

Atomic structure of wheat ribosome reveals unique features of the plant ribosomes

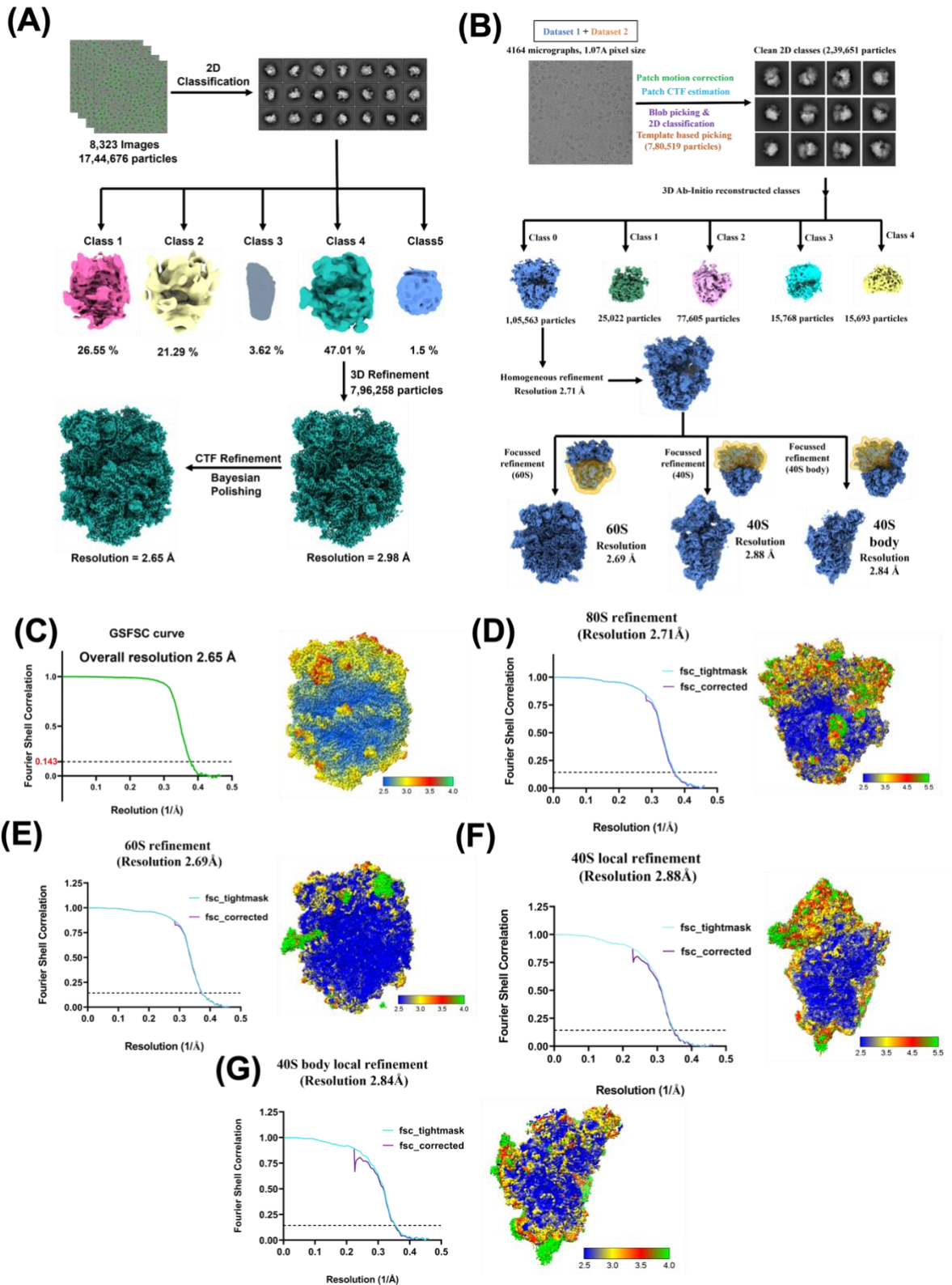
Rishi Kumar Mishra¹, Prafful Sharma¹, Faisal Tarique Khaja¹, Adwaith B. Uday¹, Tanweer Hussain^{1}*

¹Department of Developmental Biology and Genetics, Indian Institute of Science, Bengaluru, INDIA. PIN-560012

*Correspondence: hussain@iisc.ac.in

This file contains Supplementary Information: 8 Supp Figures and 8 Tables

Supp Fig 1



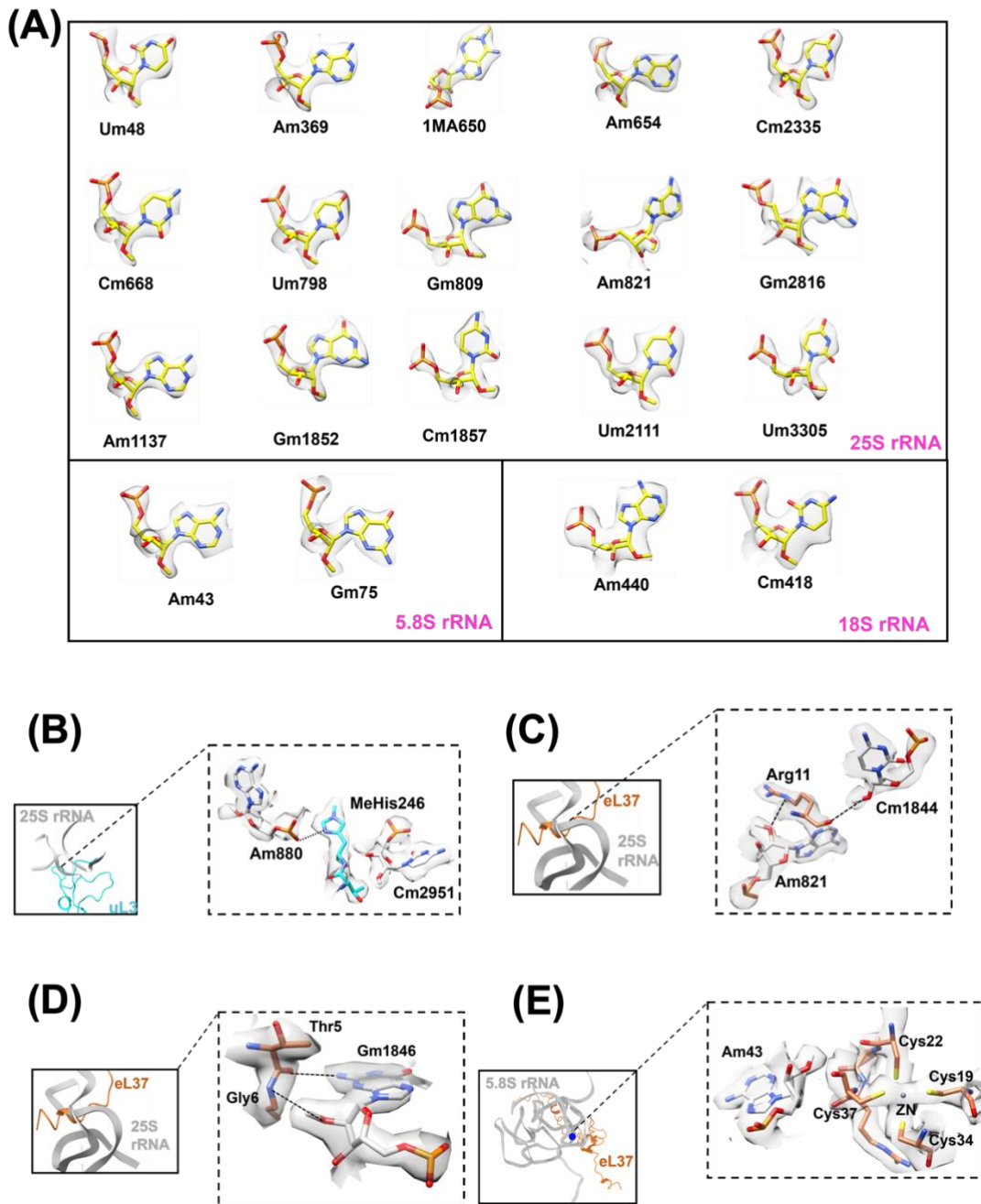
Supp Figure 1 Cryo-EM data processing for the wheat ribosome

(A) Data processing pipeline followed for the 60S dataset. The dataset of 8,323 movies was used and 17,44,676 particles were selected and subjected to 2D classification and reference-based 3D

classification leading to the removal of more than 50% of the junk particles. Class 4 obtained in 3D classification was used for 3D refinement resulting in a resolution of 2.98Å after postprocessing, followed by per-particle CTF refinement and Bayesian polishing. The polished particles were used for further 3D refinement leading to a global resolution of 2.65 Å after postprocess.

