximum setting for 20 mins in a cold room. The cell lysate was briefly centrifuged at 8500 xg for 5 mins and the lysate was transferred to a fresh tube without disturbing the glass beads. The whole-cell lysate was clarified by centrifugation at 18,000 xg for 5 mins and the soluble fraction was transferred to a fresh tube and the procedure was repeated once. Total protein was quantified using protein assay reagent (BioRad) following the manufacturer’s instruction with a BSA concentration series as protein standards. Ten micrograms of total proteins from either the WT or the 3xFLAG-expressing *U. maydis* strains were analyzed on an 8% SDS-PAGE gel and transferred to a Immun-Blot PVDF membrane (BioRad) using a Trans-Blot Turbo transfer system (BioRad) following manufacturer’s instructions. Membrane was probed with mouse monoclonal anti-FLAG M2 antibody at 1:3000 dilution (Sigma) and GAM-HRP (Goat anti-mouse horse radish peroxidase conjugated) secondary antibody at 1:10,000 dilution. The blot was visualized using Immobilon ECL (enhanced chemiluminescence) Ultra Western HRP substrate (Millipore Sigma) following manufacturer’s instructions and imaged using the Gel Logic 440 imaging system. For western blot detection of recombinant 3xFLAG-*Um*TERT expressing *U. maydis* strains were analyzed on an 8% SDS-PAGE gel and transferred to an Immum-Blot PVDF membrane (BioRad) using a Trans-Blot Turbo transfer system (BioRad) following manufacturer’s instructions. Membrane was probed with mouse monoclonal anti-FLAG M2 antibody at 1:3000 dilution (Sigma) and GAM-HRP (Goat anti-mouse horse radish peroxidase conjugated) secondary antibody at 1:10,000 dilution. The blot was visualized using Immobilon ECL (enhanced chemiluminescence) Ultra Western HRP substrate (Millipore Sigma) following manufacturer’s instructions and imaged using the Gel Logic 440 imaging system. For western blot detection of recombinant 3xFLAG-*Um*Em1, 12% SDS-PAGE gel was transferred to Immobilon-FL PVDF membrane (Millipore Sigma) and probed with anti-FLAG M2 antibody as described above, followed by incubation with IRDye 800CW Goat anti-mouse secondary antibody (LI-COR Biosciences) at 1:15,000 dilution. Internal control section of the blot was probed with anti-α-Tubulin antibody (Santa Cruz, 53030) at 1:250 dilution followed by IRDye 800CW Goat anti-Rat secondary antibody (LI-COR Biosciences). Blot was visualized using the Amersham Typhoon NIR-IP imager.
**Anti-FLAG immunoprecipitation.** Telomerase holoenzyme was affinity-purified by anti-FLAG immunoprecipitation (IP). For each IP, 10 µl of monoclonal anti-FLAG mouse M2 affinity gel (Sigma-Aldrich) was added to a low-retention 1.5 ml microcentrifuge tube. After centrifugation at 1000 xg for 15 sec, the buffer was aspirated. Beads were then washed twice with 30 µl of 1x Tris-Buffered Saline (TBS) solution (20 mM Tris-HCl, pH 7.6 and 150 mM NaCl). Following aspiration of buffer, the lysate was added immediately to the beads and mixed gently on a rotating mixer at 4°C for 1 hr. Following mixing, beads were centrifuged at 1000 xg for 30 sec and lysate was aspirated to a separate tube. Beads were washed once with 200 µl 1X TBS and transferred to a fresh tube. Beads were washed two more times with 100 µl TBS with beads transferred to a fresh tube following each wash. For RNase A treatment, bead suspension was split into two aliquots (50 µl each) and 0.5 µl RNase A solution (VWR - 10 mg/ml) was added to one aliquot and incubated at 37°C for 5 min. Following RNase A treatment, beads were centrifuged at 1000 xg for 30 sec and the supernatant was aspirated prior to telomerase activity analysis by TRAP assay.

**TRAP assay.** Anti-FLAG immunoprecipitated beads were washed once with 1X Primer Extension (PE) buffer (50 mM Tris-HCl, pH 8.0, 0.5 mM MgCl₂, 5 mM 2-mercaptoethanol, and 5 mM spermidine). Following the wash, the TRAP assay master mix (1X PE buffer, 0.5 µM telomeric DNA (TS) primer, mix of dATP+dGTP+dTTP at 50 µM each) was added to the beads and incubated at 30°C for 2 hours. Following phenol/chloroform extraction, the extended DNA primer was ethanol precipitated and washed using 70% ethanol and resuspended in 10 µl water. Aliquots of the purified DNA (5 µl) were added to the PCR master mix (1X ExTaq buffer, dNTP mix of 2.5 mM each, 0.4 µM ³²P labeled TS primer, 0.4 µM ACX primer, and 0.625 U of ExTaq DNA polymerase Hot-start version, Takara). PCR was performed under an optimized condition (94°C for 2 min followed by 30 cycles of 94°C for 30 sec, 60°C for 30 sec, and 60°C for 5 min). PCR products were resolved on a 10% native polyacrylamide gel. Dried gel was exposed to a phosphor-storage screen and analyzed with an FX Pro Molecular Imager (Bio-Rad).

**Purification and Illumina sequencing of TERT bound RNA.** Following anti-FLAG immunoprecipitation and washes, RNA co-purified with 3xFLAG-UmTER was extracted using RNA extraction buffer (20 mM Tris-HCl, pH 7.5, 10 mM EDTA and 0.5% SDS), followed by acid phenol/chloroform extraction and ethanol precipitation. RNA was quantified using a Bioanalyzer 2100 (Agilent technologies, Inc.) instrument and used for cDNA library construction using the Illumina ScriptSeq v2 RNA-Seq library preparation kit following manufacturer’s instructions. Sequencing was performed using the Illumina NextSeq 500 instrument at the Arizona State University genomics core facility (Tempe, AZ, USA).

**Bioinformatic search of UmTER candidates.** Genomic loci with putative template permutations were extracted from the *U. maydis* reference genome (assembly Umaydis521_2.0) and the Illumina sequencing reads were mapped to the loci. The putative templates were defined as 8-12 nucleotides in length and permutation of the 5’-CCCTAA-3’ sequence. A total of 782 loci of 2 kb containing a putative template (length chosen based on the length of *Neurospora crassa* TER – 2049 nt) (2) were extracted from the genome (Fig. 1b). A total of 109,069,132 Illumina 50 bp paired-end sequencing reads were mapped onto the extracted loci in a strand-specific manner using bowtie2 (3). Default parameters were used for bowtie2 mapping and the additional no-unal flag was included to eliminate unmapped reads from the output. Mapped loci were ranked based
on the number of mapped reads and read-covered regions with less than 50 bp of interrupted threshold coverage were defined as a mapped locus. The top 5 loci were used as queries for BLAST against the non-redundant nucleotide database for annotation of known genes. For homolog search, the loci were used as queries in standalone BLAST (version 2.2.31+) search against a custom BLAST database built using the *U. bromivora* genome (assembly UBRO_v3). The tablet genome viewer (4) was used to visualize reference loci and the mapped reads. Bedtools (5) was used to compute the read coverage per base.

