

SUPPLEMENTARY DATA

The association of late-acting snoRNPs with human pre-ribosomal complexes requires the RNA helicase DDX21

Katherine E. Sloan, Matthias S. Leisegang, Carmen Doebele, Ana S. Ramírez, Stefan Simm, Charlotta Safferthal, Jens Kretschmer, Tobias Schorge, Stavroula Markoutsas, Sara Haag, Michael Karas, Ingo Ebersberger, Enrico Schleiff, Nicholas J. Watkins, Markus T. Bohnsack

Supplementary Figure S1. Evolutionary conservation of DDX21 and DDX50.

Evolutionary relationships of the 160 DDX21/50 orthologues identified are shown. Metazoa and close relatives are highlighted using blue lines, fungi are indicated by red lines, and plants/green algae are visualized by green lines. The domain architectures of the individual proteins are summarized (brackets) across clades (triangles) and schematic overviews of the major domains (DEAD, Helicase, GUCT, RRM) are given. The scale bar corresponds to 0.8 substitutions per nucleotide position.

Supplementary Figure S2. The pre-ribosomal distribution and total levels of SNORD88 and SNORD14 are not affected by DDX21 depletion. (A) The major DDX21 LSU crosslinking site is indicated in red on the 2D structure of a fragment of domain IV of the 28S rRNA. **(B)** RNA from HEK293 cells treated with non-target siRNA (siNT) or siRNA targeting DDX21 (siDDX21) was separated on a denaturing polyacrylamide gel and transferred to a nitrocellulose membrane. Northern blotting was performed to analyse total snoRNA levels using probes specific for the snoRNAs indicated to the left of the panels. **(C-D)** Extracts prepared from cells transfected with a non-target siRNA (siNT) or siRNAs targeting DDX21 (siDDX21) were separated by sucrose-gradient density centrifugation. RNA was isolated from the fractions, separated by polyacrylamide gel electrophoresis and analysed by Northern blotting using a probe hybridising to SNORD88 **(C)** or SNORD14 **(D)**. Representative blots of

three independent experiments are shown. Quantification of three independent experiments are presented above the Northern blots for each fraction as percentage of the total snoRNA signal (data are presented as mean +/- SEM); fractions containing free snoRNPs or specific pre-ribosomal complexes are indicated below.

Supplementary Figure S3. Modification of U428 of the 18S rRNA sequence by SNORD68 is reduced in cells depleted of DDX21. Cells stably expressing RNAi-resistant, Flag-tagged DDX21, DDX21_{SAT} or only the FLAG-tag were transfected with non-target siRNAs (siNT) or those targeting DDX21 (siDDX21) as indicated. RNA was extracted and analysed by site-directed RNase H cleavage to monitor rRNA methylation of U428 in the 18S rRNA. RNA fragments were separated by agarose-glyoxal gel electrophoresis followed by Northern blotting using a probe hybridising to the 5' end of ITS1.

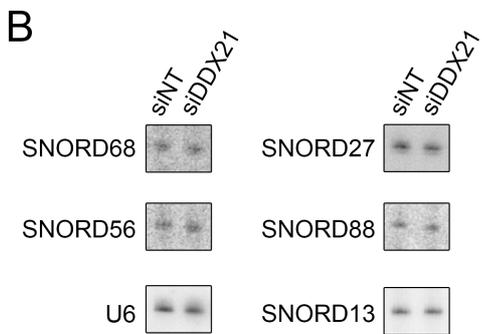
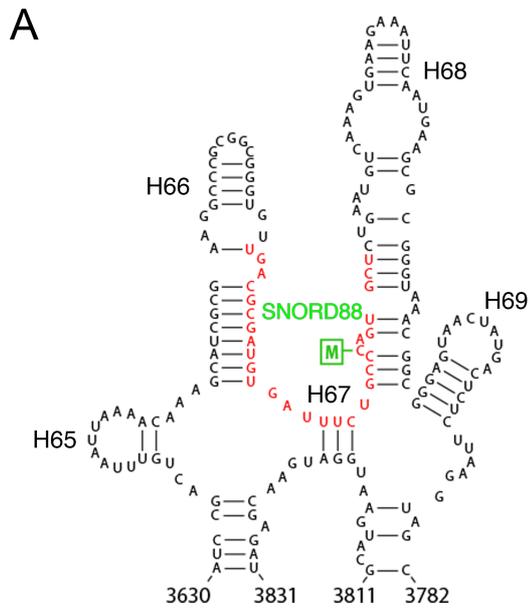
Supplementary Figure S4. Basepairing sites of selected snoRNAs with the 18S rRNA sequence. 2D structure of human 18S rRNA with the 2'-O-methylations (boxed M). Basepairing sites of snoRNAs investigated in this study are highlighted in green and additional basepairing sequences identified in (1) are shown in thick grey lines with thin grey lines linking basepairing regions belonging to a single snoRNA.