- (B) Data processing strategy followed for the 80S dataset: A total of 4,164 movies were collected and were subjected to patch motion correction and CTF estimation, and a total of 7,80,519 particles were picked using blob picking and subjected to multiple rounds of 2D classification to remove the junk particles; then an Ab-initio reconstruction into 5 classes was performed, and the class containing clean 80S was used for the final homogeneous refinement resulting into a resolution of 2.71Å followed by focused refinement on the 60S, 40S subunit and 40S body resulting into a resolution of 2.69Å, 2.88Å and 2.84Å respectively as depicted by the FSC curve and local resolution maps for Cryosparc data processing
- (C) Resolution estimation using Fourier Shell Correlation (FSC) plot for the autorefined map of the 60S dataset and Local resolution map for the 60S ribosome from wheat showing a range of resolution from 2.5Å in the core to the surface of the ribosome to 3.5Å in the peripheral region
- (D) Resolution estimation using FSC curve for the homogenous refined map of the 80S dataset and Local resolution map for the 80S ribosome from wheat showing a range of resolution from 2.5Å in the core to 5.5Å or more in the peripheral region
- (E) Resolution estimation using FSC curve for the local refined map of the 60S and Local resolution map for the 60S ribosome from wheat showing a range of resolution from 2.5Å in the core to 3.5Å or more in the peripheral region
- (F) Resolution estimation using FSC curve for the local refined map of the 40S and Local resolution map for the 40S ribosome from wheat showing a range of resolution from 2.5Å in the core to 3.5Å or more in the peripheral region
- (G) Resolution estimation using FSC curve for the local refined map of the 40S body and Local resolution map for the 40S body showing a range of resolution from 2.5Å in the core to 3.5Å or more in the peripheral region

Supp Fig 2

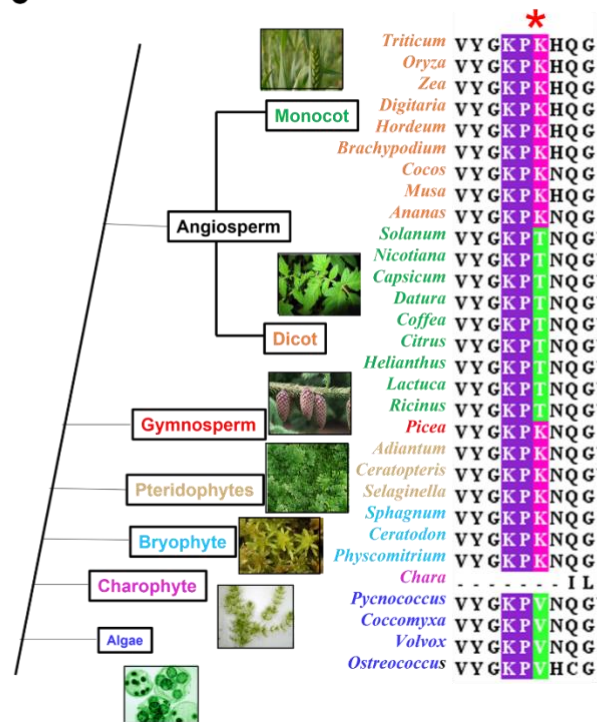


Supp Figures 2 The density for the methylation in the cryo-EM map of wheat ribosome and common features between wheat and tomato ribosome

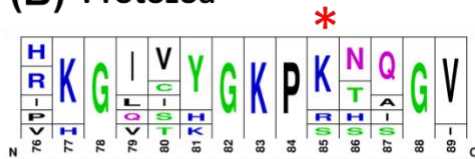
- (A) Density render diagrams for representative nucleotides that reflect the fit of chemical modification in the map
- (B) Interaction of universally conserved modified amino acid His246 with 25S rRNA bases harbouring plant specific 2'O-methylation (Am880 and Cm2951)
- (C) Plant specific Gm1846 forming interactions with N-terminal region of eL37
- (D) Arg11 of eL37 RP forms interactions with bases Cm1844 and Am821, which possess 2'O-methylation only in plants
- (E) Am43 in plant 5.8S rRNA directly interact with Zinc-finger motif of eL37

Supp Fig 3

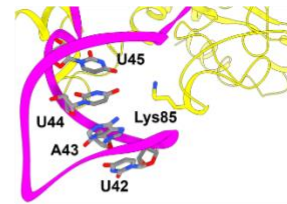
(A)



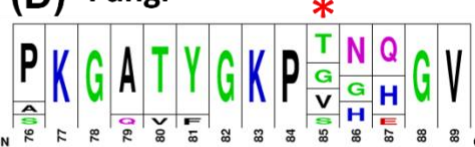
(B) Protozoa



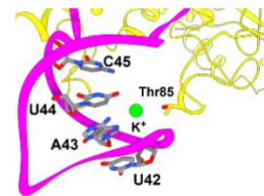
(C) *L. donovani*



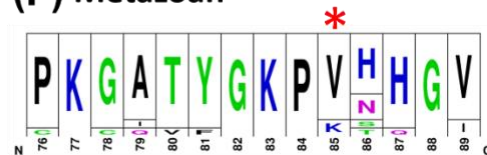
(D) Fungi



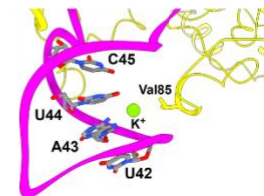
(E) *S. cerevisiae*



(F) Metazoan



(G) *H. sapiens*



Supp Figures 3 Sequence and structure analysis of eL15 from different species

- (A) The sequence analysis in plants shows the Lys85 (shown by an asterisk) to be highly conserved in Bryophyte, Pteridophyte, Gymnosperm and Angiosperm
- (B) & (C) Sequence analysis of eL15 and structural analysis of eL15-H11 interface from protozoa
- (D) & (E) Sequence analysis of eL15 and structural analysis of eL15-H11 interface from fungi
- (F) & (G) Sequence analysis of eL15 and structural analysis of eL15-H11 interface from metazoan

Supp Fig 4

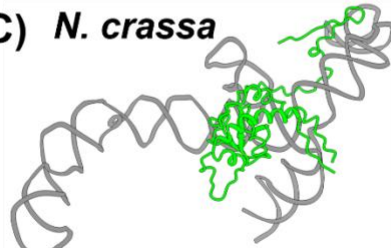
(A) *T. aestivum*



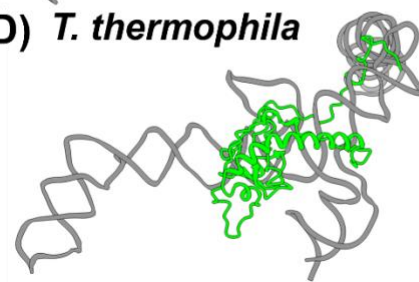
(B) *H. sapiens*



(C) *N. crassa*



(D) *T. thermophila*

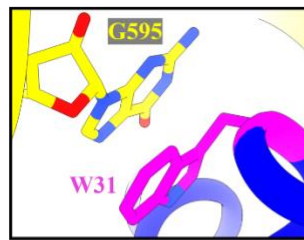


ES7 (Expansion segment)
eL6 (ribosomal protein)

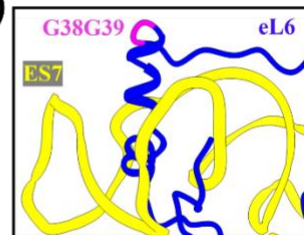
(E)

