Supplemental Material to:

Mohamed Hafez, Gertraud Burger,
Sergey V. Chteinberg, Franz B. Lang

A second eukaryotic group with mitochondrion-encoded tmRNA: In silico identification and experimental confirmation

2012; 10(8)
http://dx.doi.org/10.4161/rna.25376

www.landesbioscience.com/journals/rnabiology/article/25376/

SUPPLEMENTAL DATA

Supplement to the manuscript :

A second eukaryotic group with mitochondrion-encoded tmRNA: *In silico* identification and experimental confirmation

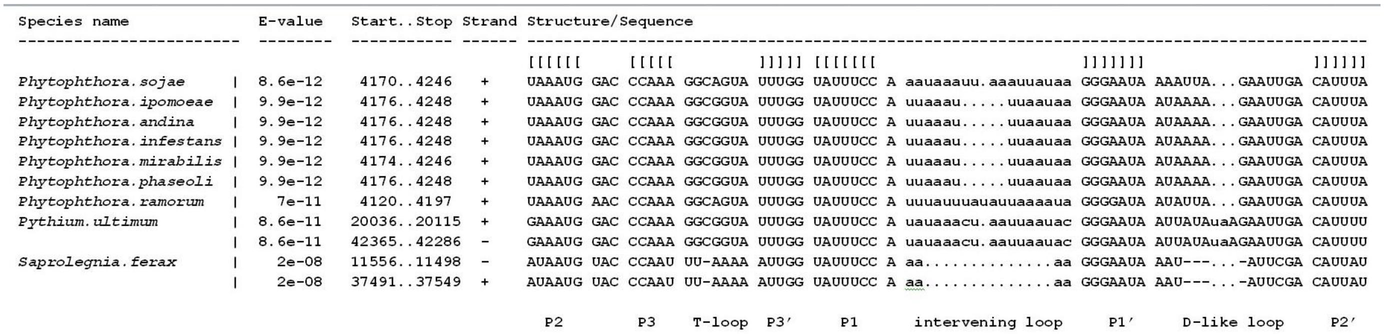
Mohamed Hafez, Gertraud Burger, Sergey V. Steinberg and B. Franz Lang

AUTHORS' CONTRIBUTIONS

BFL conceived the study, developed the mt-tmRNA covariance model, identified corresponding genes in complete oomycete mtDNAs, and evaluated RNA-Seq data for mapping of the *P. sojae* RNA termini. MH identified additional tmRNA genes in GenBank, wrote the first manuscript draft, and designed figures. SVS established RNA tertiary models. All authors participated in the discussion of results, and production of the final manuscript version.