Supplementary Figure S5. The pre-ribosomal distribution of the late-associated snoRNAs SNORD27, SNORD25 and SNORD13 are not affected by depletion of DDX21. (A-C) Extracts prepared from cells transfected with a non-target siRNA (siNT) or siRNAs targeting DDX21 (siDDX21) were separated by sucrose-gradient density centrifugation. RNA was isolated from the fractions, separated by polyacrylamide gel electrophoresis and analysed by Northern blotting using probes hybridising to SNORD27 **(A)** SNORD25 **(B)** and SNORD13 **(C)**. Representative blots of three independent experiments are shown. Quantification of three independent

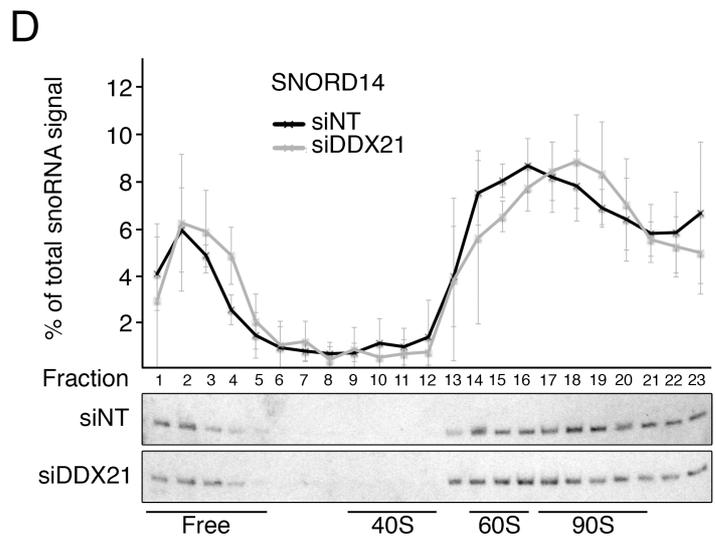
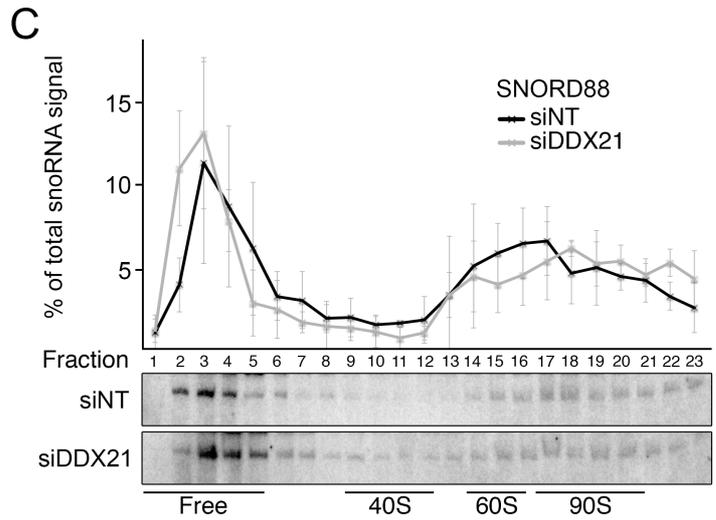
experiments are presented above the Northern blots for each fraction as percentage of the total snoRNA signal (data are presented as mean \pm SEM); fractions containing specific pre-ribosomal complexes are indicated below.