| | |
|---------------------------------|--------------------|
| <i>Triticum aestivum</i> | HRRGLVAIKAKHGC AFP |
| <i>Oryza sativa</i> | HRRGLVAIKAKHGC AFP |
| <i>Zea mays</i> | HRRGLVAIKAKNGC AFP |
| <i>Panicum virgatum</i> | HRRGLVAIKAKHGC AFP |
| <i>Setaria italica</i> | HRRGLVAIKAKHGC AFP |
| <i>Sorghum bicolor</i> | HRRGLVAIKAKHGC AFP |
| <i>Brassica napus</i> | HKRGLVAIKAKNCGVFP |
| <i>Raphanus sativus</i> | HKRGLVAIKAKNCGVFP |
| <i>Hibiscus syriacus</i> | HKRGLVAIKAKHGGFFP |
| <i>Ricinus communis</i> | HKRGLVAIKAKNCGVFP |
| <i>Gossypium gossypoides</i> | HKRGLVAIKAKHCGVFP |
| <i>Sinapis alba</i> | HKRGLVAIKAKNCGVFP |
| <i>Pisum sativum</i> | HKRGLVAIKAKHGC AFP |
| <i>Picea sitchensis</i> | HKRGLVAIKAKHGC SFP |
| <i>Adiantum capillus</i> | HKRGLVAIKAKHGC SFP |
| <i>Ceratopteris richardii</i> | HKRGLVAIKAKNCGTFF |
| <i>Sphagnum magellanicum</i> | HKRGLVAIKAKNCGTFF |
| <i>Ceratodon purpureus</i> | HKRGLVAIKAKNCGKFP |
| <i>Physcomitrium patens</i> | HKRGLVAIKAKHGC SFP |
| <i>Marchantia polymorpha</i> | KRRGLVAIKKKNCGKFP |
| <i>Coccomyxa subellipsoidea</i> | HRRGLVALKKKNCGKFP |
| <i>Chlorella ohadii</i> | KRRGLVAIKKKNCGKFP |
| <i>Volvox reticuliferus</i> | KRRGLVRLKDKKSPVVE |
| <i>Drosophila melanogaster</i> | ARKALVKKRFKAPTKI |
| <i>Xenopus laevis</i> | SRKAMKRRKYSAAKSR I |
| <i>Felis catus</i> | SRKAMKRRKYEAPETKI |
| <i>Mauremys mutica</i> | SRKAMKRRKYSAAKSR I |
| <i>Hyaena hyaena</i> | SRKAMKRRKYSAAKSR I |
| <i>Equus quagga</i> | SRKAMKRRKYSATKSRV |
| <i>Elephantulus edwardii</i> | SRKAMKRRKYSAAKSRV |
| <i>Homo sapiens</i> | SRKAMKRRKYSAAKSRV |
| <i>Saccharomyces cerevisiae</i> | ----- |
| <i>Candida albicans</i> | ----- |
| <i>Kluyveromyces lactis</i> | ----- |
| <i>Neurospora crassa</i> | ---MS SAAAPQTKTFG |
| <i>Plasmodium falciparum</i> | -----MTNTSNEL |
| <i>Leishmania donovani</i> | ----- |
| <i>Euglena gracilis</i> | ----- |
| <i>Toxoplasma gondii</i> | ----- |
| <i>Tetrahymena thermophila</i> | ----- |

(F)



(G)



| | |
|--|---|
| ■ Plantae (Monocot) | ■ Plantae (Algae) |
| ■ Plantae (Dicot) | ■ Protozoa (Sporozoa) |
| ■ Plantae (Gymnosperm) | ■ Fungi |
| ■ Plantae (Pteridophytes) | ■ Metazoan |
| ■ Plantae (Bryophytes) | |

Supp Figure 4 Interaction between eL6 and Expansion segment 7^c (ES7^c)

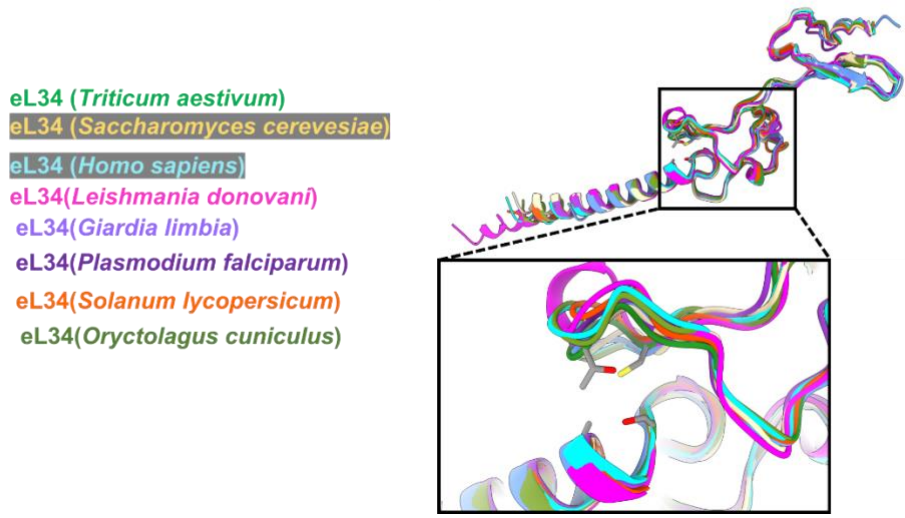
(A) & (B) Intercalation of NTT of eL6 through ES7^c in (A) plant (wheat) and (B) metazoan (human), respectively, as depicted by an asterisk to represent the position of intercalation
 (C) & (D) No intercalation through ES7^c by eL6 observed in (C) fungi and (D) protozoan, respectively
 (E) Sequence comparison of NTT of plant eL6 with protozoa, yeast, and metazoan
 (F) Stacking interaction between a plant-specific W31 in eL6 with G595 of 25S rRNA
 (G) Plant-specific Glycine dipeptide G38G39 depicted in magenta, facilitates the bending of the eL6-NTT towards ES7^c

Supp Fig 5

(A)