SUPPLEMENTAL TABLES

Table S1. Search results with a covariance mt-tmRNA model (based on jakobid sequences) in nine complete oomycete mitochondrial genome sequences. Note a high statistic support despite the short length of the gene sequence, which is due to conservation of both primary sequence and RNA structure. *S. ferax* has the shortest sequence and the most derived features, with no apparent sequence conservation in the T-loop, and a reduced-size T-loop and D-loop.



```
Species name             E-value   Start..Stop Strand Structure/Sequence
------------------------ --------  ----------- ------ -------------------------------------------------------------------------------------------
                                                      [[[[[[     [[[[[         ]]]]] [[[[[[[                       ]]]]]]]                 ]]]]]]
Phytophthora.sojae     | 8.6e-12    4170..4246   +    UAAAUG GAC CCAAA GGCAGUA UUUGG UAUUUCC A aauaaauu.aaauuauaa GGGAAUA AAAUUA...GAAUUGA CAUUUA
Phytophthora.ipomoeae  | 9.9e-12    4176..4248   +    UAAAUG GAC CCAAA GGCGGUA UUUGG UAUUUCC A uuaaau.....uuaauaa GGGAAUA AUAAAA...GAAUUGA CAUUUA
Phytophthora.andina    | 9.9e-12    4176..4248   +    UAAAUG GAC CCAAA GGCGGUA UUUGG UAUUUCC A uuaaau.....uuaauaa GGGAAUA AUAAAA...GAAUUGA CAUUUA
Phytophthora.infestans | 9.9e-12    4176..4248   +    UAAAUG GAC CCAAA GGCGGUA UUUGG UAUUUCC A uuaaau.....uuaauaa GGGAAUA AUAAAA...GAAUUGA CAUUUA
Phytophthora.mirabilis | 9.9e-12    4174..4246   +    UAAAUG GAC CCAAA GGCGGUA UUUGG UAUUUCC A uuaaau.....uuaauaa GGGAAUA AUAAAA...GAAUUGA CAUUUA
Phytophthora.phaseoli  | 9.9e-12    4176..4248   +    UAAAUG GAC CCAAA GGCGGUA UUUGG UAUUUCC A uuaaau.....uuaauaa GGGAAUA AUAAAA...GAAUUGA CAUUUA
Phytophthora.ramorum   |   7e-11    4120..4197   +    UAAAUG AAC CCAAA GGCAGUA UUUGG UAUUUCC A uuuauuuauauuaaaaua GGGGAUA AUAUUA...GAAUUGA CAUUUA
Pythium.ultimum        | 8.6e-11   20036..20115  +    GAAAUG GAC CCAAA GGCGGUA UUUGG UAUUUCC A uauaaacu.aauuaauac GGGAAUA AUUAUAuaAGAAUUGA CAUUUU
                       | 8.6e-11   42365..42286  -    GAAAUG GAC CCAAA GGCGGUA UUUGG UAUUUCC A uauaaacu.aauuaauac GGGAAUA AUUAUAuaAGAAUUGA CAUUUU
Saprolegnia.ferax      |   2e-08   11556..11498  -    AUAAUG UAC CCAAU UU-AAAA AUUGG UAUUUCC A aa..............aa GGGAAUA AAU---...-AUUCGA CAUUAU
                       |   2e-08   37491..37549  +    AUAAUG UAC CCAAU UU-AAAA AUUGG UAUUUCC A aa..............aa GGGAAUA AAU---...-AUUCGA CAUUAU

                                                        P2        P3   T-loop P3′     P1      intervening loop      P1′       D-like loop   P2′
```

Table S2. Complete list of currently known mt-tmRNAs in jakobids and oomycetes. Sequences identified by covariance search in complete mtDNAs are marked gray.

Species	**GenBank accession number**	**Note**
Andalucia godoyi	KC353352	complete mitochondrial genome
Histiona aroides	KC353353	complete mitochondrial genome
Jakoba libera	KC353355	complete mitochondrial genome
Reclinomonas americana ATCC50633	KC353356	complete mitochondrial genome
R. americana ATCC 50283	KC353357	complete mitochondrial genome
R. americana ATCC 50284	KC353358	complete mitochondrial genome
R. americana ATCC 50394	NC_001823	complete mitochondrial genome
Seculamonas ecuadoriensis	KC353359	complete mitochondrial genome
Albugo candida	CAIV01000124	whole genome shotgun sequence
A. laibachii	FR825415	Illumina short reads
Bremia lactucae	JP948723	Transcriptome Shotgun Assembly
Phytophthora alni	DQ162893	Partial sequence
P. andina	NC_015619	complete mitochondrial genome
P. cactorum	DQ162892	Partial sequence
P. cambivora	DQ162894	Partial sequence
P. capsici	DQ162899	Partial sequence
P. cinnamomi	JF746382	Partial sequence
P. citricola	DQ162900	Partial sequence
P. citrophthora	DQ162901	Partial sequence
P. cryptogea	DQ162908	Partial sequence
P. drechsleri	DQ162910	Partial sequence
P. erythroseptica	DQ162909	Partial sequence
P. europaea	DQ162895	Partial sequence
P. fragariae	DQ162896	Partial sequence
P. infestans	NC_002387	complete mitochondrial genome
P. inundata	DQ162902	Partial sequence
P. ipomoeae	NC_015622	complete mitochondrial genome
P. katsurae	DQ162904	Partial sequence
P. kernoviae	DQ162914	Partial sequence
P. lateralis	DQ162912	Partial sequence
P. medicaginis	DQ162911	Partial sequence
P. mirabilis	NC_015606	complete mitochondrial genome
P. nicotianae	JF706917	Partial sequence
P. palmivora	DQ162906	Partial sequence
P. phaseoli	NC_015616	complete mitochondrial genome
P. pseudosyringae	DQ162907	Partial sequence
P. quercina	DQ162905	Partial sequence
P. ramorum	NC_009384	complete mitochondrial genome
P. sojae	NC_009385	complete mitochondrial genome
Pseudoperonospora cubensis	AHJF01014480	whole genome shotgun sequence
Pythium ultimum	NC_014280	complete mitochondrial genome
Saprolegnia ferax	NC_005984	complete mitochondrial genome

