

Supplementary Figure S1

Sequence alignment showing conservation across various organisms. The alignment is color-coded by residue type.

| Organism | Sequence | Length |
|-------------------------|--|--------|
| <i>Reclinomonas</i> | MIRICPIVRSKVPLLGTFLRSDSWLAPHALALRRAICKNVALRSYSVNSEQPKHTFDISKLTRNEIQQLRELKRAKERFKDRTVAFYFSSVAVLFLGLAYAAVPLYRAICARTGFGGT | 42 |
| <i>Saccharomyces</i> | + + MIRICPIVRSKVPLLGTFLRSDSWLAPHALALRRAICKNVALRSYSVNSEQPKHTFDISKLTRNEIQQLRELKRAKERFKDRTVAFYFSSVAVLFLGLAYAAVPLYRAICARTGFGGT | 119 |
| <i>Rhizobium</i> | | |
| <i>Sinorhizobium</i> | | |
| <i>Brucella</i> | | |
| <i>Rhodospaerooides</i> | | |
| <i>Ehrlichia</i> | | |
| <i>Anaplasma</i> | | |
| <i>Wolbachia</i> | | |
| <i>Rickettsia</i> | | |
| <i>Reclinomonas</i> | + + + MFKNRKSIAILIAAVSITMIGFSYGSVPLYRIFCQVTGFGGT | 42 |
| <i>Saccharomyces</i> | | |
| <i>Rhizobium</i> | MTNTPQTPPKERANGVIVGACLAFFVAGMVGMAYAAVPLYDMFCRVTGNGT | 51 |
| <i>Sinorhizobium</i> | MADNGQADRKRERSNGVIVGTCLAFFVAGMIGMAYAAVPLYDMFCRVTGNGT | 51 |
| <i>Brucella</i> | MTDQGENEKKQRNSNATIAVACLSFFVCMIGAAYASVPLYRIFCQVTGNGT | 52 |
| <i>Rhodospaerooides</i> | MSNQAPTPKSIDPKSKTAQALSLLVVFMGALAWASVPFYDWFCRVTFGGV | 52 |
| <i>Ehrlichia</i> | | |
| <i>Anaplasma</i> | MKPRIRNTIYGLIAILSMICLVYASVPLYSIFCKVTGNGT | 42 |
| <i>Wolbachia</i> | | |
| <i>Rickettsia</i> | MFKTKRNNNEVLTLLVSVIASMLCLTYASVPLYRIFCKATGNGT | 45 |
| <i>Reclinomonas</i> | MFSFLRKGSKNSIVFFLVLVSLVVSMICLAYASVPLYSIFCKATGNGT | 47 |
| <i>Saccharomyces</i> | | |
| <i>Rhizobium</i> | MSKSNKNLAFSLLGLMMMSVLLSFASVPIYNLFCKVTGNGT | 43 |
| <i>Sinorhizobium</i> | | |
| <i>Brucella</i> | | |
| <i>Rhodospaerooides</i> | | |
| <i>Ehrlichia</i> | | |
| <i>Anaplasma</i> | | |
| <i>Wolbachia</i> | | |
| <i>Rickettsia</i> | | |
| <i>Reclinomonas</i> | --TQVADLESIDLTLKDEQQENRIITVRFNGDVSDTMPWKFHPIQQIEKVMVGETALAFYSAENPTDSSIIIGISTYNVPQQAGIYFNKIQCFCEEQRKPHETIDMPVFFFIDPAILD | 160 |
| <i>Saccharomyces</i> | PITDERRKFTDDKLIPVDTEKR--IRISFTSEVSQILPWFVQQREVYVLPGETALAFYKAKNYSDKDIIGMATSYIAPGEAAQYFNKIQCFCEEQKLAAGEEIDMPVFFFIDPDFAS | 236 |