| | | 80 | 90 | 100 | 110 | 120 |
|------------------------------|-------------------------------------|----------|-------|---|---------------------|--------|
| Mammals | <i>Elephantulus edwardii</i> | VGKAPKSA | GM | P GRLRGVRAVRPKVL | MRLSKTKKHVSRAYGGSM | AK VR |
| | <i>Macaca fascicularis</i> | VGKAPKSA | GM | P GRLRGVRAVRPKVL | MRLSKTKKHVSRAYGGSM | AK VR |
| | <i>Oryctolagus cuniculus</i> | VGKAPKSA | GM | P GRLRGVRAVRPKVL | MRLSKTKKHVSRAYGGSM | AK VR |
| Fungi | <i>Rattus rattus</i> | VGKAPKSA | GM | P GRLRGVRAVRPKVL | MRLSKTKKHVSRAYGGSM | AK VR |
| | <i>Homo sapiens</i> | VGKAPKSA | GM | P GRLRGVRAVRPKVL | MRLSKTKKHVSRAYGGSM | AK VR |
| | <i>Saccharomyces cerevesiae</i> | LATRPK | GD | GSALQGISLTPRQY | ATVSKTHKTVSRAYGGSRA | AN VK |
| | <i>Podila horticola</i> | PGTAPK | GD | GVALP GVPALRPTEY | ARISRRQKSVSRAYGGSRA | AN VK |
| | <i>Rhizopus delemar</i> | PVKAPR | GD | GEALAGIKALRPREF | ATVSKTKKTVSRAYGGSRA | AH VR |
| Plantae (Monocot) | <i>Kluyveromyces lactis</i> | LATRPK | GD | GSALQGISLTPRQY | ATVSKTHKTVSRAYGGSRA | AN VK |
| | <i>Candida glabrata</i> | LATRPK | GD | GIALP GIATLRPRQY | ASISKTHKTVSRVYGGSR | AN VK |
| | <i>Triticum aestivum</i> | RASGPK | PVT | GKKIQGIPHLRPT EYKRPRLSRNRRTVNRPTGGVL | GGP VR | |
| | <i>Zostera marina</i> | RASGPK | PVT | GKRIQGIPHLRPT EYKRSRLSRNRRTVNRAYGGVL | GGC VR | |
| | <i>Cocos nucifera</i> | RASGPK | PVT | GKRIQGIPHLRPAEYKRSRLSRNRRTVNRAYGGVL | GGV VR | |
| Plantae (Dicot) | <i>Zingiber officinale</i> | RASGPK | PVT | GKRIQGIPHLRPTQYKRSRLSRNRRTVNRAYGGVL | GGG VR | |
| | <i>Zea mays</i> | RASGPK | PVT | GKKIQGIPHLRPAEYKRSRLSRNRRTVNRPTGGVL | SGT VR | |
| | <i>Coffea arabica</i> | RASGPK | PVT | GKRIQGIPHLRPAEYKRSRLSRNRRTVNRAYGGVL | GGV VR | |
| Plantae (Gymnosperm) | <i>Helianthus annuus</i> | RASGPK | PVT | GKRIQGIPHLRPAEYKRSRLSRNRRTVNRAYGGAL | SAG VR | |
| | <i>Pisum sativum</i> | RASGPK | PVT | GKRIQGIPHLRPT EYKRSRLSRNRRTVNRAYGGVL | GGG VR | |
| Plantae (Pteridophytes) | <i>Solanum lycopersicum</i> | RANGPK | PVT | GKRIQGIPHLRPAEYKRSRLSRNRRTVNRAYGGVL | SGS VR | |
| | <i>Picea sitchensis</i> | RASGPK | PVT | GKRIHGIPHLRPAEYKRSRLSRNRRTVNRAYGGVL | SGS VR | |
| Plantae (Bryophytes) | <i>Adiantum capillare</i> | RAQGPK | CAIT | GKRIQGIPHLRPAEYKRSRLSRNRRTVNRAYGGAL | SGS VR | |
| | <i>Ceratopteris richardii</i> | KANGPK | CALT | GKRIQGIPHLRPAEYKRSRLSKRQKT VNRAYGGTL | SGE VR | |
| | <i>Sphagnum magellanicum</i> | RAKGPK | SIT | GKRIAGIPHLRPT EYKTSRLSRNRRTVNRAYGGNL | SGS VR | |
| Plantae (Algae) | <i>Ceratodon purpureus</i> | RANGPK | PVT | GKRIAGIPHLRPT EYKTSRLSRNRRTVNRAYGGNL | AGS VR | |
| | <i>Physcomitrium patens</i> | RARGPK | PVT | GKRIAGIPHLRPT EYKTSRLSRNRRTVNRAYGGNL | AGS VR | |
| | <i>Pycnococcus provasolii</i> | RVSGPK | CAKS | GQVIHGKVVHVPWEMS KNRNMKKEKT VHRAYGGCL | SHG VR | |
| | <i>Coccomyxa sp</i> | RPSPRI | GAT | GVKLHGIPSLRPKEMS NRRLSRP SKTVHRIYGGHL | SHA VR | |
| Protozoa (Sporozoa) | <i>Volvox africanus</i> | QESHPK | AVS | GARLHGFAAVPHTQL | HTLSKRAKKVNR IYGGHL | SHK VR |
| | <i>Ostreococcus tauri</i> | TTKGAQ | TPS | DNGRIHGVP RVATQKYSRKHM SKNKS VSRAYGGVL | GGG VR | |
| | <i>Micromonas commode</i> | KTKGPQ | TPS | DHGRIHGVP RVAQVQKYS SKHMAKNKS VTRAYGGAL | SAG VR | |
| Protozoa (Cnidospora) | <i>Emeria necatrix</i> | QASRQK | GGC | GRLLP GIPARRPPQF | RLKKRERTVNRAYGGTR | HS VR |
| | <i>Plasmodium malariae</i> | KAGKPK | ADC | KTAIQGVKALRPADN | RRARKNRTVSRAYGGTS | AK VR |
| | <i>Toxoplasma gondii</i> | QPSRPK | GNC | HRALP GIPAVAPHRL | RLKKRERTVHRAYGGSRA | HA VR |
| | <i>Babesia ovata</i> | VAQGPK | GDC | KRRLAGIAALRP HLY | RNLKRRRTVSRAYGGVR | HG VR |
| Protozoa (Sarcomastigophora) | <i>Cryptosporidium ubiquitum</i> | VYSRPK | GDC | KKPLAGIPACAPYEM | KHLKKRERTVARAYGGTK | ST VR |
| | <i>Mycobolus squamulis</i> | KGTQPK | GDC | KKTLAGISAVRCKL | RGMSKRQKT VNRSYGGSL | GK VR |
| | <i>Nosema bombycis</i> | HSKVRR | HEC | NAKLLS IARMRPAEL | SQKVS SKRVCRPYGGKF | GN VR |
| Protozoa (Ciliophora) | <i>Encephalitozoon cuniculi</i> | HSKKHR | HEC | NAILGS IARMRPAEF | SQKVS ARR VNRPTGATT | GR VR |
| | <i>Paranosema locustae</i> | PGRVPK | CVK | RSKLRGIDI CRPAAF | ARLRSKQRTVARTYGGNL | GS VR |
| | <i>Englena gracilis</i> | LPKGPHTP | VSLG | HKP I P GVKRLRS IQR | KSA PKRHLTVSRAYGGCL | HD VR |
| Archaea | <i>Leishmania</i> | RSQGIHTP | VWLG | HKRLGGTKALRHIDA | RLASRHEKSVSRAYGGVL | SHD VR |
| | <i>Trichomonas vaginalis</i> | RQNGPH | CAET | GKRIINGIKCVKTCEL | RRMKNRQTVSRPTGGVY | GS VR |
| | <i>Trypanosoma cruzi</i> | RSQGPHTP | VWLG | HKRLAGTKALRHTEA | RLASRHEKSTSRPTGGVL | SHD VR |
| | <i>Blepharisma stoltei</i> | IAKGPH | CKET | GERLAGIPALRPKEY | SRI NKKDRTVSRAYGGVL | SHK VR |
| Archaea | <i>Paramecium sonneborni</i> | KTSAST | ADSNL | SVVLNGLKRI RPTKL | KQLSRQKTVSRPTGGVL | SAS VR |
| | <i>Stylonychia lemnae</i> | SA | --- | GKT VAGIPKLRSPAL | SRLTVTKRTVSRAYGGKL | HA VR |
| | <i>Tetrahymena thermophila</i> | VVNYTK | SEAGC | NVALNGIAQVRPAEY | ATIARS AKTVSRVYGGEL | HT VR |
| | <i>Ichthyophthirius multifiliis</i> | VVNQK | AEPGC | GALLNGIANVRASAL | STMSRRQKTVSRPTGGH | HT VR |
| | <i>Methanothermobacter sp.</i> | VVNYTK | SEAGC | NVALNGIAQVRPAEY | ATIARS AKTVSRVYGGEL | HT VR |
| | <i>Nanoarchaeota archaeon</i> | QPSKAK | GGC | GKVLGAVARARPHKM | RKMAKTKRPTRPYGGNL | SP VR |
| | <i>Methanomicrobia archaeon</i> | KHSKPR | CAEC | GAELHGVP RGSPT EI | KKLSKSKKTPTRPYGGYL | SK VR |
| | <i>Thermococci archaeon</i> | KVDWAK | CANC | GSILNGVPRLRPSEM | RKLSKSEERRPNRPYGGYL | PR VR |
| | <i>Euryarchaeota archaeon</i> | KPSKHV | CVHC | RKPLHAVARGRPYQI | KKLSKSKRPNRPYGGYL | PE VR |

(B)



Supp Figure 5 Sequence alignment of eL34 from various eukaryotes

(A) Sequence comparison shows an absence of the otherwise conserved cysteine in the case of plants and protozoa

(B) Structure comparison of eL34 from species with and without Zinc-finger motif showing no overall difference in the conformation