Figure S1. Standard tRNA secondary structure

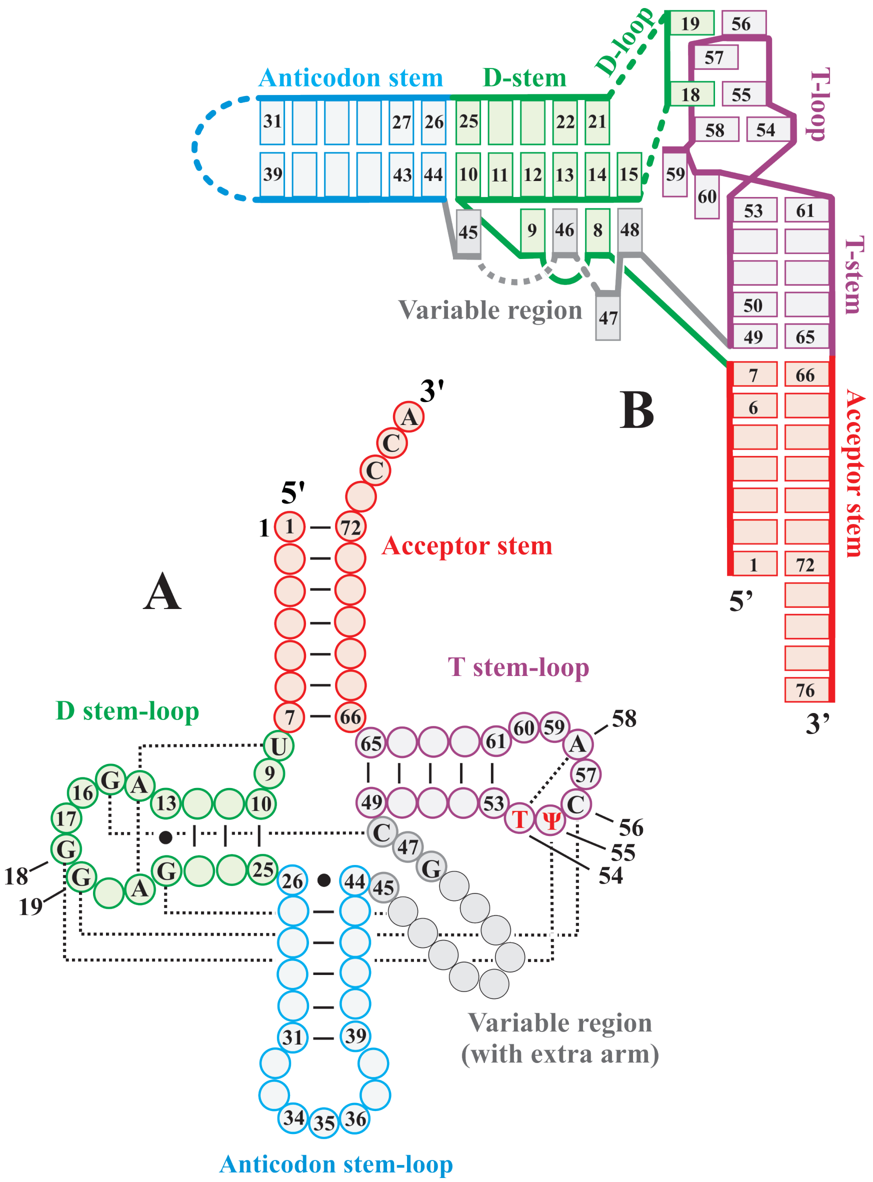
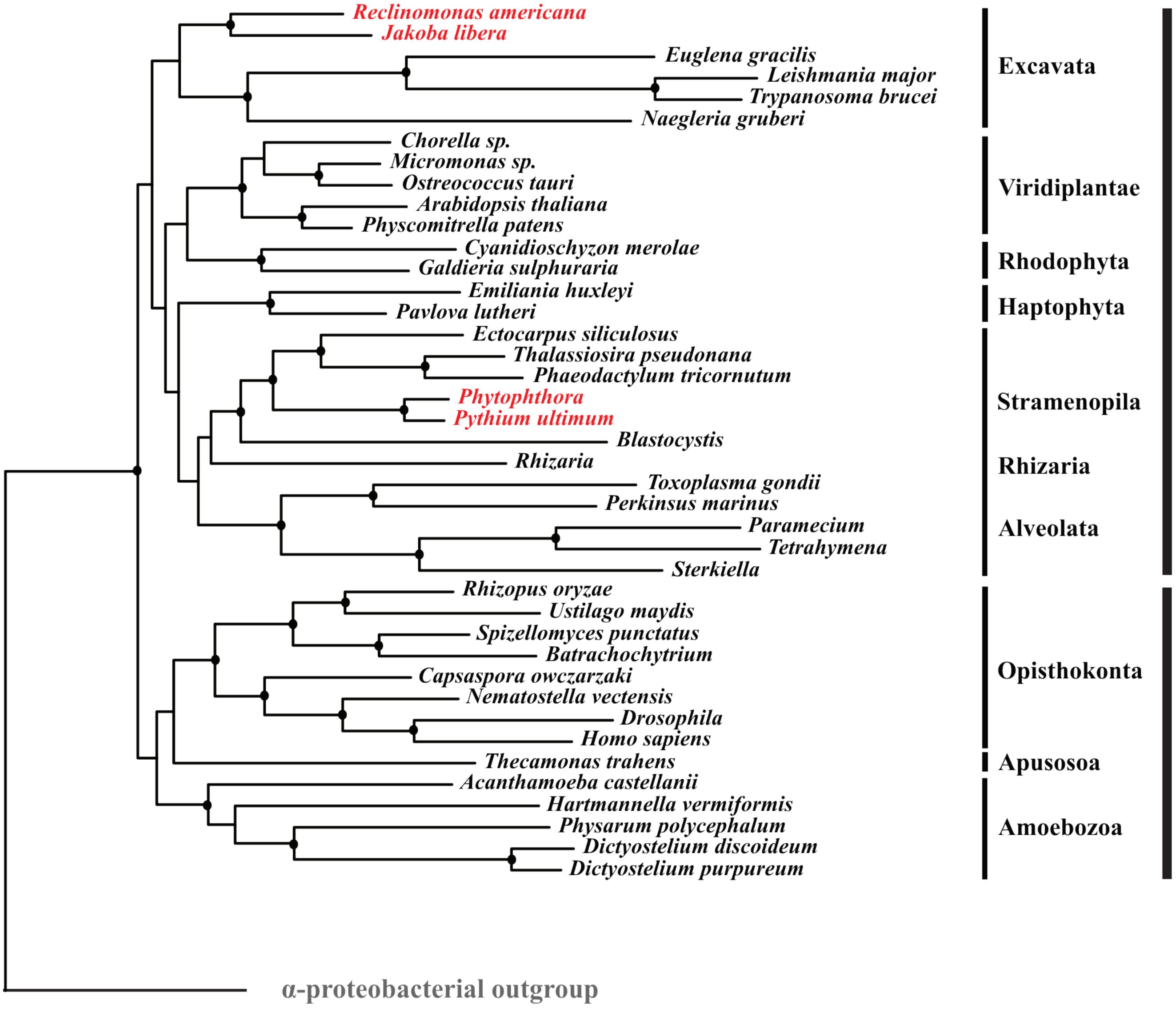