**Identification and phylogenetic comparative analysis of UmTER homologs.** For TER identification based on gene synten, the Emi1 and Oat homologs were identified by performing standalone BLAST searches against respective Ustilaginomycetes genomes. Both protein-coding genes from the identified species were manually annotated and the intervening sequence was searched for a conserved template. Identified putative TER sequences which show gene synteny conservation were used in the sequence alignment and phylogenetic comparative analysis to determine a secondary structure model for *UmTER* core domains. Multiple sequence alignment of Ustilaginomycetes TERs was performed initially using the ClustalW algorithm of the BioEdit program (6). Manual refinements were made to improve the preliminary alignments using highly conserved regions and invariant primary sequence motifs as anchor points.

**Total RNA isolation.** *U. maydis* cells were collected and homogenized into a fine powder in liquid nitrogen using a mortar and pestle. Approximately 1 ml of TRI-reagent (Molecular Research Center) was added to 50 milligrams of homogenized powder, followed by acid phenol/chloroform extraction and isopropanol precipitation. Isolated total RNA was pre-treated with Turbo DNase I (Ambion) following manufacturer instructions to remove trace genomic DNA contamination.

**5’ and 3’ RACE.** The 5’ RACE was performed using the template-switching RT enzyme mix (NEB) with the template-switching oligo (TSO) following manufacturer’s instructions. The 3’ RACE was performed using the guanosine/inosine (G/I) or poly(A) tailing strategy of the USB poly(A) tail-length assay kit (Affymetrix). RLM-RACE was performed using the FirstChoice RLM-RACE kit (Invitrogen) following manufacturer’s instructions except for the de-capping step which was performed using the RNA 5’ Pyrophosphohydrolase (RppH) enzyme (NEB) following manufacturer’s instructions.

**Telomere Restriction Fragment (TRF) analysis.** Genomic DNA was purified using DNAzol (Molecular Research Centre, MRC) following liquid nitrogen homogenization on a micro tissue homogenizer (Kimble biomasher II). For each reaction, 5 µg genomic DNA was digested with 10 U each of *Alu*1 and *Rsa*1 at 37 °C overnight. Digested DNA was electrophoresed on a 25 cm long 1.5 % agarose/1X TAE gel along with 1 kb plus DNA ladder (NEB) pre-mixed with gel green dye (Biotium) at a constant 35 V for 14 hrs at 4 °C. The gel was imaged under blue light and denatured in 1.5 M NaCl/0.5 N NaOH for 30 mins followed by neutralization in 1.5 M NaCl/0.5 M Tris-HCl pH 7.0 for 30 mins. The gel was dried at 50 °C for 1 hour and pre-hybridized in hybridization solution (5X SSC, 5X Denhardt’s reagent, 0.1% SDS, and 20 mM NaH2PO4) for at least 1 hour. The gel was then in-gel hybridized with 15 pmol of synthetic 5’ IRDye800 near infra-red fluorophore-labeled TTA(GGGTTA)₅ oligo (IDT) in 15 ml fresh hybridization solution at 45 °C overnight. The gel was then washed with 3X SSC at room temperature for 20 min twice, and once
with 3X SSC/0.1% SDS at 55 °C for 20 min and scanned using a Typhoon NIR-IP imager at the slow scan setting. The image was resized and overlayed on the gel green-fluorescent image to obtain marker bands and analyzed using the WALTER tool (7).

**Enzymatic treatments of RNA.** Calf intestine alkaline phosphatase (CIP) from NEB, Terminator 5'-Phosphate-dependent exonuclease from Lucigen and T4 polynucleotide kinase from NEB were used following manufacturer’s instructions.

**5' cap-specific immunoprecipitation.** Thirty µg of *U. maydis* total RNA was preincubated with 3 µg anti-m⁷G (MBL Life Science, RN016M) / anti-TMG (Santa Cruz, K121) / IgG (Sigma I3181) antibody in 1x RNA IP buffer (25 mM Tris-HCl, pH 8.0, 200 mM NaCl, and 0.2 mM EDTA) at 4°C for 1.5 hr. The RNA-antibody mixtures were transferred to 20 µl of pre-washed Protein A/G agarose beads (Pierce #20421) in 1x RNA IP buffer and mixed at 4°C for 1.5 hrs. The beads were then washed three times with 1x RNA wash buffer (25 mM Tris-HCl, pH 8.0, 300 mM NaCl, and 0.2 mM EDTA). Immunoprecipitated RNA was purified by acid phenol/chloroform extraction and ethanol precipitation, and resuspended in 20 µl H₂O. The RNA was directly used for quantitative reverse transcription PCR.

**Reverse Transcription quantitative PCR (RT-qPCR).** DNase-treated total RNA was treated with specific enzymatic reaction and then used in RT reaction to make cDNA library using random nonamer or gene-specific primer with ProtoScript II (NEB) reverse transcriptase following manufacturer instructions. The resulting cDNA was PCR amplified with PowerUp SYBR Green MasterMix (Thermo Fisher) on a StepOnePlus real-time PCR system (Applied biosystems). All the real-time PCR primers were validated to produce a single specific PCR product by 3% agarose gel electrophoresis. Technical triplicates were performed for each set of real-time PCR. RNA abundance levels were calculated using ∆∆Ct method (8) by the StepOnePlus software. Prism (GraphPad Software) was used to perform statistical analysis.

**Northern blot analysis.** Ten µg of total RNA was resolved on an agarose-formaldehyde gel using a pKa-matched buffer system as described previously (9). Following electrophoresis, gel was capillary transferred onto a Hybond-XL membrane (Cytiva) for 3 hours and RNA was UV crosslinked to the membrane under optimal crosslink setting on a Spectrolinker XL-1500 UV crosslinker (Spectroline). The membrane was pre-hybridized in the ULTRAhyb ultrasensitive hybridization buffer (Invitrogen) at 65°C for 45 mins and hybridized overnight with a ³²P radiolabeled riboprobe. The blot was washed with 1X SSC solution (150 mM sodium chloride and 15 mM sodium citrate) with 0.2% SDS at 65°C twice for 10 min each and once with 0.2X SSC / 0.2% SDS for 30 min, exposed to a phosphor-storage screen overnight and imaged on an FX Pro phosphoimager (BioRad). For ³²P end-labeled oligonucleotide probes, pre-hybridization and hybridization were performed using ULTRAhyb-Oligo hybridization buffer (Invitrogen) at 42°C. Washes were also performed at 42°C.

**Generation of *U. maydis* recombinant strains.** The *U. maydis* recombinant strains were generated via protoplast transformation as described previously (10) with minor modifications. Briefly, a single *U. maydis* colony was grown in YPD (Yeast extract, peptone, dextrose) media and harvested between OD₆₀₀ of 0.2-0.6. Approximately 75 million cells were centrifuged and
washed once with SCS buffer (20 mM Sodium citrate pH 5.8 and 1 M Sorbitol). Cells were resuspended in 1 ml SCS buffer containing 12.5 mg/ml lysing enzymes from Trichoderma harzianum (Millipore Sigma – L1412) and incubated at 30°C for approximately 5-20 minutes with mixing or until ~30-40% of the cells have been protoplasted (pinheads) as estimated under light microscopy. Enzyme-treated samples were centrifuged at 1000 xg for 10 min at 4°C and washed twice with cold SCS buffer followed by a wash with STC buffer (10 mM Tris-HCl pH 7.5, 0.1 M CaCl₂, and 1 M Sorbitol). Protoplasts were resuspended at ~2×10⁸ protoplasts / ml in cold STC buffer. A 50 µl aliquot of protoplasts was mixed with 15 µg heparin and 5 µg plasmid DNA or linear DNA to a total volume no more than 60 µl and incubated on ice for 10 mins. Five hundred microliter of STC with 40% W/V PEG4000 was added to the protoplast, heparin and plasmid DNA mixture and incubated for an additional 15 mins. The entire mixture was spread onto a dual layered regeneration agar plate. The top layer is 1/2 ml YEPS agar (2% peptone, 1% yeast extract, 2% sucrose, 1 M sorbitol and 1.5% agar) poured freshly immediately before plating and the bottom layer is 12 ml YEPS agar supplemented with 400 µg/ml Hygromycin or 4 µg/ml Carboxin. Plates were incubated at 28°C for 5-7 days until transformants appear. Isolated single transformant colonies were spotted onto fresh YEPS agar plates and grown at 28°C for 2 days or until colonies were big enough for colony PCR. Colony PCR of U. maydis was performed as described (11). Successfully verified colonies were grown in YEPS liquid culture and cryopreserved.