Supp Fig 6

(A)

```

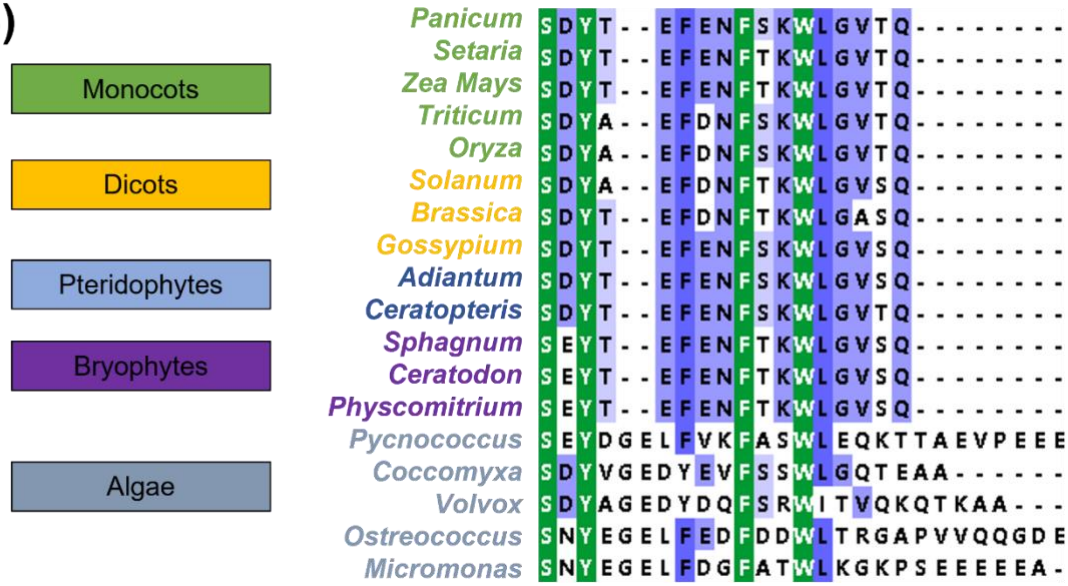
Saccharomyces_cerevisiae1 1...MSRPOVTVHSLTGEAT...ANALPLPAVFSAPTRPDI VHTVFTSVNKRKRDQAVS...EKAGHOTSAESWGTGRAVARI PRVGGGTHSGDGFAGMCRGGRMFAPTRITWRWMI 110
Chaetomium_thermophilum1 1...MASRPTVTYTGADIKPT...GATEVLPWFESAPTRPDI VHNHTTGMARNRKGFYAVS...EKAGHOTSAESWGTGRAVARI PRVSGGTHRA DGFAGMCRS GRMFAPTRITWRWMI 111
Drosophila_melanogaster1- 1MSLGNARRLVSYTERNEPKAD...KMICLPAFKAPTRPDI VYVMEVOLLRRNRKQAVS...ELAGHOTSAESWGTGRAVARI PRVGGGTHSGDGFAGMCRGGRMFAPTRITWRWMI 115
Triticum_aestivum1-404 1...MATTARPLVSKALGDMPDAD...AGVPMHVMKAPTRPDI VTFHRLVSCSRDPYAVS...RKAGHOTSAESWGTGRAVARI PRVGGGTHSGDGFAGMCRGGRMFAPTRITWRWMI 115
solanium_lycopersicon1-40k 1MATAAARITITIMOSLENDMATDS...AAVPLPAWKKAPTRPDI VYVYVHSHISRAHRDPYAVS...RKAGHOTSAESWGTGRAVARI PRVGGGTHRA DGFAGMCRGGRMFAPTRITWRWMI 116
Arabidopsis_thaliana1-405 1MAAAARPLVFTIOTLIDIMSTDOSSVVEVRODMTAPVVRPDI VNFVHQAQISNRSRPFYAVS...RKAGHOTSAESWGTGRAVARI PRVGGGTHRA DGFAGMCRGGRMFAPTRITWRWMI 117
Tetrahymena_thermophila1- 1...MTRDQIHNDAKCAN...KOTATKLLPAVETAPTRIDIVHKYFTDLNRKRKASGKKI STRGTADMGHSASWGTGRAVARI PRVGGGTHSGDGFAGMCRGGRMFAPTRITWRWMI 118
Plasmodium_falciparum1-41 1...MATIRPVANVYSTNGKNV...VGEVEI PVVFTPIRNDLIDQSYVTNMSKRHRHPYAVK...LGAQYVESWGTGRAVARI PRVGGGTHRA DGFAGMCRGGRMFAPTRITWRWMI 112
Mus_musculus1-419 1...MACARPLISYSEKDES...SG...KNVILPAFKAPTRPDI VMTHTLRRNRKDPYAVS...ELAGHOTSAESWGTGRAVARI PRVGGGTHSGDGFAGMCRGGRMFAPTRITWRWMI 112
homo_sapiens1-427 1...MACARPLISYSEKDES...SG...KNVILPAFKAPTRPDI VMTHTLRRNRKDPYAVS...ELAGHOTSAESWGTGRAVARI PRVGGGTHSGDGFAGMCRGGRMFAPTRITWRWMI 112

Saccharomyces_cerevisiae1 111VRVHNEKRYTASAIKRTAVASLVARGRHVEKPIPIPLVSTDL...ESIQTKERVAALHVGASHLLKVLKSKLBAQDKYRNRWITRGRPLVYAE...NGIVKALRHVPGV 225
Chaetomium_thermophilum1 112VKIINGORRITASALAKSAPAPLMDKRDVSTVPLVLDVDSAAVAGAVATAARYKLRAGQPVKRVKSKLBAQDKYRNRWITRGRPLVYAE...NGIVKALRHVPGV 232
Drosophila_melanogaster1- 116RRVNVNQRVYLVKAIKSSCPALVQSHVVEGGSTFPLVSDV...QKQVYKQKQVIFRRLKVMHIOKQYKQRTSAGSETMQRRIARIGPLVYDQ...EGLRKRTRIPG 230
Triticum_aestivum1-404 116RRVNI LRHRVAVASALAKTAVPALVTRAGHRIEVSPEFPLVSDSA...EGIEKTSQAKVVKOLGAYABADKAKDSVIRPDKGMMNRRIYINIKKPLIVYATE...GKLVKAFRNI PGV 231
solanium_lycopersicon1-40k 117RRIPVNOHRYVASAIKSSVPSLVLARGRHIEVSPELPLVSDSI...EGIEKTSVLIKALKOI GAYPAEKADSHAI RPKDKGMMNRRIYISKIPLVYATE...GAKLVKAFRNI PGV 232
Arabidopsis_thaliana1-405 118RRVNVNQRVYLVKAIKSSCPALVQSHVVEGGSTFPLVSDSA...EAVEKTSALIKVROI GAYDPAEKAKNIGI RPKDKGMMNRRIYISKIPLVYATE...GKLVKAFRNI PGV 233
Tetrahymena_thermophila1- 118RRVNVNQRHVAASALAKSALVPLVFRGRHISNVOLLVYVSDSV...EYERKQVAFNRVGYDVLVRAETKALBAQDKLRNRKLRIGPLVYVGN...KSTLLRALRNI PGV 233
Plasmodium_falciparum1-41 113RRVNLKERRYVCSGISASGVTLSLARGRHISHLKVLPLVSDSI...EELSKTKERVNFVSLGLKDEVNRVLRKSKIRAGKQMMNRRIYIKIPILVYEND...NGVKAFRNI PGV 227
Mus_musculus1-419 113RRVNTIQKRYICSAALAKSALPALVMSKDRHIEEVEPLPLVSDV...EYKTKERVQLKRLKAWNIRKQVYASQRMBAQDKGMMNRRIYIRDFICLVYEND...NGIIRAFRNI PGV 227
homo_sapiens1-427 113RRVNTIQKRYICSAALAKSALPALVMSKDRHIEEVEPLPLVSDV...EYKTKERVQLKRLKAWNIRKQVYASQRMBAQDKGMMNRRIYIRDFICLVYEND...NGIIRAFRNI PGV 227

Saccharomyces_cerevisiae1 226ETANVASLLEOLAPGHLGRVIVWESAFKIKLQDQWSETVA...SSKQVITSHIISTSDVTRINSSEI DSAIR...PAGGATGRHVLKRPKPKKQVLRNLNRYAKVFAEKLGSK...342
Chaetomium_thermophilum1 233ETCPDVALNLLDLPAGGHLGRVIVWESAFIKQDAVYE...SKKQVITPANI VSOADLSLINSSTI BSVLRAPGERTKACVOKKIPMLRNGIMLRYNPYASTFAEKLGVE...344
Drosophila_melanogaster1- 231ETIINDNLLKLPAGGGRVGRVWESAFKIKLQDQWSETVA...SSKQVITSHIISTSDVTRINSSEI DSAIR...PAGGATGRHVLKRPKPKKQVLRNLNRYAKVFAEKLGSK...342
Triticum_aestivum1-404 232DVANVERLLDLPAGGHLGRVIVWESAFKIKLQDQWSETVA...SSKQVITSHIISTSDVTRINSSEI DSAIR...PAGGATGRHVLKRPKPKKQVLRNLNRYAKVFAEKLGSK...342
solanium_lycopersicon1-40k 233EICHVDRI NLLKLPAGGHLGRVIVWESAFKIKLQDQWSETVA...SSKQVITSHIISTSDVTRINSSEI DSAIR...PAGGATGRHVLKRPKPKKQVLRNLNRYAKVFAEKLGSK...342
Arabidopsis_thaliana1-405 234ELCHVERL NLLKLPAGGHLGRVIVWESAFKIKLQDQWSETVA...SSKQVITSHIISTSDVTRINSSEI DSAIR...PAGGATGRHVLKRPKPKKQVLRNLNRYAKVFAEKLGSK...342
Tetrahymena_thermophila1- 234DYCNKRL NLLKLPAGGGRVGRVWESAFKIKLQDQWSETVA...SSKQVITSHIISTSDVTRINSSEI DSAIR...PAGGATGRHVLKRPKPKKQVLRNLNRYAKVFAEKLGSK...342
Plasmodium_falciparum1-41 228DLCKVTK NLLKLPAGGSI GRLCWSESAFKKLDVYKGIHEKVTKNKIYIPKSI VHPDIYRIINSOKVASELAKKRC...KRLKLNKRSITFAVRCVLPAYKLRSLAVRRMR 345
Mus_musculus1-419 228TLLNRSK NLLKLPAGGGRVGRVWESAFKIKLQDQWSETVA...SSKQVITSHIISTSDVTRINSSEI DSAIR...PAGGATGRHVLKRPKPKKQVLRNLNRYAKVFAEKLGSK...342
homo_sapiens1-427 228TLLNRSK NLLKLPAGGGRVGRVWESAFKIKLQDQWSETVA...SSKQVITSHIISTSDVTRINSSEI DSAIR...PAGGATGRHVLKRPKPKKQVLRNLNRYAKVFAEKLGSK...342

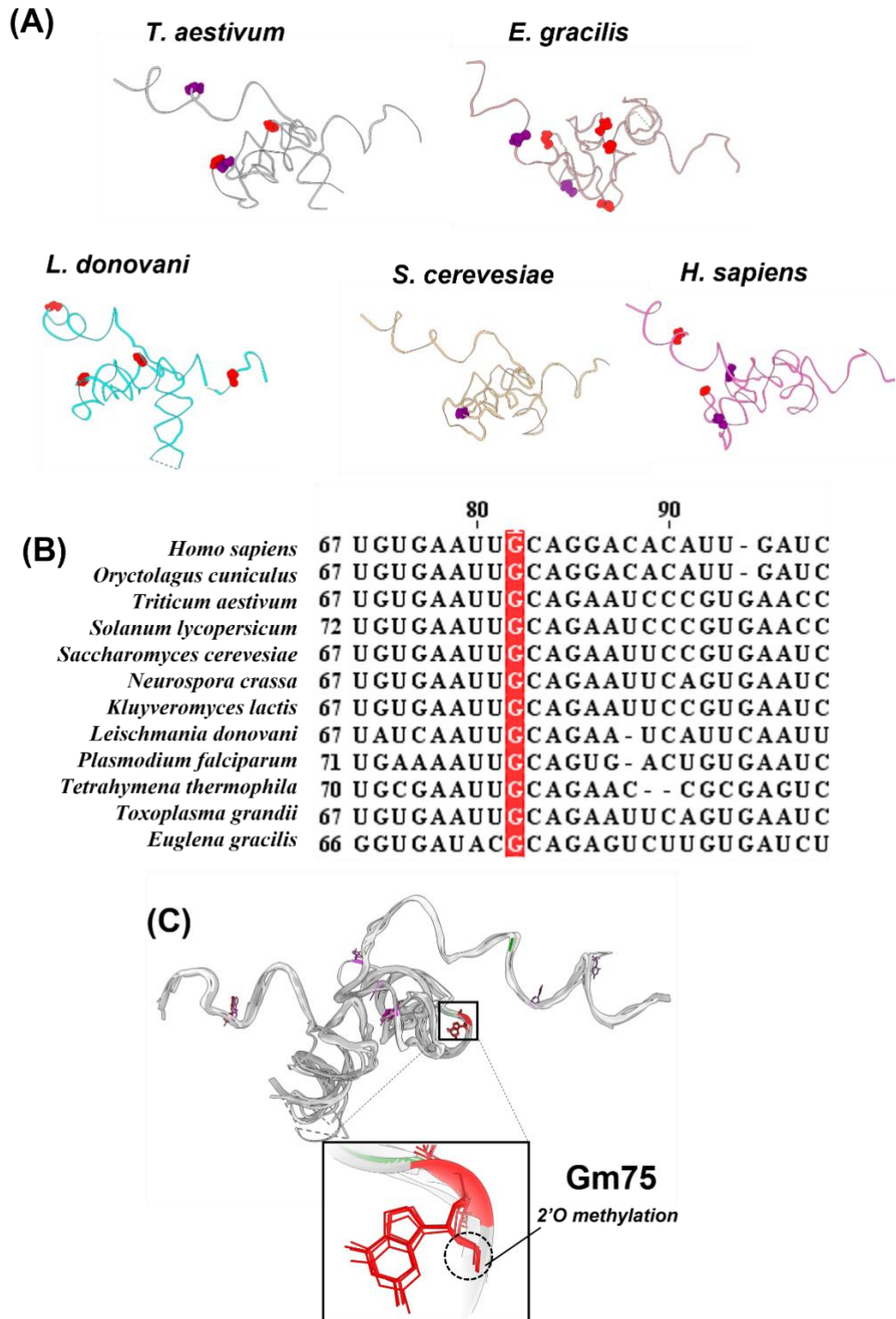
Saccharomyces_cerevisiae1 343...KAEKTGTPK...AVFTTELLKHD... 362
Chaetomium_thermophilum1 345...KAEKGPPK...VPASFKELLHA... 365
Drosophila_melanogaster1- 349RTVNRVLAARAKKQVLA...KSHFANVA...TRKA...ANRALKLA...ARKKKVAARPKPAK... 401
Triticum_aestivum1-404 350RVKARKDKINSKRTKLSVEEASKIKAAGKAWYQTIMSDSYMEF...DVSFQWLG...VSO... 404
solanium_lycopersicon1-40k 352RVKARKEKLDKRRHQITKEEASAIKGAASHWYKTIMSDSYAEF...DNFTIWLG...VSO... 406
Arabidopsis_thaliana1-405 352RVKARKEKLDKRRKIVTRKEALAIKAAGKAWYKTIMSDSYIEF...DNFTIWLG...ASO... 406
Tetrahymena_thermophila1- 351SKRRQRTLKAAR...KARTIKKGGQAWIATFNKNECAIKARQEDADFIA...OGQEI...EGDE... 410
Plasmodium_falciparum1-41 346SILEKSNKKEKRVQKIQKLEKQI NHDYVYKGAQVAKVKKKK...REEKASKRTANQAVINVA...AE... 411
Mus_musculus1-419 346NHILRVKLE-AAATALA...TSEKVVV...EGOT...ADKKPVAGKGGKVDVAKK...KPGKGVAKKPEKPKPTTEEKPPAA... 419
homo_sapiens1-427 346NHILRVKLE-AAATALQ...AKSDERA...AV...AGKPVVGRGKGAQVGVKQRKPLVGGKAAATKPKPEKPKPKPTTEEKPPAA... 427
    
```