Figure adapted from (Lang et al. 2011). **A, cloverleaf.** Each stem-loop is shown in a different color. Short lines connect nucleotides forming Watson-Crick base pairs within stems. Black filled circles (base pairs 13-22 and 26-44) indicate that many tRNAs have non-Watson-Crick interactions. Dotted lines connect nucleotides involved in conserved tertiary interactions. Nucleotides are numbered in accordance with the standard tRNA nomenclature, based on the yeast Phe-tRNA (Rich and RajBhandary 1976). Note that numerous cytosolic tRNAs do not fit this structure. For instance, nucleotide 47 does not exist in some tRNAs, and position 17 may be either empty or additional nucleotides may be inserted between positions 17 and 18. Further, some tRNAs have one or two additional nucleotides between positions 20 and 21. Finally, the variable region located between the anticodon and T stem-loop, which minimally contains three nucleotides, may be extended by eight or more nucleotides and form an extra arm.

B, L-form. Rectangles represent individual nucleotides. Nucleotides of stem-loops are shown in the same color and with the same numbering as in A. Nucleotides of the anticodon loop and non-stacked nucleotides of the D-loop (16, 17 and 20) are not shown (indicated by dashed lines). The dashed gray line connecting nucleotides 46 and 47 of the variable region indicate that additional non-numbered nucleotides may be inserted and form the extra arm.

Tertiary interactions U8-A14 and 15-48 are found in the D-stem-loop region of all cytosolic tRNAs. The presence of a Watson-Crick or U-G base pair 13-22 and of the tertiary contacts 9-23 and 22-46 constitutes the standard pattern of tertiary interactions. This pattern may also be accompanied by the tertiary contact 10-45. In the DT-region, there are two inter-loop base pairs G18-Ψ55 and G19-C56 separated by a purine-57. The dinucleotide 59-60 bulges between base pairs 53-61 and 54-58. Nucleotide 59 stacks to the tertiary base pair 15-48, which constitutes the last layer of the D/anticodon helical domain. This interaction stabilizes the perpendicular arrangement of the two helical domains of the tRNA L-shape.