Cloning of U. maydis TER and TERT. The UmTER sequence was PCR amplified from genomic DNA and cloned into the EcoRV digested TOPO vector. The plasmid with the cloned UmTER was verified by Sanger sequencing. To obtain the UmTERT coding sequence, 2 µg DNase I-treated total RNA was annealed with 100 pmol oligo d(T)₂₅ DNA primer in the presence of 10 mM dNTPs (each) by incubation at 65°C for 3 min and quickly cooled on ice. A 20 µl reverse-transcription reaction was performed with the annealed RNA in 1X Protoscript II Reverse Transcriptase Reaction buffer (50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl₂, 10 mM DTT and 200 units ProtoScript II Reverse Transcriptase, NEB) at 48 ºC for 30 min. The reaction was terminated by incubation at 80 ºC for 5 min. The first-strand cDNA was used as a template for PCR, performed in a 25 µl reaction template in 1X Q5 Reaction buffer [50 mM Tris-Cl (pH 8.3), 2mM DTT, 0.5mM MgCl₂ and 1mM β-mercaptoethanol], 0.2 mM of each dNTP, 0.5 unit of Q5 DNA Polymerase, 0.2 µM forward and reverse primer each. PCR products were resolved on a 1% agarose gel to verify size. Restriction digested PCR DNA was ligated to pCITE-3xFLAG vector and UmTERT coding sequence was verified by Sanger sequencing.

Telomerase in vitro reconstitution and direct primer-extension assay. Recombinant 3xFLAG-UmTERT protein was in vitro synthesized in rabbit reticulocyte lysate (RRL) using the TnT Quick Coupled Transcription/Translation System (Promega). In vitro transcribed full-length or fragments of UmTER were gel purified and assembled in RRL with the synthetic 3xFLAG-UmTERT protein. The mixture was incubated at 30°C for 30 min. The telomerase primer-extension assay was carried out with 2 µl of in vitro reconstituted telomerase in a 10 µl reaction in 1X PE buffer [50 mM Tris–HCl (pH 8.3), 2mM DTT, 0.5mM MgCl₂ and 1mM spermidine], 1mM dTTP, 1 mM dATP, 5 mM dGTP, 0.165 mM α-³²P-dGTP (3000 Ci/mmol, 10 mCi/ml, PerkinElmer) and 1 µM telomeric primer. The reactions were incubated at 30°C for 60 min and terminated by phenol/chloroform extraction, followed by ethanol precipitation. Telomerase-extended products
were resolved on a denaturing 8M urea/10% polyacrylamide gel. The dried gel was exposed to a phosphor storage screen and analyzed with a Typhoon NIR-IP imager.

**Library preparation for Nanopore sequencing.** Total RNA integrity and concentration was determined using Agilent Tapestation 4200 at the Arizona State University, Genomics core facility. Reverse transcription was carried out using 65 ng of total RNA with gene specific oligo annealing downstream (precursor only library) or upstream (mature and precursor library) of the 3’ end of mature TER using 100 Units Maxima H minus RT (Invitrogen) in a 10 µl reaction at 42 °C for 1.5 hours in the presence of strand switching primer (SSP-Oxford Nanopore technologies). For generating the library to analyze 3’ ends of recombinant UmTER transcripts, G/I-tailing was performed on total RNA using reagents from a USB poly(A) tail length assay kit (Affymetrix). Reverse transcription was performed using an oligo-dC reverse primer with SSP in the reaction. An aliquot of the RT reaction was diluted 10-fold and 1 µl was used as template in a 10 µl PCR reaction with 0.1 µM primers, 2.5 mM each dNTPs using LongAmp Taq 2X master mix (NEB). For preparing a precursor-only library, an aliquot of this PCR reaction was further 10-fold diluted and 1 µl used in PCR reaction with 1 µM primers using the same conditions as above. Resultant PCR DNA was purified using a DNA clean and concentrator kit (Zymo research) using the manufacturer’s instructions and quantified using Agilent Tapestation 4200. Purified PCR DNA (250 pg) was used for Nanopore library preparation using SQK-PCS109 kit (Oxford Nanopore Technologies) following the manufacturer’s instructions beginning at the “Selecting for full-length transcripts by PCR” step. The resultant library was quantified using Agilent Tapestation 4200 followed by Rapid adapter addition and loaded onto a primed FLO-MIN106D flow cell and sequenced using a miniION device with data collection using the MinKNOW interface.

**Construction of recombinant UmTER precursor expression cassette.** The hsp70 promoter and 5’ UTR sequence, *U. maydis* glyceraldehyde-3-phosphate dehydrogenase (GAPDH) partial ORF, Umter including 200 bp upstream and 264 bp downstream flanking sequences and hsp70 terminator were PCR amplified from *U. maydis* genomic DNA using primers containing compatible overhangs with adjacent segments. PCR reaction conditions described above (“Cloning of *U. maydis* TER and TERT” section) for Q5 DNA polymerase (NEB) were used. Individual PCR fragments were gel extracted and purified using the Wizard SV gel and PCR cleanup kit (Promega) and used in overlap extension PCR using primers flanking the entire cassette. The assembled cassette was gel purified, digested with KpnI (NEB) and Afl II (NEB), and ligated to pPotef-LN22*-3xGNLS-cbx plasmid digested with the same enzymes (12). The expression cassette sequence was fully verified to be free of mutations by Sanger sequencing. The plasmid was linearized with Sspl (NEB) restriction enzyme and purified by phenol/chloroform extraction and ethanol precipitation prior to transformation of *U. maydis* cells.

**Phylogenetic, sequence and structural analysis of Emi1 and homologs.** Multiple sequence alignment was performed using the T-COFFEE web server’s M-Coffee aligner (13). Minor manual adjustments were made to the alignments with BioEdit. Alignment figure was generated using the ESPript 3.0 utility (14). Protein structure prediction was performed using Alphafold2 (15) with default parameters on the Agave high-performance computing cluster of Arizona State University. The highest confidence model was chosen for structural comparisons. Phylogenetic tree construction was performed with MEGA X (16) software using the Maximum likelihood method.
and Le-Gascuel model as recommended by MEGA X (17). Bootstrap support values were inferred from 200 replicates and expressed as percentage values (18).

**Nanopore long-read sequencing data analysis.** Raw fast5 data was base-called on a RTX2080 GPU using the high accuracy model (configuration file - dna_r9.4.1_450bps_hac.cfg) using guppy version 5.0.16 (Oxford Nanopore technologies). Quality control, adapter trimming, read orientation and classification of raw fastq reads was performed using Pychopper v2's cdna_classifier.py script (Oxford Nanopore Technologies). Reads were length filtered where applicable using fastp (19). Reads that pass the quality score threshold were mapped to the reference UmTER locus using minimap 2 (20) followed by transcript assembly using StringTie 2 (21). Isoform relative abundance estimation was performed using liqa (22). Alignments were processed using Samtools suite (23). Alignments were visualized using Tablet (4) and Integrative Genomics Viewer (IGV) (24). Nucleotide position-specific read counts were obtained using Tablet.