(B)



Supp Figure 6 Sequence analysis of uL4 from plants and other species
 (A) Multiple sequence alignment of uL4 from different eukaryotic organisms showing varying lengths of uL4 CTT
 (B) Multiple sequence alignments of uL4 CTT from higher plants to lower algae showing conservation of residues (highlighted in green) involved in the interaction with neighbouring eL20 and eL21

Supp Fig 7



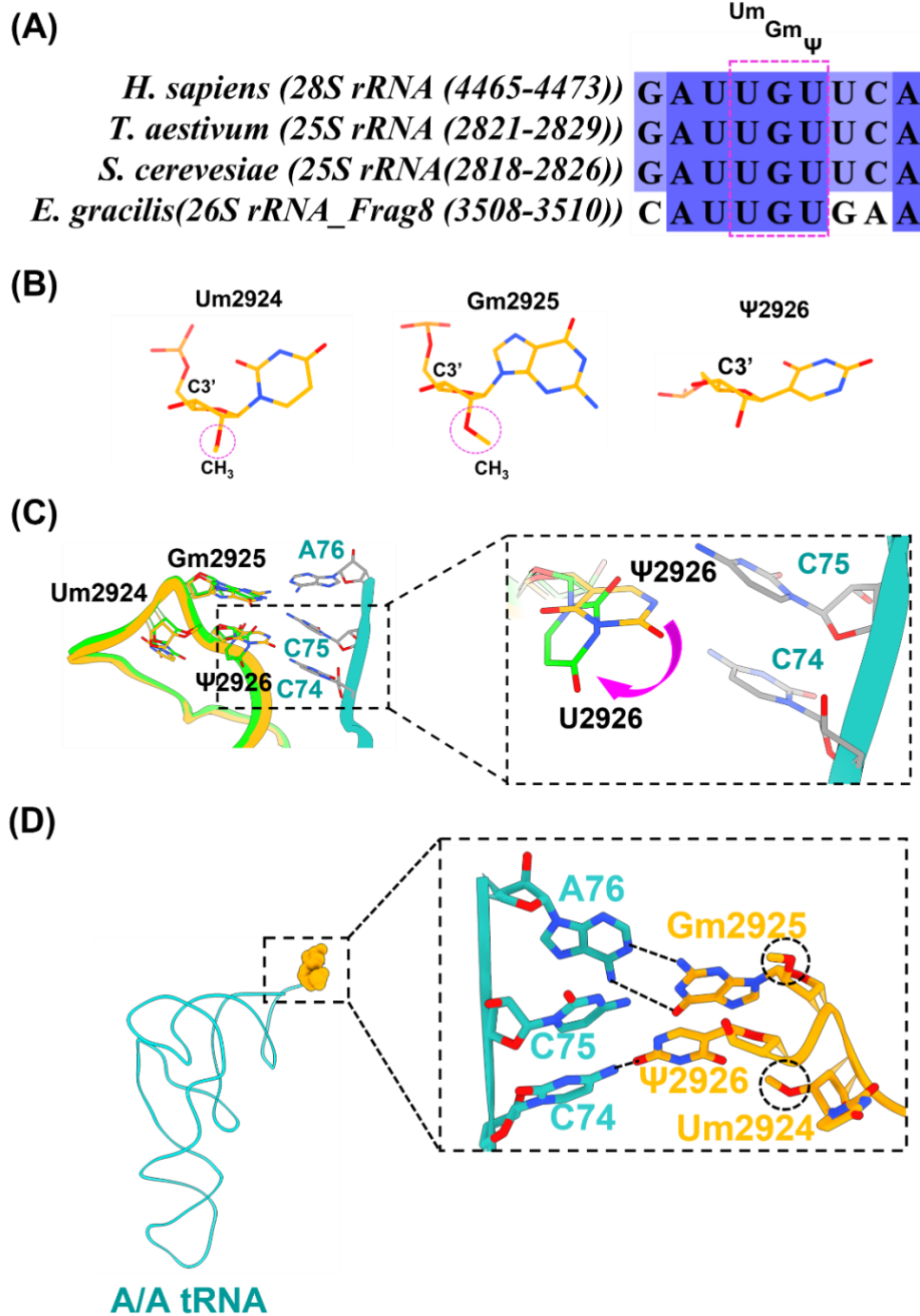
Supp Figure 7 Potential role of conserved Gm75 in translation