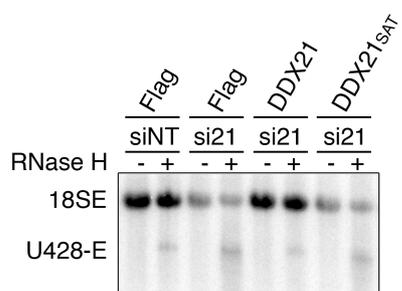


Sloan et al., Figure S1

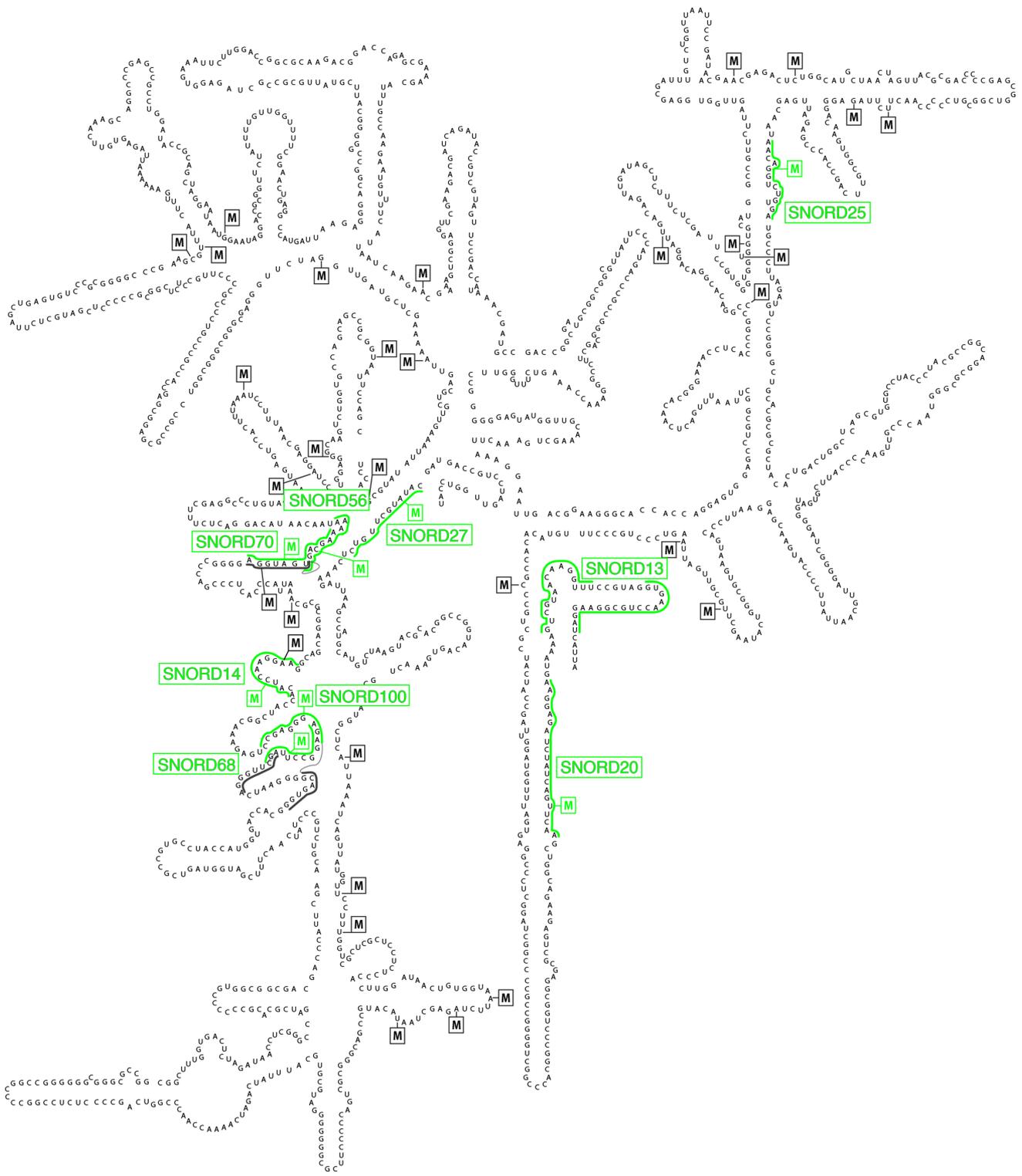


Sloan et al., Figure S2

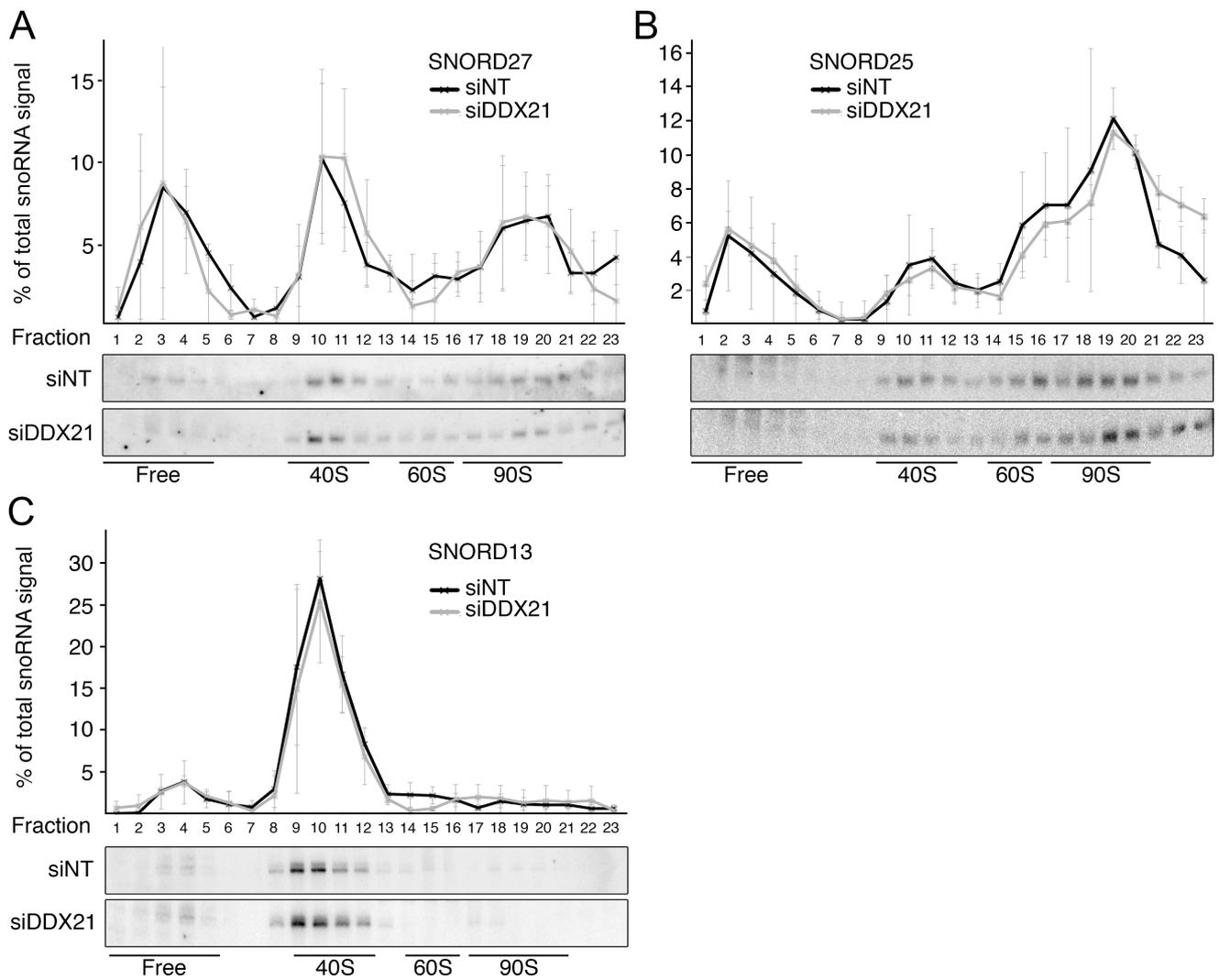




Sloan et al., Figure S3



Sloan et al., Figure S4



Sloan et al., Figure S5

Supplementary Table S1. Proteins identified by mass spectrometry after PWP2-Flag pulldown using human cell nuclear extracts. Proteins identified by mass spectrometry after PWP2-Flag pulldown using human cell nuclear extracts. Nuclear extracts from HEK293 cells stably expressing Flag-tagged PWP2 or only the tag were incubated with anti-Flag magnetic beads and eluates were analysed by mass spectrometry (MS). The table lists the proteins identified for PWP2-Flag (but not the Flag control) and how frequently they were identified in either three MS/MS or nano-LC MS/MS experiments.

Gene	Protein	Yeast homolog	No. of peptides found	Nano-LC MS/MS	MS/MS	Uniprot Acc. Number
PWP2	Periodic tryptophan protein 2 homolog	Pwp2	44	3 of 3	3 of 3	Q15269