| <i>Rhizobium</i> | TQRVE--QASDVILDEK-----IKVTFDANVGPGLAWDKFVQRDIDVRIGETVQVMYRAKNLSSKPATGQATFNVT PMAAGAYFNKVQCFCFTETTLEPGEEMEMPVVFFDPEIVK | 162 |
| <i>Sinorhizobium</i> | TQRVE--QASDLILDEK-----IKVTFDANVAAGLPWEFVFPVQRDIDVRIGETVQIMYRAKNLASTPTTGQATFNVT PMAAGAYFNKVQCFCFTETTLEPGEEMEMPVVFFDPEIVK | 162 |
| <i>Brucella</i> | TQRVE--QYSDTILDKT-----IKVRFDANIANGLPWDFKPMQREVTVRIGETTMIKYEAHNLFGEEETYGRASFNVAPGRAGAYFNKVECFCTDNTLKPGEDLELPVVFFDPEFVN | 163 |
| <i>Rhodospaerooides</i> | TNTAD--TGSDEILDQT-----ITVRFDASKERGMPWEFKPMVREIEMRIGETGLVFYEAYNPTDRAVAGQASYNVAPYQAGYFFNKIAFCFCFEEQVLIQPGERVEREMPVTFYDPEIYT | 163 |
| <i>Ehrlichia</i> | VRTSN--ISNSKIGNTI-----IKVRFNADIHKQLPWFVYFPEVSHVVFVKPGEQKLIIFYRAENLLDEDTSAGMAVYNVTPHKVGKYFNKVCACFCFTKQTLYPHQKTIMPVSSFFIDPAIET | 153 |
| <i>Anaplasma</i> | TKEAASVAPVVKKTGSYN-----IKVRLNADTHG-IPLLHPEQPYVMVTPGVQQLAFYVAENPSDKEISGMAYNVT PFGQAGKYFYKIAFCFCFSQQTFAFGLKITMPVSSFFIAPEIE | 157 |
| <i>Wolbachia</i> | IRKV--NTTTNITNQK-----IRVHFNADIMS DLPWEFKSETNYVDVNIQEQLAFYYAKNLSQPSFGMAVYNVTPFKAGKYFNKVCACFCFEEQMILLPKQKAAMPVSFYIDPEIML | 158 |
| <i>Rickettsia</i> | TIKETVS VY SKVKGTKA-----IIIIFDANVDPNLPWHFIPRQKRVQIVPGQNTLVFYEAENLSN KDIIGTSIYNVTPNKAGKYFVKIHCFCFEEQLLKAREKVLMPPVTFYIDNDFER | 156 |
| <i>Reclinomonas</i> | DPKMSDIDSITLSYTFFNVEDI | 182 |
| <i>Saccharomyces</i> | DPAMRNIDDIILHYTFFRAHYGDGTAVSDSKKEPEMNADEKAASLANAAILSPEVIDTRKDNSN | 300 |
| <i>Rhizobium</i> | PVETKDIKTITLSYTIFYPREPSKPPIAAVKDGETENRL | 199 |
| <i>Sinorhizobium</i> | PVETQGIKTLITLSYTIFYPREPSKPVAQVK-AKAENKL | 198 |
| <i>Brucella</i> | DPLKDVKTITLSYTFFPIDKPRPVN AKAVGSTRNGG | 201 |
| <i>Rhodospaerooides</i> | DRDAKYTQEITLSYTIFYEIDLPEEVAQAALEQPADTINN | 202 |
| <i>Ehrlichia</i> | DPETADVKLITLSYTFFKYKE | 174 |
| <i>Anaplasma</i> | DPQTEDIREITLSYTFFPTEQESSPAAAATPQNT PQANATGAQKHTDGAA | 207 |
| <i>Wolbachia</i> | DSNTKDLGEITLSYTFFKIK | 178 |
| <i>Rickettsia</i> | DPEMENIKVITLSYSFFKIREL | 178 |