(A) Chemical modifications on 5.8S rRNA of different species is represented where the position as well as number of modification varies across eukaryotes

(B) Sequence comparison of 5.8S rRNA from different species shows high conservation of Guanine at this position throughout evolution

(C) Superposition of the 5.8S rRNA shows the conservation of modification on G75 in different species (yeast ribosome shows no modification (PDB-ID: 4V88))

Supp Fig 8



Supp Figure 8 Conservation of modified triplet and potential of 2' O ribose methylation to induce C3' endo conformation

(A) Sequence comparison shows the conservation of modified nucleotides UmGmΨ triplet (Um2924Gm2925Ψ2926) through eukaryotic evolution

(B) A C3' endo conformation is induced in Um2924Gm2925Ψ2926 of H92 by the presence of methylation at 2' O ribose of the UmGmΨ triplet

(C) Comparison of H92 between wheat ribosome structure (orange) and pseudouridine-less ribosome structure (green) from yeast (PDB ID 7MPI) showing the altered conformation of U2926

(D) Planar arrangement of the modified residues Um2924Gm2925Ψ2926 involved in interaction with CCA at the acceptor arm of tRNA

Table 1: Data Collection and refinement statistics

| Data collection and processing | 60S subunit | 40S subunit |
|--|---------------------|---------------------|
| Microscope | Titan Krios | Titan Krios |
| Voltage | 300 kV | 300 kV |
| Magnification | 75,000x | 75,000x |
| Detector | Falcon III | Falcon III |
| Sampling (Å/pixel) | 1.07 | 1.07 |
| Total electron dose [$e^-/\text{Å}^2$] | 44.60 | 44.60 |
| Symmetry imposed | C1 | C1 |
| Average resolution (Å) | 2.7 | 2.9 |
| Atomic Model composition | | |
| Chains RNA/Protein | 3/44 | 1/31 |
| Non-hydrogen atoms | 116711 | 69046 |
| Amino Acids | 6191 | 4453 |
| Nucleotides | 3134 | 1571 |
| Number of ligand atoms | 317 | 87 |
| Zn ²⁺ /Mg ²⁺ /K ⁺ | 3/242/72 | 1/70/16 |
| Refinement Statistics | | |
| Model to map CC (CC _{mask} /CC _{box} /CC _{peaks} /CC _{volume}) | 0.82/0.73/0.72/0.82 | 0.77/0.67/0.76/0.72 |
| Resolution [Å] by model-to-map FSC, threshold 0.50 (masked/unmasked) | 2.6/2.5 | 3.3/3.4 |
| Bond length rmsd | 0.006 | 0.003 |
| Bond angle rmsd | 0.696 | 0.636 |
| Validation | | |
| Clash score | 6.57 | 12.43 |
| Rotamer outliers [%] | 3.32 | 0.47 |
| Ramachandran plot [%] (Favored/allowed/disallowed) | 95.76/4.24/0 | 93.56/6.35/0.09 |
| CaBLAM outliers [%] | 2.32 | 3.50 |
| MolProbity Score | 2.32 | 2.03 |
| Accession ID | | |
| EMDB ID | | |
| PDB ID | | |

Table 2: Accession IDs for eL15 ribosomal protein

| Organism | Accession ID |
|------------------------------------|---------------------|
| <i>Triticum aestivum</i> | XP_044323743.1 |
| <i>Oryza brachyantha</i> | XP_040380504.1 |
| <i>Zea mays</i> | NP_001150386.2 |
| <i>Solanum lycopersicum</i> | XP_004241917.1 |
| <i>Nicotiana tomentosiformis</i> | XP_009601449.1 |
| <i>Capsicum annuum</i> | XP_016582172.1 |
| <i>Picea sitchensis</i> | ABK22257.1 |
| <i>Adiantum capillus</i> | KAI5058505.1 |
| <i>Ceratopteris richardii</i> | KAH7288671.1 |
| <i>Selaginella moellendorffii</i> | XP_002961462.1 |
| <i>Sphagnum magellanicum</i> | KAH9540499.1 |
| <i>Ceratodon purpureus</i> | KAG0554928.1 |
| <i>Physcomitrium patens</i> | XP_024357573.1 |
| <i>Coccomyxa subellipsoidea</i> | XP_005646118.1 |
| <i>Volvox africanus</i> | GIL45611.1 |
| <i>Cyanidioschyzon merolae</i> | XP_005534852.1 |
| <i>Plasmodium vivax</i> | XP_001615040.1 |
| <i>Leishmania donovani</i> | XP_003863082.1 |
| <i>Euglena gracilis</i> | QLA09621.1 |
| <i>Giardia intestinalis</i> | XP_001708727.1 |
| <i>Paramecium octaurelia</i> | CAD8175268.1 |
| <i>Toxoplasma gondii</i> | XP_002366469.1 |
| <i>Stentor coeruleus</i> | OMJ69253.1 |
| <i>Saccharomyces cerevisiae</i> | AJS80338.1 |
| <i>Kluyveromyces lactis</i> | QEU62382.1 |
| <i>Candida viswanathii</i> | RCK59573.1 |
| <i>Rhizopus arrhizus</i> | KAG0748686.1 |
| <i>Podila verticillata</i> | KAF9367241.1 |
| <i>Rhizoctonia solani</i> | CAE6457639.1 |
| <i>Cladochytrium replicatum</i> | KAI8800011.1 |
| <i>Lobosporangium transversale</i> | XP_021882443.1 |
| <i>Homo sapiens</i> | NP_001240308.1 |
| <i>Oryctolagus cuniculus</i> | XP_002716262.1 |
| <i>Bos taurus</i> | DAA16199.1 |
| <i>Xenopus laevis</i> | XP_018124768.1 |
| <i>Chelydra serpentina</i> | KAG6940002.1 |
| <i>Mauremys mutica</i> | XP_044863280.1 |
| <i>Gallus gallus</i> | NP_001292094.1 |
| <i>Anas platyrhynchos</i> | XP_005029295.3 |
| <i>Hydra vulgaris</i> | P61368.2 |
| <i>Periplaneta americana</i> | KAJ4450415.1 |
| <i>Caenorhabditis elegans</i> | NP_499964.1 |

Table 3: Accession IDs for eL6 ribosomal protein sequences

| Organism | Accession ID |
|---------------------------------|---------------------|
| <i>Triticum aestivum</i> | XP_044403216.1 |
| <i>Oryza sativa</i> | XP_015627578.1 |
| <i>Zea mays</i> | NP_001130513.1 |
| <i>Panicum virgatum</i> | XP_039847586.1 |
| <i>Setaria italica</i> | XP_004975938.1 |
| <i>Sorghum bicolor</i> | XP_002448023.1 |
| <i>Brassica napus</i> | XP_013664611.2 |
| <i>Raphanus sativus</i> | XP_018433501.1 |
| <i>Hibiscus syriacus</i> | XP_039007682.1 |
| <i>Ricinus communis</i> | XP_002513192.1 |
| <i>Gossypium gossypioides</i> | MBA0750743.1 |
| <i>Sinapis alba</i> | KAF8114319.1 |
| <i>Pisum sativum</i> | KAI5405181.1 |
| <i>Picea sitchensis</i> | ABK21588.1 |
| <i>Adiantum capillus</i> | KAI5075325.1 |
| <i>Ceratopteris richardii</i> | KAH7291547.1 |
| <i>Sphagnum magellanicum</i> | KAH9535645.1 |
| <i>Ceratodon purpureus</i> | KAG0570454.1 |
| <i>Physcomitrium patens</i> | XP_024369198.1 |
| <i>Marchantia polymorpha</i> | OAE31965.1 |
| <i>Coccomyxa subellipsoidea</i> | XP_005647799.1 |
| <i>Chlorella ohadii</i> | KAI7837103.1 |
| <i>Volvox reticuliferus</i> | GIL85813.1 |
| <i>Felis catus</i> | XP_006938548.1 |
| <i>Hyaena hyaena</i> | XP_039102100.1 |
| <i>Equus quagga</i> | XP_046497158.1 |
| <i>Elephantulus edwardii</i> | XP_006901041.1 |
| <i>Homo sapiens</i> | 6QZP (PDB) |
| <i>Mauremys mutica</i> | XP_044845807.1 |
| <i>Xenopus laevis</i> | XP_018119120.1 |
| <i>Drosophila melanogaster</i> | NP_651876.1 |
| <i>Candida albicans</i> | AJV67506.1 |
| <i>Kluyveromyces lactis</i> | XP_451742.1 |
| <i>Neurospora crassa</i> | CAE76504.1 |
| <i>Plasmodium falciparum</i> | XP_001350109.1 |
| <i>Leishmania donovani</i> | 3JCS (PDB ID) |
| <i>Euglena gracilis</i> | QLA09617.1 |
| <i>Toxoplasma gondii</i> | PIM00564.1 |
| <i>Tetrahymena thermophila</i> | XP_001019693.2 |