**Sample preparation for proteomics analysis.** For detection of native Emi1 from U. maydis cell lysate, 800 µg of soluble total protein as determined by Bradford assay was mixed with 4X Bolt LDS sample buffer (ThermoFisher) to a final concentration of ~1.2 X supplemented with 90 mM DTT and incubated at 70 °C for 10 mins. A sample volume of 55 µl was loaded per well for a total of 6 wells onto a Bolt 12%, Bis-Tris 10 well gel (ThermoFisher). Electrophoresis was performed in 1X Bolt MES running buffer until the dye front reached close to the end of the gel. Gel was stained immediately for 1 hour using One-Step Blue protein gel stain (Biotium) followed by destaining in Millipore water for 1 hour. Gel slices were excised inside a clean bench using a new blade and immersed in 0.1% formic acid solution until processing. From ~900 µg of total soluble protein as input, anti-FLAG immunoprecipitation was performed using anti-FLAG M2 magnetic beads (Sigma-M8823) according to the manufacturer’s instructions, and beads were treated with 2X Bolt LDS sample buffer without DTT supplementation. Samples were placed on a magnet and pipetted to fresh tubes for gel loading and excision of gel slices as described above. Gel bands were washed in 100 mM Ammonium Bicarbonate (AmBic) / Acetonitrile (ACN) and reduced with 10 mM dithiothreitol at 50 °C for 45 minutes. Cysteines were alkylated with 100 mM iodoacetamide in the dark for 45 minutes in room temperature (RT). Gel bands were washed in 100mM AmBic / ACN prior to adding 1 µg trypsin (Promega #V5111) for overnight incubation at 37°C. Supernatant-containing peptides were collected into a new tube. Gel pieces were washed with gentle shaking in 50% ACN / 1% Formic acid (FA) at RT for ten minutes, and supernatant was collected in the previous tubes. The final peptide extraction step was done with 80% ACN / 1% FA, and 100% ACN, and all supernatant was collected. The peptides were dried in a SpeedVac and reconstituted with 5% ACN / 0.1% FA in water before injecting into the LC-MS/MS.

**LC-MS / MS Analysis.** Peptides were analyzed by LC-MS / MS using a Dionex UltiMate 3000 Rapid Separation nanoLC coupled to a Orbitrap Elite Mass Spectrometer (Thermo Fisher Scientific Inc, San Jose, CA). Samples were loaded onto the trap column, which was 150 µm x 3 cm in-house packed with 3 um ReproSil-Pur® beads. The analytical column was a 75 µm x 10.5 cm PicoChip column packed with 3 um ReproSil-Pur® beads (New Objective, Inc. Woburn, MA). The flow rate was kept at 300 nL/min. All fractions were eluted from the analytical column at a flow rate of 300 nL/min using an initial gradient elution of 5% B from 0 to 5 min, transitioned to 40% over 100 min, 60% for 4 mins, ramping up to 90% B for 3 min, holding 90% B for 3 min,
followed by re-equilibration of 5% B at 10 min with a total run time of 120 min. Mass spectra (MS) and tandem mass spectra (MS/MS) were recorded in positive-ion and high-sensitivity mode with a resolution of ~60,000 full-width half-maximum. The 15 most abundant precursor ions in each MS1 scan were selected for fragmentation by collision-induced dissociation (CID) at 35% normalized collision energy in the ion trap. Previously selected ions were dynamically excluded from re-selection for 60 s. The collected raw files spectra were stored in raw format.

**Mass spectrometric data analysis.** Proteins were identified from the MS raw files using the Mascot search engine (Matrix Science, London, UK. version 2.5.1). MS/MS spectra were searched against the Ustilago-maydis_3custom_20220523 database. All searches included carbamidomethyl cysteine as a fixed modification and oxidized methionine, deamidated asparagine and aspartic acid, and acetylated N-term as variable modifications. Three missed tryptic cleavages were allowed. A 1% false discovery rate cutoff was applied at the peptide level. Only proteins with a minimum of two peptides above the cutoff were considered for further study. Identified peptides/protein were visualized by Scaffold software (version 5.0, Proteome Software Inc., Portland, OR).

**References**

1. M. Kojic, C. F. Kostrub, A. R. Buchman, W. K. Holloman, BRCA2 homolog required for proficiency in DNA repair, recombination, and genome stability in Ustilago maydis. *Mol Cell* **10**, 683-691 (2002).
2. X. Qi, Y. Li, S. Honda, S. Hoffmann, M. Marz, A. Mosig, J. D. Podlevsky, P. F. Stadler, E. U. Selker, J. J.-L. Chen, The common ancestral core of vertebrate and fungal telomerase RNAs. *Nucleic Acids Res* **41**, 450-462 (2013).
3. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).
4. I. Milne, M. Bayer, L. Cardle, P. Shaw, G. Stephen, F. Wright, D. Marshall, Tablet--next generation sequence assembly visualization. *Bioinformatics* **26**, 401-402 (2010).
5. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
6. T. A. Hall, BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Sym Ser* **41**, 95-98 (1999).
7. M. Lycka, V. Peska, M. Demko, I. Spyroglou, A. Kilar, J. Fajkus, M. Fojtova, WALTER: an easy way to online evaluate telomere lengths from terminal restriction fragment analysis. *BMC Bioinformatics* **22**, 145 (2021).
8. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408 (2001).
9. F. H. Mansour, D. G. Pestov, Separation of long RNA by agarose-formaldehyde gel electrophoresis. *Anal Biochem* **441**, 18-20 (2013).
10. T. Tsukuda, S. Carleton, S. Fotheringham, W. K. Holloman, Isolation and characterization of an autonomously replicating sequence from Ustilago maydis. *Mol Cell Biol* **8**, 3703-3709 (1988).
11. K. Bosch, L. Frantzeskakis, M. Vranes, J. Kamper, K. Schipper, V. Gohre, Genetic Manipulation of the Plant Pathogen Ustilago maydis to Study Fungal Biology and Plant Microbe Interactions. *J Vis Exp* **115**, 54522 (2016).
12. T. Pohlmann, S. Baumann, C. Haag, M. Albrecht, M. Feldbrugge, A FYVE zinc finger domain protein specifically links mRNA transport to endosome trafficking. *Elife* 4 (2015).

13. P. Di Tommaso, S. Moretti, I. Xenarios, M. Orobitg, A. Montanyola, J. M. Chang, J. F. Taly, C. Notredame, T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res* 39, W13-17 (2011).

14. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42, W320-324 (2014).

15. J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-589 (2021).

16. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* 35, 1547-1549 (2018).

17. S. Q. Le, O. Gascuel, An improved general amino acid replacement matrix. *Mol Biol Evol* 25, 1307-1320 (2008).

18. J. Felsenstein, Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* 39, 783-791 (1985).

19. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884-i890 (2018).

20. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094-3100 (2018).

21. S. Kovaka, A. V. Zimin, G. M. Pertea, R. Razaghi, S. L. Salzberg, M. Pertea, Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol* 20, 278 (2019).

22. Y. Hu, L. Fang, X. Chen, J. F. Zhong, M. Li, K. Wang, LIQA: long-read isoform quantification and analysis. *Genome Biol* 22, 182 (2021).

23. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, S. Genome Project Data Processing, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079 (2009).

24. J. T. Robinson, H. Thorvaldsdottir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer. *Nat Biotechnol* 29, 24-26 (2011).

25. E. Y. Yu, M. Hsu, W. K. Holloman, N. F. Lue, Contributions of recombination and repair proteins to telomere maintenance in telomerase-positive and negative Ustilago maydis. *Mol Microbiol* 107, 81-93 (2018).

26. T. L. Bailey, C. Elkan, Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 2, 28-36 (1994).
Fig. S1. Purification of active *U. maydis* telomerase holoenzyme. **(A)** Schematic showing immunoprecipitation (IP) of ectopically expressed 3xFLAG-*UmTERT* bound to cellular telomerase RNA. Co-immunoprecipitated RNA are extracted and subject to Illumina next-generation sequencing. Bioinformatics analysis of sequencing data is performed to determine TER candidates. **(B)** Anti-FLAG Western blot of the soluble fraction of lysates from *U. maydis* expressing 3xFLAG-*UmTERT* or wild-type (WT) cells. **(C)** Telomerase activity detected by TRAP assay after IP using anti-FLAG antibody-conjugated beads. Telomerase was treated with or without RNase A. Human cell lysate was included as a positive control.
Fig. S2. Determining the 5’ and 3’ ends of UmTER using RACE analyses. (A) Schematic for cap-independent template-switching reverse transcriptase (TSRT)-mediated 5’ RACE of UmTER. The UmTER-specific reverse primer, RT-R, was used by TSRT for reverse transcription, which adds non-templated nucleotides to the 3’ end of first strand cDNA. The presence of a template switching oligo (TSO) in the RT reaction allows for further addition of a known sequence at the 3’ end of first strand cDNA, followed by PCR, cloning and sequencing to determine the 5’-end nucleotide identity of UmTER. (B) Agarose gel electrophoresis of 5’ RACE products. PCR was performed for RT (+) or RT (-) reactions using forward primer (F) annealing to the TSO sequence and a reverse primer (R1) annealing to UmTER region of cDNA followed by gel electrophoresis. (C) Sanger sequencing chromatogram of 5’ RACE PCR product. Sanger sequencing chromatogram sequences are compared to corresponding genomic sequences with sequence elements indicated below the chromatogram. (D) Schematic showing the G/I-tailing mediated 3’ RACE of UmTER. U. maydis total RNA G/I tailed, and reverse transcribed using an oligo dC reverse primer
(Oligo dC-R) containing a unique 5’ adapter sequence followed by PCR, cloning and sequencing to determine the 3’ nucleotide identity of UmTER. (E) Agarose gel electrophoresis of 3’ RACE products. PCR was performed for RT (+) or RT (-) reactions using forward primers (F0 or F1) that anneal to the UmTER region and reverse primer (R) that anneals to the adapter sequence of the oligo dC-R primer region of the cDNA. (F) Sanger sequencing chromatogram of 3’ RACE PCR product. (G) Schematic showing co-IP of 3xFLAG-tagged UmTERT protein bound RNA. Co-immunoprecipitated UmTER was analyzed by 5’ and 3’ RACE assays. (H) Agarose gel electrophoresis of PCR products of 5’ RACE was performed on TERT bound TER for RT (+) or RT (-) reactions. (I) Sanger sequencing chromatogram of TERT bound TER specific 5’ RACE PCR product. (J) Sanger sequencing chromatogram of TERT bound TER specific 3’ RACE PCR product. (K) Comprehensive analysis of mature UmTER 3’ end. Mature UmTER 3’ region nucleotides and their position are shown to the left with the dominant 3’ end determined based on 3’ RACE highlighted in red. Number of 3’ RACE clones derived from total RNA (grey bars) or TERT bound TER (red bars) corresponding to each 3’ end nucleotide is plotted in a horizontal bar plot to the right.
**Fig. S3.** Conserved gene synteny of the *Umter* gene locus in Class Ustilaginomycetes. Evolutionary relationships between the species and respective orders are shown to the left. Branch length is not proportional to evolutionary distance. Schematic of the gene synteny of the *ter* gene locus in Ustilaginomycetes species with coding sequences flanking *ter* and their transcription orientations shown to the right. The open reading frame (ORF) immediately 5' of *ter* (red) is of Early meiotic induction protein (*emi1*) (orange). Homologous sequences were found in all species. A putative Zn finger protein gene ORF (green) and a CAR2-ornithine amino-transferase protein ORF (*oat*) shown (purple). Scale bar shown on bottom right.
The UmTER gene is essential for telomere maintenance. **(A)** Gene-knockout strategy of Umter. The Umter gene is shown as blue arrow with flanking regions shown in grey. The construct that contains the cassette which provides hygromycin resistance following successful gene replacement is shown as black arrow. Right and left flanks are shown in grey with lengths of respective elements indicated below each element. **(B)** Knockout validation via northern blot. (top) Northern blot of total RNA isolated from either WT or ΔUmter clones using UmTER-specific oligonucleotide probe. Marker band size is indicated on the left and UmTER band shown with a red arrowhead. (middle) trt1 specific probing as internal control. (bottom) Ethidium bromide staining of agarose gel prior to transfer, with 28S and 18S rRNA bands indicated to the left. **(C)** Telomere Restriction Fragment (TRF) analysis of wild-type (WT) and two ΔUmter clones (C1, C2). Strains were propagated and harvested for genomic DNA isolation at indicated generations (Gen). The numbers of generation were estimated based on approximately 11 doubling/day for WT and 8 doubling/day for the ΔUmter strains, similar to that of the trt1Δ strain (25). Position and marker band sizes are indicated to the left and right of the gel. **(D)** Box and whiskers plot of restricted telomere spread corresponding to each lane from (C).
Fig. S5. *In vitro* validation of UmTER and identification of the minimal UmTER regions essential for telomerase activity. Numbers to the right or left of a gel indicate nucleotides added to the primer. (A) *In vitro* reconstitution of *U. maydis* telomerase. (*top*) Sequence of the UmTER template (open box) with the annealing positions of six permuted telomeric DNA primers (1-6) shown. Predicted primer extended products are shown in lowercase and aligned with the template. Telomerase synthesized radiolabeled ‘dG’ nucleotides are shown in lowercase red. (*bottom*) Direct primer-extension assay of *in vitro* reconstituted telomerase from synthetic UmTERT and TR. A $^{32}$P end-labeled oligonucleotide is added as the recovery control (r.c.) to the reaction prior
to ethanol precipitation of DNA products. (B) Truncation analysis of *UmTER*. *(top)* Schematic of full-length *UmTER* and nucleotide ranges of the 5' (a1-a5) and respective 3' (b1-b5) fragments shown. Activity levels of fragment combinations compared to *UmTER* full-length RNA (FL) is shown to the right with (+) indicating equal or higher activity than FL while (-) indicating lower activity. *(bottom)* Activity assay of *UmTER* 5' and respective 3' fragment combinations reconstituted with synthetic *UmTERT* with FL included as a control. *(C)* Serial truncation analysis of *UmTER* 5' and 3' fragments. *(left)* Schematic showing the 5' and 3' fragments 5F1 (a5) and 3F1 (b5) respectively, were individually terminally truncated further to generate the 5F and 3F fragments. Level of activity of each fragment combination is shown to the right. Nucleotide ranges of minimal 5' and 3' fragments based on fine terminal truncations are shown at the bottom. *(right)* Activity assay of telomerase reconstituted from the *UmTER* 5F and 3F truncated fragments in various combinations compared to FL. *The 3' end position was revised to be 1291 following comprehensive 3' RACE analysis on *UmTER*. This does not affect the conclusions derived from this figure.*
Fig. S6. Multiple sequence alignment of 18 TERs from Order Ustilaginales of Phylum Basidiomycota. Multiple sequence alignment was performed initially using the ClustalW algorithm and then improved manually in the BioEdit program. Highly conserved regions and motifs were aligned first followed by alignment of intervening sequences using conserved regions as anchors. Individual nucleotides are colored by identity (A; green, G; black, U; red, C; blue) and nucleotides that are conserved in ≥ 75% of given fungal species are shaded (White text on colored background). The template and base-paired helices in the secondary structures are denoted within white boxes above the alignment. Number of intervening nucleotides between TER elements are indicated for each TER in the alignment. (A) Alignment of the template pseudoknot region. (B) Alignment of the Conserved region 4/5 (CR4/5).