WDR36	WD repeat-containing protein 36	Utp21	24	3 of 3	2 of 3	Q9NR30
UTP12	WD repeat-containing protein 3	Utp12	5	3 of 3	0 of 3	Q8NI36
TBL3	Transducin beta-like protein 3	Utp13	9	3 of 3	0 of 3	Q9UNX4
DDX21	Nucleolar RNA helicase 2	not found	6	2 of 3	2 of 3	Q12788
NOP2	NOP2 nucleolar protein homolog	Nop2	12	1 of 3	1 of 3	P46087
NPM1	Nucleophosmin	not found	11	2 of 3	2 of 3	P06748
FBL	rRNA 2'-O-methyltransferase Fibrillarlin	Nop1	12	1 of 3	2 of 3	P22087

SUPPLEMENTARY MATERIALS AND METHODS

Complex analysis by Western blotting and mass spectrometry

To analyse eluates of pulldown experiments Western blotting was performed according to standard protocols using the following antibodies: DDX21 (Bethyl A300-628A), PWP2 (GeneTex GTX105344), NOP2 (Bethyl A302-018A), UTP18 (Bethyl A301-551A), UTP6 (Sigma-Aldrich HPA025936), WDR36 (Sigma-Aldrich WHO0134430), TBL3 (Sigma-Aldrich HPS042562), Nucleophosmin (Sigma-Aldrich B0556), Fibrillarin (Santa Cruz H140) and Tubulin (Sigma-Aldrich T6199). MS/MS of proteins extracted from Coomassie-stained bands after in-gel tryptic