Table 4: Accession IDs for uL22 ribosomal Protein

| Organism | Accession ID |
|---------------------------------|---------------------|
| <i>Triticum aestivum</i> | XP_044431146.1 |
| <i>Oryza sativa</i> | XP_015612256.1 |
| <i>Solanum lycopersicum</i> | XP_004236338.1 |
| <i>Arabidopsis thaliana</i> | NP_174060.1 |
| <i>Saccharomyces cerevisiae</i> | AJS47412.1 |
| <i>Candida albicans</i> | XP_712916.1 |
| <i>Drosophila melanogaster</i> | NP_001284961.1 |
| <i>Caenorhabditis elegans</i> | NP_740781.1 |
| <i>Homo sapiens</i> | KAI2586824.1 |
| <i>Elephantulus edwardii</i> | XP_006903542.1 |
| <i>Plasmodium falciparum</i> | XP_001350227.1 |
| <i>Euglena gracilis</i> | QLA09612.1 |
| <i>Escherichia coli</i> | WP_266146827.1 |
| <i>Salmonella enterica</i> | NRK38637.1 |

Table 5: Interacting residues at eL34 and 28S rRNA interface

| eL34 residues | 25S rRNA residues |
|----------------------|----------------------------------|
| K49 | C1708 |
| K50 | A1706 |
| Q52 | A1707 |
| G53 | G1639 (through K ⁺)* |
| R58 | G1591 |
| H56 | A1740 |
| R56 | A1655 |
| T60 | A1592 |
| K63 | G1616 |
| R64 | U1822 |
| R66 | A1593 |
| S68 | A1642 |
| R69 | G1644 |
| N70 | G1639 (through K ⁺)* |
| R71 | C1739 |
| R72 | U1822 |
| N75 | C1632 |
| R76 | C1638 |

*The interaction between the eL34 residues and rRNA at these positions is coordinated through K⁺ ion

Table 6: Accession IDs for eL34 ribosomal protein

| Organism | Accession ID |
|----------------------------------|---------------------|
| <i>Elephantulus edwardii</i> | XP_006884412 |
| <i>Macaca fascicularis</i> | XP_045253972.1 |
| <i>Oryctolagus cuniculus</i> | PDB: 7NFX_g |
| <i>Rattus rattus</i> | CAA32574.1 |
| <i>Homo sapiens</i> | NP_000986.2 |
| <i>Saccharomyces cerevesiae</i> | PDB: 4V6I_Bi |
| <i>Podila horticola</i> | KAG0023736.1 |
| <i>Rhizopus delemar</i> | EIE75655.1 |
| <i>Kluyveromyces lactis</i> | NP_010977.2 |
| <i>Candida glabrata</i> | XP_449882.1 |
| <i>Triticum aestivum</i> | AAW50987.1 |
| <i>Cocos nucifera</i> | XP_008802610.1 |
| <i>Zingiber officinale</i> | XP_042388246.1 |
| <i>Zea mays</i> | NP_001132454.1 |
| <i>Coffea arabica</i> | XP_027075604.1 |
| <i>Helianthus annuus</i> | KCW64962.1 |
| <i>Pisum sativum</i> | KAI5420024.1 |
| <i>Solanum lycopersicum</i> | XP_004238799.1 |
| <i>Picea sitchensis</i> | ABK21342.1 |
| <i>Adiantum capillus</i> | KAI5062804.1 |
| <i>Ceratopteris richardii</i> | KAH7284748.1 |
| <i>Sphagnum magellanicum</i> | KAH9559090.1 |
| <i>Ceratodon purpureus</i> | KAG0605459.1 |
| <i>Physcomitrium patens</i> | XP_024360062.1 |
| <i>Pycnococcus provasolii</i> | GHP03976.1 |
| <i>Coccomyxa sp</i> | BDA42094.1 |
| <i>Volvox africanus</i> | GIL50355.1 |
| <i>Ostreococcus tauri</i> | XP_022838836.1 |
| <i>Micromonas commoda</i> | XP_002504814.1 |
| <i>Euglena gracilis</i> | QLA09635.1 |
| <i>Leishmania donovani</i> | XP_001686957.1 |
| <i>Trichomonas vaginalis</i> | XP_001303734.1 |
| <i>Trypanosoma cruzi</i> | XP_815813.1 |
| <i>Eimeria necatrix</i> | XP_013335039.1 |
| <i>Plasmodium malariae</i> | SBT74739.1 |
| <i>Toxoplasma gondii</i> | XP_002366393.1 |
| <i>Babesia ovata</i> | XP_028868707.1 |
| <i>Cryptosporidium ubiquitum</i> | XP_028875035.1 |
| <i>Myxobolus squamalis</i> | KAF1744176.1 |
| <i>Nosema bombycis</i> | ADZ95696.1 |
| <i>Encephalitozoon cuniculi</i> | NP_597582.1 |

| | |
|------------------------------------|----------------|
| <i>Paranosema locustae</i> | PDB: 6ZU5_LGG |
| <i>Tetrahymena thermophila</i> | XP_001021598.1 |
| <i>Blepharisma stoltei</i> | CAG9333327.1 |
| <i>Paramecium sonneborni</i> | CAD8077493.1 |
| <i>Ichthyophthirius multifilii</i> | XP_004037728.1 |
| <i>Stylonychia lemnae</i> | CDW83538.1 |
| <i>Methanothermobacter sp.</i> | XP_001021598.1 |
| <i>Nanoarchaeota sp.</i> | MCK4589265.1 |
| <i>Methanomicrobia sp.</i> | MCD6128027.1 |
| <i>Thermococci sp.</i> | RLF91603.1 |
| <i>Euryarchaeota sp.</i> | MBM4240109.1 |

Table 7: Accession IDs for uL4 ribosomal protein

| Organism | Accession ID |
|-------------------------------------|---------------------|
| <i>Panicum miliaceum</i> | RLM84596.1 |
| <i>Setaria italica</i> | XP_004981434.1 |
| <i>Zea mays</i> | ACG25262.1 |
| <i>Oryza sativa Japonica Group</i> | XP_015647011.1 |
| <i>Triticum aestivum</i> | 4V7E_CC |
| <i>Brassica rapa</i> | XP_009146912.2 |
| <i>Gossypium arboreum</i> | XP_017649246.1 |
| <i>Adiantum capillus-veneris</i> | KAI5082907.1 |
| <i>Ceratopteris richardii</i> | KAH7365535.1 |
| <i>Sphagnum magellanicum</i> | KAH9563930.1 |
| <i>Ceratodon purpureus</i> | KAG0597514.1 |
| <i>Physcomitrium patens</i> | XP_024362557.1 |
| <i>Pycnococcus provasolii</i> | GHP11437.1 |
| <i>Ostreococcus tauri</i> | XP_022840336.1 |
| <i>Volvox africanus</i> | GIL42455.1 |
| <i>Coccomyxa sp. Obi</i> | BDA44074.1 |
| <i>Micromonas commoda</i> | XP_002506564.1 |
| <i>Homo sapiens</i> | NP_000959.2 |
| <i>Mus musculus</i> | NP_077174.1 |
| <i>Drosophila melanogaster</i> | NP_524538.2 |
| <i>Arabidopsis thaliana</i> | NP_187574.1 |
| <i>Tetrahymena thermophila</i> | XP_001017488.2 |
| <i>Plasmodium falciparum</i> | XP_001351629.1 |
| <i>Saccharomyces cerevisiae</i> | NP_009587.1 |
| <i>Thermochaetoides thermophila</i> | XP_006696470.1 |

Table 8: Source of the sequences of 5.8S rRNA

| Organism | Source of Sequence |
|---------------------------------|---------------------------|
| <i>Triticum aestivum</i> | 4V7E |
| <i>Solanum lycopersicum</i> | 7QIW |
| <i>Homo sapiens</i> | 6QZP |
| <i>Oryctolagus cuniculus</i> | 7O7Y |
| <i>Saccharomyces cerevesiae</i> | 4V88 |
| <i>Neurospora crassa</i> | 6YWS |
| <i>Kluyveromyces lactis</i> | 6UZ7 |
| <i>Plasmodium falciparum</i> | 5UMD |
| <i>Tetrahymena thermophila</i> | 4V8P |
| <i>Toxoplasma gondii</i> | 5XXB |
| <i>Euglena gracillis</i> | 6ZJ3 |