Sequence alignment of the Cox11 protein family. Sequences representing Cox11 from a diversity of organisms were aligned using ClustalW version 1.81 with default parameters (<http://www.ebi.ac.uk/clustalw/>). Amino acid residues are colored to show chemically-similar residues (pink, basic; blue, acidic; green, polar; red, non-polar), with the predicted transmembrane segment shaded grey. Across the alignment, asterisks denote absolutely conserved residues; colons represent residues highly-conserved across species. The positively-charged residues (+) and putative stop-transfer enhancement signal (underlined) in the predicted mitochondrial targeting sequence of *ScCox11* are shown.

Footnotes:

1 – While it lacks the N-terminal extension seen in the nuclear-encoded Cox11 from *Saccharomyces*, the N-terminus of *RaCox11* could form a short basic, amphipathic helix (with residues K₃, R₅ and K₆) and is followed by a helical transmembrane segment (grey shaded) of amphipathic character (S₇, S₁₆, T₁₈, S₂₃ and S₂₆) that could provide for the observed “minimalist importability”. Note that the N-terminal sequence of *RaCox11* does not conform to sequence motifs defined for the outer membrane proteins Tom20 and Tom70 (Chan et al 2006; Hulett et al 2008).

References:

- Chan NC, Likić VA, Waller RF, Mulhern TD and Lithgow T. 2006. The C-terminal TPR domain of Tom70 defines a family of mitochondrial protein import receptors found only in animals and fungi. *J Mol Biol.* **358**:1010-22
Hulett JM, Lueder F, Chan NC, Perry AJ, Wolynec P, Likić VA, Gooley PR and Lithgow T. 2008. The transmembrane segment of Tom20 is recognized by Mim1 for docking to the mitochondrial TOM complex. *J Mol Biol* **376**:694-704

Table S1. Membrane proteins encoded on the mitochondrial genome of *Reclinomonas americana*.

| Gene | Protein function | Hydrophobicity ¹ | | Signal Peptide ² | TMs ³ |
|------------------|---|-----------------------------|-------|-----------------------------|------------------|
| | | H17 | mesoH | | |
| <i>atp1</i> | F ₁ F ₀ -ATP synthase, peripheral membrane protein | 0.8 | -0.1 | N | - |
| <i>atp3</i> | F ₁ F ₀ -ATP synthase, integral membrane protein | 0.7 | -0.9 | N | - |
| <i>atp4</i> | F ₁ F ₀ -ATP synthase, integral membrane protein | 2.5 | -0.1 | Y | 2 |
| <i>atp6</i> | F ₁ F ₀ -ATP synthase, integral membrane protein | 2.5 | 1.5 | N | - |
| <i>atp8</i> | F ₁ F ₀ -ATP synthase, integral membrane protein | 2.2 | -0.5 | N | 1 |
| <i>atp9</i> | F ₁ F ₀ -ATP synthase, integral membrane protein | 2.8 | 0.2 | Y | 1-2 |
| <i>cob</i> | Cytochrome c reductase (Complex III) | 2.8 | 1.2 | N | 8 |
| <i>cox1</i> | Cytochrome c oxidase (Complex IV) | 2.5 | 1.3 | N | 5 |
| <i>cox2</i> | Cytochrome c oxidase (Complex IV) | 2.8 | 0.8 | N | 3 |
| <i>cox3</i> | Cytochrome c oxidase (Complex IV) | 2.4 | 1.0 | N | 6 |
| <i>cox11</i> | Cytochrome oxidase assembly factor | 2.2 | -0.5 | Y | 1 |
| <i>nad1</i> | NADH dehydrogenase (Complex I) | 2.6 | 1.4 | N | 7-8 |
| <i>nad2</i> | NADH dehydrogenase (Complex I) | 2.5 | 1.2 | N | 11 |
| <i>nad3</i> | NADH dehydrogenase (Complex I) | 2.7 | 0.6 | Y | 3 |
| <i>nad4</i> | NADH dehydrogenase (Complex I) | 2.5 | 1.3 | N | 6 |
| <i>nad4L</i> | NADH dehydrogenase (Complex I) | 2.3 | 1.0 | N | 3-4 |
| <i>nad5</i> | NADH dehydrogenase (Complex I) | 2.6 | 1.1 | N | 9 |
| <i>nad6</i> | NADH dehydrogenase (Complex I) | 2.6 | 1.8 | Y | 5 |
| <i>nad7</i> | NADH dehydrogenase (Complex I) | 1.1 | -0.5 | N | - |
| <i>nad8</i> | NADH dehydrogenase (Complex I) | 0.6 | -1.4 | N | - |
| <i>nad9</i> | NADH dehydrogenase (Complex I) | 1.5 | -1.1 | N | - |
| <i>nad10</i> | NADH dehydrogenase (Complex I) | 1.5 | -0.6 | N | 1 |
| <i>nad11</i> | NADH dehydrogenase (Complex I) | 0.8 | -0.6 | N | - |
| <i>orf169</i> | Putative protein of unknown function | 2.2 | 0.75 | N | 1-2 |
| <i>orf717</i> | Putative protein of unknown function | 1.6 | 0.5 | N | 3 |
| <i>sdh2</i> | Succinate dehydrogenase (Complex II) | 1.1 | -1.2 | N | - |
| <i>sdh3</i> | Succinate dehydrogenase (Complex II) | 2.7 | 1.3 | N | 3 |
| <i>sdh4</i> | Succinate dehydrogenase (Complex II) | 1.7 | 0.4 | N | 3 |
| <i>secy</i> | SecY subunit of translocon. | 2.2 | 0.9 | N | 10 |
| <i>tatA</i> | TatA subunit of TAT protein transporter. | 2.4 | -4.4 | N | 1 |
| <i>tatC</i> | TatC subunit of TAT protein transporter. | 2.1 | 1.1 | N | 6 |
| <i>ccmF/yejR</i> | CcmF subunit of Ccm complex. Integral membrane protein ⁴ . | 3 | 1.2 | N | 12 |
| <i>ccmC/yejU</i> | CcmC subunit of Ccm complex. Integral membrane protein ⁴ . | 2.5 | 1.3 | N | 5-6 |
| <i>ccmB/yejV</i> | CcmB subunit of Ccm complex. Integral membrane protein channel subunit ⁴ . | 2.5 | 1.6 | N | 6 |
| <i>ccmA/</i> | CcmA subunit of Ccm complex. Peripheral | 0.5 | -0.6 | N | - |

yejW membrane protein, ATP-binding subunit⁴.