**Fig. S7.** Promoter sequence alignment of TER precursors from 17 Ustilaginomycetes species. The 225 bp region upstream of the start codon of *emi1* homologs were extracted. The MEME tool from the MEME (Multiple Em for Motif Elicitation) suite was used to identify shared sequence motifs from the extracted sequences (26). The highest confidence motif was aligned using the BioEdit. Individual nucleotides are colored by identity (A; green, G; black, U; red, C; blue) and nucleotides that are conserved in ≥80% of given fungal species are shaded (White text on colored background). Number of nucleotides intervening the putative transcription start site (TSS) and aligned motif is indicated and the start codon is shown. *U. maydis* TER precursor’s TSS was experimentally determined.
Fig. S8. Canonical intron defining elements in alternatively spliced UmTER precursor isoforms and U2 snRNA secondary structural model. (A) Sequence and relative positions of intron defining sequence elements in UmTER precursor isoforms B-E. Exons 1 and 2 are shown as green boxes. The 5’ and 3’ splice sites are shown with red and blue arrows respectively. Highly conserved sequences flanking the 5’ and 3’ splice junctions are shown in red and less conserved sequences shown in blue. Branch point “adenosine” is shown in red and indicated with an arrow with flanking residues required for base pairing with the U2 snRNA shown in open box. Lengths between intron defining sequence elements are indicated for each isoform. (B) The secondary structural model of U. maydis U2 snRNA (EnsemblFungi Transcript ID – ENSRNA0496731139-T1) is shown. The U2 sequence that base-pairs with the branch site sequence in the intron is shown in blue and boxed.
**Fig. S9.** The *UmTER* precursor has a 3’ poly A tail. (top) Sanger sequencing chromatogram of 3’ RACE sequence compared to corresponding genomic sequence with sequence elements indicated below the chromatogram. (bottom) Splicing configurations of *UmTER* precursor isoforms B-E. Isoform A is un-spliced while isoforms C, D have alternative 5’ splice sites but the same 3’ splice site. Isoforms B and E share alternative 5’ splice sites but the same 3’ splice site yet distinct from the 3’ splice site shared between C and D. Nucleotide positions of precursor and mature *UmTER* indicated below the transcript. The presence of a 3’ poly(A) tail is shown.
**Fig. S10.** Generation of mutant UmTER recombinant expression cassette and validation of successful genome integration. Lengths of each element in base pairs indicated below where applicable. (A) Schematic showing organization of recombinant expression cassette. The cassette is regulated by the hsp70 promoter (P\textsubscript{hsp70}) and terminator (T\textsubscript{hsp70}). In addition, the first 390 bp of the *U. maydis* GAPDH protein-coding sequence (NCBI RefSeq – XM_011390465.1) was fused to the hsp70 promoter, followed by the mature UmTER sequence with 200 bp of upstream and 264 downstream flanking sequences. The recombinant UmTER includes a marker sequence (Marker, purple) for northern blot probing. The template is shown in red. The complete cassette is cloned into the genome integration plasmid pPotef-LN22*-3xGNLS-cbx that confers carboxin resistance (ip\textsuperscript{r}) for selection in *U. maydis*. Genomic integration site (ip\textsuperscript{s}) and unique linearization restriction site (SspI) are shown. Primer annealing positions, F and R, used for recombinant screening via PCR are shown. (B) Agarose gel electrophoresis of recombinant integration screening via PCR. Wild-type genomic DNA is included as a negative control. (C) The organization of the wild-type UmTER genomic locus is shown. Wild-type UmTER-specific northern blot probe P (WT) annealing region is indicated.
**Fig. S11.** Sequence and structural homology analysis of UmTER precursor-encoded protein Emi1. For multiple sequence alignments, species are ordered by sub-phyla and the respective subphylum is shown to the left. The first and last amino acid positions of each sequence are shown to the left and right of the sequence respectively. The consensus sequence is shown in the bottom row of each alignment and defined as having >60% sequence identity for a given amino acid position. Positions that are 100% conserved are highlighted in red background with white text and amino acids that share >60% similarity for a given position are colored in red and shown in blue open box. Amino acids corresponding to Alphafold 2 predicted alpha helices are indicated above the alignment with horizontal lines and the alpha helix numbers are shown above the line. (A) Multiple sequence alignment of 17 Emi1 homologs from Phylum Basidiomycota representative species. (B) Multiple sequence alignment of Emi1 protein (GenBank:ONH79050.1) and 13 homologs from Phylum Ascomycota representative species. (C) Phylogenetic tree constructed from cumulative multiple sequence alignment of Emi1 homologs. Evolutionary analysis was performed by the Maximum Likelihood method and evolutionary history was inferred using the Maximum Likelihood method and Le_Gascuel_2008 model. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site as shown by the scale bar. Bootstrap support values inferred from 200 replicates are shown as a percentage for corresponding nodes. (D) Alphafold2 predicted structural model of *U. maydis* Emi1 protein for the amino acid range 10-122. Alpha helices α1-α5 are labeled with residues conserved between both Basidiomycota and Ascomycota species highlighted and labeled in red. (E) Alphafold2 predicted structural model of *Saccharomyces cerevisiae* s288c Emi1 protein (Uniprot ID – Q04406) for the amino acid range 80-187 as deposited in the Uniprot database. Alpha helices α1-α5 are labeled with residues conserved between both Basidiomycota and Ascomycota species highlighted and labeled in red. (F) Amino acid sequence comparison of *U. maydis* Emi1 isoforms. Alphafold2 predicted helices shown above the comparison. The length of each ORF indicated to the right and positions numbered at the bottom.
Fig. S12. Tandem mass spectra of (A) native Emi1- and (B) 3xFLAG-Emi1-specific tryptic peptide fragment ESAENVVELRR. The b- and y- ion fragments and their derivative peaks are labeled.
Fig. S13. TERs from Class Ustilaginomycetes share a putative Sm binding site proximal to the 3’ end. Multiple sequence alignment of 5’ region, template, and 3’ regions of class Ustilaginales species telomerase RNAs. Residues conserved 100% are shaded with white text on a colored background. The number of intervening nucleotides between sequence elements is shown. Start and end positions experimentally determined for U. maydis TER shown above the alignment. The estimated lengths of each TER are shown to the right. Template and the putative Sm binding site are indicated above the alignment.
| Subphylum          | Species/ Strain               | Protein ID            | GenBank accession no. | Length |
|-------------------|-------------------------------|-----------------------|-----------------------|--------|
| Agaricomycotina   | *Pleurotus ostreatus*         | Uncharacterized protein PC9H_009394 | XP_036628287.1*       | 132    |
|                   | *Hygrophoropsis aurantica*   | hypothetical protein BJ138DRAFT_1014460 | KAH7907454.1          | 130    |
|                   | *Suillus bovinus*             | uncharacterized protein EDB93DRAFT_1177345 | XP_041302598.1*       | 127    |
|                   | *Steccherinum ochraceum*      | hypothetical protein EIP91_003540 | TCD64874.1            | 133    |
|                   | *Amanita poly pymaris / BW_CC*| hypothetical protein AX15_005731 | KAF8624851.1          | 125    |
|                   | *Rhizoctonia solani*          | unnamed protein product | CAE6502813.1          | 137    |
|                   | *Cryptococcus floricola*      | hypothetical protein B9479_000077 | TYJ59088.1            | 167    |
|                   | *Filobasidium floriforme*     | uncharacterized protein HD553DRAFT_308342 | XP_046038264.1*       | 175    |
| Pucciniomycotina  | *Mixia osmundae / IAM 14324*  | hypothetical protein L969DRAFT_92007 | XP_014571169.1*       | 126    |
|                   | *Atractiella rhizophila*      | hypothetical protein BT69DRAFT_1261676 | KAH8924231.1          | 120    |
|                   | *Microbotryum intermedium*    | BQ2448_3206           | SCV71618.1            | 141    |
|                   | *Rhodotorula toruloides*      | hypothetical protein Rl10032_c14g4500 | GEM10483.1            | 149    |
|                   | *Leucosporidium creatinivorum*| hypothetical protein BCR35DRAFT_288241 | ORY88967.1            | 160    |
| Ustilaginomycotina| *Urocystis primulicola*       | Unannotated           | N/A                   | 110    |
|                   | *Violaceomyces palustris*     | hypothetical protein IE53DRAFT_370055 | PWN49096.1            | 128    |
|                   | *Testicularia cyprei*         | hypothetical protein BCV7DRAFT_217114 | PWZ00901.1            | 128    |
|                   | *Ustilago maydis*             | hypothetical protein UMAG_03168 | KIS68597.1            | 134    |
|                   | *Sporisorium scitamineum*     | uncharacterized protein SPSC_04188 | CDR88361.1            | 133    |