Figure S2. Occurrence of mt-tmRNAs across eukaryotes



The tree (adapted from (Derelle and Lang 2011)) has been inferred with mtDNA- plus nucleus-encoded mitochondrial proteins, and is rooted with α-Proteobacteria (from which mitochondria derive). Species representatives that have mtDNA-encoded tmRNAs are marked in red. Note that tmRNAs are only identified in sub-lineages of two very distant organismal groups, jakobids (Excavata) and oomycetes (Stramenopila).

```
# STOCKHOLM 1.0

Recli.amer.94      ACTATGGAACCGAGGGCGGATCTCGGCATCTCCAtaaaaatgtttg-----------ataacattttttGGGGATGTTTTTA----GTATTCGACATAGT
Recli.amer.83      ACTATGGACCCGAGGGCAGTTCTCGGCATCTCCAtttagatat-----------------tgtttttaaGGGGATGTTTTTA----GGATTCGACATAGT
Recli.amer.33      GCTATGGACCCGAGGGCAGTTCTCGGCATCTCCAtctaaagata---------------ttctttttaaGGGGATGTTTTTA----GGATTCGACATAGT
Recli.amer.84      ACTATGGACCCGAGGGCAGTTCTCGGCATCTCCAtctaaaaaaa---------------ttttttttaaGGGGATGTTTTTA----GAATTCGACATAGT
Histiona.aroides   ACTATGGACCCGTGGGTGGAACACGGCATCTCCAggccctg------------------ctattttatgGGGGGTGTTAATA-----AATTCGACATAGT
Andalucia          AGAAAGGAACCGATCT-CATCTTCGGCATTTCCAattct-----------------------atctttcGGGAATGAATTG------TTTTCGACTTTCT
Seculamonas        CGAATGGACCCATGGGCAGTTCATGGCATTCCCAgtaacaacatcagagta-gtttttatattgttttaGGGAATGATATATactaGAGATCGACATTCA
Phyt.medicaginis   TAAATGGACCCAAAGGCGGTATTTGGTATTTCCAtttaaaa---------------------attaaaaGGGAATAATAATA----GAATT-GACATTTA
Phyt.sojae         TAAATGGACCCAAAGGCAGTATTTGGTATTTCCAaataaatt------------------aaattataaGGGAATAAAATTA----GAATT-GACATTTA
Phyt.infestans     TAAATGGACCCAAAGGCGGTATTTGGTATTTCCAttaaat----------------------ttaataaGGGAATAATAAAA----GAATT-GACATTTA
Phyt.cinnamomi     TAAATGGACCCAGAGGCAGTTTTTGGTATTTCCAtataaatt-------------------aaatgactGGGGATAATAATA----GAATT-GACATTTA
Phyt.citrophthora  AAAATGGACCCAAAGGCAGTATTTGGTATTTCCAtattgta--------------------aattaataGGGAATAATAATA----GAATT-GACATTTA
Phyt.ramorum       TAAATGAACCCAAAGGCAGTATTTGGTATTTCCAtttatttat-----------------attaaaataGGGGATAATATTA----GAATT-GACATTTA
Pythium.ultimum    GAAATGGACCCAAAGGCGGTATTTGGTATTTCCAtataaact------------------aattaatacGGGAATAATTATAt-aaGAATT-GACATTTT
Bremia.lactucae    TAAATGGACTCGAGGGCAGTATTTGATATTTCCAttttaa----------------------ttaatagGGGGATAATAATA----GAATT GACATTTA
Saprolegnia.ferax  ATAATGTACCCAATTT-AAAAATTGGTATTTCCAaa-------------------------------aaGGGAATAAAT---------ATTCGACATTAT
Phyt.katsurae      TAAATGGACTCAAAGGCAGTATTTGATATTTCCAtttatattatttattaattaaaaattaataaataaGGGAATAAAAATA----G-ATT GACATTTA
#=GC SS_cons       <<<<<<...<<<<<<.....>>>>>><<<<<<<....................................>>>>>>>..................>>>>>>
#=GC RF            BBBBBBbbbCCCCCCcccccCCCCCCDDDDDDDe-----------------------------------DDDDDDDdddddd----ffffffffBBBBBB
```

Dataset 1. Alignment used for building the mt-tmRNA CM model. Sequences that are not aligned with confidence are in lower case, corresponding sequence columns were masked. cmbuild was executed using '—hand'.

SUPPLEMENTAL REFERENCES

Derelle, R., and B. F. Lang. 2011. Rooting the eukaryotic tree with mitochondrial and bacterial proteins. Mol Biol Evol **29**:1277-1289.

Lang, B. F., D. Lavrov, N. Beck, and V. Steinberg. 2011. Mitochondrial tRNA structure, identity and evolution of the genetic code. Pp. 431-474 *in* C. E. Bullerwell, ed. Organelle Genetics. Springer, Berlin, Heidelberg, 2012.