Footnotes:

- 1 - Hydropathy scores were predicted according to Claros *et al* (1995).
- 2 - SignalP3.0 (Bendtsen *et al* 2004) suggests several proteins have signal sequences that would direct them to the Sec translocon for integration into the mitochondrial inner membrane.
- 3 - Prediction of membrane proteins made use of (i) DAS (Cserzo *et al*. 1997) to detect putative transmembrane domains – TMs - the number predicted is shown, and (ii) functional annotations for peripheral membrane proteins.
- 4 - The mitochondrial genome of *R. americana* encodes the CcmA, CcmB, CcmC and CcmF subunits of a Type II cytochrome *c* assembly pathway. This system is used to present and conjugate heme to apocytochrome *c* in the intermembrane space and is found in various protists as well as plants (Allen *et al*. 2008; Hamel *et al*. 2009). The multiple transmembrane segments predicted in CcmB (6 transmembrane segments), CcmC (6 transmembrane segments) and CcmF (11-12 transmembrane segments) make them putative substrates for the Sec translocon and both CcmC and CcmF predict strongly for a signal-anchor targeting sequence that would initiate their assembly by the Sec translocon.

References:

- Allen JW, Jackson AP, Rigden DJ, Willis AC, Ferguson SJ, Ginger ML. 2008. Order within a mosaic distribution of mitochondrial c-type cytochrome biogenesis systems? FEBS J. **275**:2385-2402
- Bendtsen JD, Nielsen H, von Heijne G and Brunak S. 2004. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol. **340**:783-795.
- Claros MG, Perea J, Shu Y, Samatey FA, Popot JL and Jacq C. 1995. Limitations to *in vivo* import of hydrophobic proteins into yeast mitochondria. The case of a cytoplasmically synthesized apocytochrome b. Eur J Biochem **228**:762-771.
- Cserzö M, Wallin E, Simon I, von Heijne G and Elofsson A. 1997. Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng. **10**:673-676.
- Hamel P, Corvest V, Giegé P, Bonnard G. 2009. Biochemical requirements for the maturation of mitochondrial c-type cytochromes. Biochim Biophys Acta **1793**:125-138.

Table S2. Matrix proteins encoded on the mitochondrial genome of *Reclinomonas americana*.

| Gene name | Protein function | Hydrophobicity¹ | |
|------------------|---|-----------------------------------|--------------|
| | | H17 | mesoH |
| <i>rp11</i> | 50S ribosomal subunit L1 | 0.4 | -1.0 |
| <i>rp12</i> | 50S ribosomal subunit L2 | 0.2 | -0.8 |
| <i>rp15</i> | 50S ribosomal subunit L5 | 0.6 | -0.7 |
| <i>rp16</i> | 50S ribosomal subunit L6 | 0.3 | -0.8 |
| <i>rp110</i> | 50S ribosomal subunit L10 | 0.8 | -1.0 |
| <i>rp111</i> | 50S ribosomal subunit L11 | 0.4 | -1.0 |
| <i>rp114</i> | 50S ribosomal subunit L14 | -0.1 | -1.8 |
| <i>rp116</i> | 50S ribosomal subunit L16 | 0.7 | -1.1 |
| <i>rp118</i> | 50S ribosomal subunit L18 | -0.1 | -1.6 |
| <i>rp119</i> | 50S ribosomal subunit L19 | 0.1 | -1.5 |
| <i>rp120</i> | 50S ribosomal subunit L20 | 0.9 | -1.1 |
| <i>rp127</i> | 50S ribosomal subunit L27 | 0.6 | -1.7 |
| <i>rp131</i> | 50S ribosomal subunit L31 | 0.9 | -2.6 |
| <i>rp132</i> | 50S ribosomal subunit L32 | -1 | -7.1 |
| <i>rp134</i> | 50S ribosomal subunit L34 | -1.8 | -3.3 |
| <i>rps1</i> | 30S ribosomal subunit S1 | 1.6 | -0.8 |
| <i>rps2</i> | 30S ribosomal subunit S2 | 0.9 | -0.3 |
| <i>rps3</i> | 30S ribosomal subunit S3 | 0.3 | -0.9 |
| <i>rps4</i> | 30S ribosomal subunit S4 | 0.3 | -1.1 |
| <i>rps7</i> | 30S ribosomal subunit S7 | 0 | -1.4 |
| <i>rps8</i> | 30S ribosomal subunit S8 | 1.1 | -1.1 |
| <i>rps10</i> | 30S ribosomal subunit S10 | 0.3 | -1.7 |
| <i>rps11</i> | 30S ribosomal subunit S11 | 0.8 | -0.7 |
| <i>rps12</i> | 30S ribosomal subunit S12 | -0.4 | -1.7 |
| <i>rps13</i> | 30S ribosomal subunit S13 | -0.4 | -1.5 |
| <i>rps14</i> | 30S ribosomal subunit S14 | -1.1 | -2.6 |
| <i>rps19</i> | 30S ribosomal subunit S19 | 0.8 | -1.5 |
| <i>rpoA</i> | RNA polymerase subunit RNA polymerase subunit alpha | 1.1 | -0.5 |
| <i>rpoB</i> | RNA polymerase subunit RNA polymerase subunit beta | 1.5 | -0.4 |
| <i>rroc</i> | RNA polymerase subunit RNA polymerase subunit beta' | 1.6 | 0.4 |
| <i>rpod</i> | RNA polymerase, transcription initiation factor sigma | 0.8 | -0.8 |
| <i>tufA</i> | Elongation factor Tu | 1.1 | -0.7 |