* RefSeq
N/A – Not Available
| Subphylum                        | Species / Strain                      | Protein ID                  | GenBank accession no. | Length |
|---------------------------------|--------------------------------------|----------------------------|-----------------------|--------|
| **Taphrinomycotina**            |                                      |                            |                       |        |
| Saitoella complicata            |                                      | uncharacterized protein    | XP_019024316.1*       | 167    |
| Schizosaccharomyces pombe /     |                                      | uncharacterized protein    | NP_592971.1*          | 121    |
| NRRL Y-17804                    |                                      | SPAC227.17c                |                       |        |
| Schizosaccharomyces cryophilus /|                                      | hypothetical protein       | XP_013021753.1*       | 112    |
| OY26                            |                                      | SPOG_04037                 |                       |        |
| **Pezizomycotina**              |                                      |                            |                       |        |
| Neurospora crassa               |                                      | hypothetical protein       | KHE78267.1            | 283    |
| Fusarium nygamai                |                                      | hypothetical protein       | PNP83368.1            | 224    |
| Fusarium tricinctum             |                                      | hypothetical protein       | KAH7256687.1          | 213    |
| Aspergillus fumigatus           |                                      | hypothetical protein       | KAH1871667.1          | 227    |
| Ascobolus immerses / RN42       |                                      | hypothetical protein       | RPA84169.1            | 193    |
|                                |                                      | hypothetical protein       |                       |        |
| **Saccharomycotina**            |                                      |                            |                       |        |
| Trichomonascus ciferri          |                                      | hypothetical protein       | KAA8907417.1          | 126    |
| Komagataella phaffii            |                                      | GQ67_05264T0               | AOA65092.1            | 128    |
| Cyberlindera jadinii            |                                      | Early meiotic induction    | CEP22764.1            | 86     |
| Kluveromyces lactis             |                                      | Early meiotic induction    | QEU60272.1            | 86     |
| Saccharomyces cerevisiae        |                                      | protein 1                  | ONH79050.1            | 83     |
| Lachancea meyersii / CBS 8951   |                                      | LAME_0G01178g1_1          | SCU98948.1            | 88     |

* RefSeq
Table S3. Mass to charge ratios (m/z) of predicted and observed fragmentation ions from Emi1-specific peptides. The m/z values of all predicted peptide ions derived from the three tryptic peptide fragments of the Emi1 protein are shown. The m/z values of the detected peptide fragment ions are indicated as bold and colored (red for -b ions and blue for -y ions). The underlined m/z values indicate the peptide fragment ions detected in both native Emi1 and the recombinant 3xFLAG-Emi1 protein samples. Ions that are not chemically possible are left blank. Double charged b/y ion (b/y<sup>2H</sup>), b/y ion with loss of ammonia (b/y<sub>NH3</sub>), or b/Y ion with loss of water (b/y<sub>H2O</sub>) are shown. The identities of the terminal amino acid (aa) of each peptide ion are indicated.

### ESAENVWELR

|   | b   | b ions     | b<sup>2H</sup> | b<sub>NH3</sub> | b<sub>H2O</sub> | aa | y ions    | y<sup>2H</sup> | y<sub>NH3</sub> | y<sub>H2O</sub> | y     |
|---|-----|------------|----------------|---------------|---------------|----|-----------|--------------|---------------|---------------|-------|
| 1 | 130.05 | 112.04    | E             | 1,232.59     | 616.8         | 1,215.56 | 1,214.58  | 10    |
| 2 | 217.08 | 199.07    | S             | 1,103.55     | 552.28        | 1,086.52 | 1,085.54  | 9     |
| 3 | 288.12 | 270.11    | A             | 1,016.52     | 508.76        | 999.49   | 998.51    | 8     |
| 4 | 417.16 | 399.15    | E             | 945.48       | 473.24        | 928.45   | 927.47    | 7     |
| 5 | 531.2  | 514.18    | N             | 816.44       | 408.72        | 799.41   | 798.43    | 6     |
| 6 | 630.27 | 315.64    | D             | 612.26       | 470.85        | 685.37   | 684.38    | 5     |
| 7 | 816.35 | 408.68    | V             | 702.39       | 365.51        | 586.3    | 585.31    | 4     |
| 8 | 945.39 | 473.2     | A             | 603.32       | 270.11        | 1,172.62 | 1,155.59  | 9     |
| 9 | 1,058.48 | 529.74    | L             | 288.2        | 209.08        | 1,084.55 | 1,083.57  | 8     |
| 10| 1,232.59 | 616.8     | R             | 175.12       | 158.09        | 158.09   |            |       |