digestion was performed as previously described (2). In addition, nano-LC MS/MS was performed on PWP2-Flag pulldown eluates as in (3). Proteins were identified using mascot v2.2.03 (Matrix Science Ltd.).

Northern Blotting

For snoRNA analysis, RNA was separated on a 10% denaturing polyacrylamide gel and transferred to nitrocellulose membrane. For Northern blotting, ³²P-labelled antisense transcripts of the full length snoRNAs were used for the detection of SNORD56, SNORD68, SNORD25, SNORD20, SNORD27 and SNORD88. Additional snoRNAs were detected using 5' end ³²P-labelled oligonucleotides (SNORD14 5'-CTCAGACATCCAAGGAAGGTTTACCC-3' and 5'-AGCACTTCTGGTGGAACTACGAATGGTTT-3'; SNORD70 5'-TATTCGTCACTACCACTGAGACAACGATGA-3' and 5'-TTCATCGTTGTCTCAGTGG-3'; SNORD100 5'-GTAGAGGGAGCCAGTTGTCATCATGTAC-3' and 5'-GGGTGACATGGCAGTTTCCTCATGGC-3'; SNORD13 5'-GGTCAGACGGGTAATGTGCCC-3' and 5'-CGTAACAAGGTTCAAGGGTGGC-3'). Note, oligonucleotide probes were used in combination after verification that each one was specific for the targeted snoRNA.