Footnotes:

1 - Hydropathy scores were predicted according to Claros *et al* (1995).

References:

Claros MG, Perea J, Shu Y, Samatey FA, Popot JL and Jacq C. 1995. Limitations to *in vivo* import of hydrophobic proteins into yeast mitochondria. The case of a cytoplasmically synthesized apocytochrome b. Eur J Biochem **228**:762-771.

Table S3. EST analysis.

| Cluster ID | No. of ESTs | AutoFACT Annotation¹ | Rapid Annotation Name¹ | Rapid Annotation Score |
|-------------------|--------------------|--|--|-------------------------------|
| RAL00003938 | 3 | Mitochondrial import inner membrane translocase subunit TIM10 related cluster | Mitochondrial import inner membrane translocase subunit TIM10 related cluster | 1.00e-08 |
| RAL00003758 | 2 | Mitochondrial intermembrane space translocase subunit Tim9, putative related cluster | Translocase of inner membrane TIM9 related cluster | 4.00e-05 |
| RAL00006751 | 1 | Mitochondrial intermembrane space translocase subunit Tim9, putative related cluster | Translocase of inner membrane TIM9 related cluster | 4.00e-05 |
| RAL00003642 | 3 | Probable mitochondrial processing peptidase alpha subunit, mitochondrial precursor related cluster | Mitochondrial processing peptidase alpha subunit, mitochondrial precursor related cluster | 2.00e-41 |
| RAL00006561 | 2 | Putative mitochondrial processing peptidase related cluster | F1C9.12 mitochondrial processing peptidase beta subunit, putative | 4.00e-94 |
| RAL00006644 | 1 | Tim17 domain containing protein | | 2.00e-18 |
| RAL00003686 | 1 | Similar to sp Q12328 Saccharomyces cerevisiae YDL217c TIM22 related cluster | T22K18.6 mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23 family protein | 2.00e-16 |
| RAL00001769 | 1 | T24P15.12 mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23 family protein | T24P15.12 mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23 family protein | 2.00e-06 |
| RAL00001943 | 1 | Inner mitochondrial membrane peptidase 2 related cluster | Inner mitochondrial membrane peptidase 2 related cluster | 5.00e-19 |

| | | | | |
|-------------|---|---|---|----------|
| RAL00004675 | 1 | Inner mitochondrial membrane peptidase 2 related cluster | Inner mitochondrial membrane peptidase 2 related cluster | 1.00e-20 |
|-------------|---|---|---|----------|

Footnotes:

1 - Annotations in the Taxonomically Broad expressed sequence tag Database (TBESTDB; <http://tbestdb.bcm.umontreal.ca/>) were used to identify candidate TOMs and TIMs. The candidate sequences were extracted and analyzed for conserved sequence features using hidden Markov models previously constructed for each of the TOM and TIM components (Dolezal et al. 2006; Likic et al. 2010).

References:

- Dolezal P, Likic V, Tachezy J and Lithgow T. 2006. Evolution of the molecular machines for protein import into mitochondria. *Science* **313**:314-318.
- Likic VA, Dolezal P, Celik N, Dagley M and Lithgow T. 2010. Using hidden markov models to discover new protein transport machines. *Methods Mol Biol.* **619**: 271-284.