### ESAENVWELRR

|   | b   | b ions     | b<sup>2H</sup> | b<sub>NH3</sub> | b<sub>H2O</sub> | aa | y ions    | y<sup>2H</sup> | y<sub>NH3</sub> | y<sub>H2O</sub> | y     |
|---|-----|------------|----------------|---------------|---------------|----|-----------|--------------|---------------|---------------|-------|
| 1 | 130.05 | 65.53     | E             | 1,388.69     | 694.85        | 1,371.67 | 1,370.68  | 11    |
| 2 | 217.08 | 109.04    | S             | 1,259.65     | 630.33        | 1,242.62 | 1,241.64  | 10    |
| 3 | 288.12 | 144.56    | A             | 1,172.62     | 586.81        | 1,155.59 | 1,154.61  | 9     |
| 4 | 417.16 | 209.08    | A             | 1,011.58     | 551.29        | 1,084.55 | 1,083.57  | 8     |
| 5 | 531.2  | 266.11    | N             | 972.54       | 486.77        | 955.51   | 954.53    | 7     |
| 6 | 630.27 | 315.64    | V             | 858.49       | 429.75        | 841.47   | 840.48    | 6     |
| 7 | 816.35 | 408.68    | W             | 759.43       | 380.22        | 742.4    | 741.42    | 5     |
| 8 | 945.39 | 473.2     | V             | 573.35       | 287.18        | 556.32   | 555.34    | 4     |
| 9 | 1,058.48 | 529.74    | L             | 443.3        | 222.66        | 427.28   | 427.28    | 3     |
| 10| 1,214.58 | 607.79    | 1,197.55     | 1,196.57     | 314.19        | 314.19   |            |       |
| 11| 1,388.69 | 694.85    | 1,371.67     | 1,370.68     | 158.09        | 158.09   |            |       |

### YEYQVAEDVAYHK

|   | b   | b ions     | b<sup>2H</sup> | b<sub>NH3</sub> | b<sub>H2O</sub> | aa | y ions    | y<sup>2H</sup> | y<sub>NH3</sub> | y<sub>H2O</sub> | y     |
|---|-----|------------|----------------|---------------|---------------|----|-----------|--------------|---------------|---------------|-------|
| 1 | 164.07 | 146.06   | Y             | 1,614.74     | 807.88        | 1,597.72 | 1,596.73  | 12    |
| 2 | 293.11 | 275.1    | E             | 1,451.68     | 726.34        | 1,434.65 | 1,433.67  | 11    |
| 3 | 421.17 | 404.15   | Q             | 1,322.64     | 661.82        | 1,305.61 | 1,304.63  | 11    |
| 4 | 584.24 | 567.21   | Y             | 1,194.58     | 597.79        | 1,177.55 | 1,176.57  | 10    |
| 5 | 683.3  | 666.28   | Y             | 1,031.52     | 516.26        | 1,014.49 | 1,013.51  | 9     |
| 6 | 754.34 | 377.67   | A             | 932.45       | 466.73        | 915.42   | 914.44    | 8     |
| 7 | 883.38 | 442.2    | E             | 861.41       | 431.21        | 844.38   | 843.4     | 7     |
| 8 | 998.41 | 499.71   | D             | 732.37       | 366.69        | 715.34   | 714.36    | 6     |
| 9 | 1,097.48 | 549.24   | V             | 617.34       | 309.17        | 600.31   | 599.33    | 5     |
| 10| 1,168.52 | 584.76   | A             | 518.27       | 259.64        | 501.25   | 500.26    | 4     |
| 11| 1,331.58 | 666.29   | Y             | 447.24       | 224.12        | 430.21   | 429.22    | 3     |
| 12| 1,468.64 | 734.82   | H             | 284.17       | 142.59        | 267.15   | 266.16    | 2     |
| 13| 1,614.74 | 807.88   | K             | 147.11       | 130.09        | 129.1    |            |       |
### Table S4. List of oligonucleotides and northern blot probes used in this study

| Oligo sequence (5'→3') | ID (ASU#) |
|------------------------|-----------|
| **Figure 3 A, B**<br>**RLM Race**<br>RLM RNA adaptor | GCUGAUGCGAUGAAGGACACUCGCUUGCUUGGUUGUUGAAGAA | 3442 |
| RLM RACE-F | GCTGATGGCGATGAATTACAG | 6091 |
| UmTER RACE-R1 | CGTCCGCTCAGACGTAAG | 6941 |
| UmTER RACE-R2 | GGTTCTGGATGCTTTTCTGAAATAC | 6942 |
| **Figure 3 C**<br>**RLM Race**<br>18s rRNA-R | GCATGTATTAGCTCTAGAATTACCACAG | 7359 |
| UmTR-R | CGTTCGCGTTCAGACGTAAG | 6941 |
| 5s rRNA-R | GGTTACCTAATACGACTCATATAGGAACTGCA | 6486 |
| **Figure 3 D, F**<br>**qPCR**<br>UmTR-F | GATGACGACTATCGGAGCAG | 7120 |
| UmTR-R | GCTGAGGTCTGAGCTTCTTCTG | 7121 |
| U2 snRNA-F | CTCAAGGCGAAGTGAAG | 7114 |
| U2 snRNA-R | TCTCAAGGCGAAGTGAAG | 7115 |
| GAPDH-F | TGCCAAGAAGGTCGTCATCTC | 7112 |
| GAPDH-R | GACATGGCTCAGGACACTC | 7113 |
| **Figure 3 E**<br>**Riboprobes**<br>P1 | AGCACUCCUCUCCCCACACUCCACUGCUUCCACUCUGCUUCUCUCUCGCAGGAAAUUUCAGAACGCUUUAACCCCUUUCAGGG | - |
| P2 | UACCGCAACUCCUCGCCCAACAUUGCUAGCAGACAGAUGUAAGGCGUUGAUGAAGGCGCAU | - |
| P3 | AGAUGCGUCUCCAAUGGAGGAGAGACUGGACUAGACGCUUUGUGCCCAGAC | - |
| P4 | GGGAGCCGCGAGACGUCUGGAGGAGGGAUGUGUGUUCCCACAGGAGCAUCUG | - |
| **Figure 4A**<br>Nanopore library prep<br>SSP | TTTCTGTTGTTGCTGATATTGCTmGmGmG | - |
| RT1 | ATTGCAGCATGTCCTC | 8113 |
| R | GACGAGCTAGCAGGCA | 8114 |
| **Figure 4 C**<br>F | TTTCTGTTGTTGCTGATATTGCTTGGTCAACCCGTTTTCGTAATGCTTC | 8293 |
| RT2 | GAGATGGAAATTTGC | 8227 |
| **Figure 4D**<br>Northern probes<br>P (WT) | CTGGTTTCGCTGAGACGCAGCTTCAAACCACTCTCTGATCCTGC | 8213 |
| P (Mutant) | CTAGGCGTCAAGCTGGATGCGGGAGCATGATTGAGATGAGGG | 8195 |
| **Figure S1 C**<br>TRAP<br>TS primer | AATCCGTCGAGCAGAGGGT | 371 |
| ACX primer | GCGCGCGCTTACCCTTACCCTACCCTAC | 372 |
| **Figure S2 A, B**<br>SMART-RT<br>TSO | GCCTAATCTGGCAAGGACGTTTATCAACGCAAGATCACTGACTTrGrG | 7215 |
| TSO specific PCR-F | CATTGCAAGCAGTGGTATCAAC | 7216 |
| TSO specific PCR-R1 | CGTCGCCTTCAGACAGTAAG | 6941 |
|---------------------|----------------------|------|

**Figure S2 G, H**

**G/I-tailing RT-PCR**

| F0 | GCTTTGTCATCCTTCTCTGC | 7096 |
|----|----------------------|------|
| F1 | CACACGTGACCAAGCTTC   | 7253 |
| R  | CTTGGTAATACGACTACTATAGCGAGA | 7310 |

**Figure S11 A, B**

**Integration verification**

| PCR-F | CATACACATACGATTTAGTGACACTATAG | 7223 |
|-------|--------------------------------|------|
| PCR-R | GGGTGCTGGATGCTTTCTGTAATAC     | 6942 |