Phylogenetic analysis of DDX21

HaMStR-OneSeq (4) was used to search for orthologues of DDX21 and its paralogue DDX50 in 282 eukaryotes, 12 archaea, and 16 bacteria. Phylogenetic tracing of DDX21/50 in the eukaryotic tree was performed using only the 160 orthologues that contain the DEAD/DEAH box helicase, helicase_C and GUCT domain that characterise the human proteins as identified by FACT (5). These sequences were aligned with MUSCLE (6) using default settings. The alignment was post-processed by removing all alignment columns with more than 50% of gaps or undetermined amino acids. The resulting alignment was used as an input for maximum likelihood tree reconstruction using RaXML (7) with the PROTGAMMAILGF model of sequence evolution. While the resulting tree did not support the monophyly of the opisthokonts (animals and fungi) a subsequent tree evaluation using the tests implemented in Tree-Puzzle (8) revealed that the tree with monophyletic opisthokonts is not significantly worse than the ML tree and was, therefore, used. The tree was processed with FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>) and individual clades were collapsed to improve visualisation. Notably, DDX21 and DDX50 were only distinguishable within the Theria, suggesting that the gene duplication that gave rise to these two genes took place in the last common ancestor of the Theria. Therefore we refer to the proteins in all clades as DDX21/50. Pfam-domains in the 160 proteins were annotated with hmmscan from the hmmer package v. 3.b1 (<http://hmmer.janelia.org>) and assigned to the corresponding nodes in the tree. While the vast majority of sequences agreed in their domain architecture with that of the human proteins, additional domains were identified in two specific clades: in flowering plants and the bryophyte *Physcomitrella patens* an additional zinc finger motif at the C-terminus was detected and DDX21/50 in *Cryptosporidium parvum* and *Toxoplasma gondii* contain an additional RNA recognition motif (RRM_1; PF00076) at the C-terminus. Together these data suggest that human DDX21/50 was already present, with the same functional domains, in the last common ancestor of all

eukaryotes.

SUPPLEMENTARY REFERENCES

1. van Nues,R.W., Granneman,S., Kudla,G., Sloan,K.E., Chicken,M., Tollervey,D. and Watkins,N.J. (2011) Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing. *EMBO J.*, **30**, 2420-2430.
2. Ladig,R., Sommer,M.S., Hahn,A., Leisegang,M.S., Papatirou,D.G., Ibrahim,M., Elkehal,R., Karas,M., Zickermann,V., Gutensohn,M., Brandt,U., Klösgen,R.B. and Schleiff,E. (2011) A high-definition native polyacrylamide gel electrophoresis system for the analysis of membrane complexes. *Plant J.*, **67**, 181-94.
3. Simm,S., Papatirou,D.G., Ibrahim,M., Leisegang,M.S., Muller,B., Schorge,T., Karas,M., Mirus,O., Sommer,M.S. and Schleiff,E. (2013) Defining the core proteome of the chloroplast envelope membranes. *Front. Plant Sci.*, **4**, 11.
4. Ebersberger,I., Simm,S., Leisegang,M.S., Schmitzberger,P., Mirus,O., von Haeseler,A., Bohnsack,M.T. and Schleiff,E. (2014) The evolution of the ribosome biogenesis pathway from a yeast perspective. *Nucleic Acids Res.*, **42**, 1509-1523.
5. Koestler,T., von Haeseler,A. and Ebersberger,I. (2010) FACT: functional annotation transfer between proteins with similar feature architectures. *BMC Bioinform.*, **11**, 417.
6. Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792-1797.
7. Stamatakis,A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688-2690.
8. Schmidt,H.A., Strimmer,K., Vingron,M. and von Haeseler,A. (2002) TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics*, **18**, 502-504.