

Supplementary Materials for:

Massive and widespread organelle genomic expansion in the green algal genus *Dunaliella*

Michael Del Vasto¹, Francisco Figueroa-Martinez², Jonathan Featherston^{3,4}, Mariela A. Gonzalez⁵,
Adrian Reyes-Prieto², Pierre M. Durand^{3,6}, David Roy Smith^{1,Φ}

¹Department of Biology, University of Western Ontario, London, ON, N6A 5B7, Canada.

²Department of Biology, Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program, University of New Brunswick, Fredericton, NB, E3B 5A3, Canada.

³Department of Molecular Medicine and Sydney Brenner Institute for Molecular Biosciences, University of the Witwatersrand, Johannesburg 2193, South Africa.

⁴Agricultural Research Council, Biotechnology Platform, Pretoria 0040, South Africa.

⁵Departamento de Botánica, Facultad de Ciencias Naturales y Oceanográficas. Universidad de Concepción, Casilla 160-C, Concepción, Chile.

⁶Department of Biodiversity and Conservation Biology, Faculty of Natural Sciences, University of the Western Cape, Private Bag X17, Belville, Cape Town, 7530, South Africa.

^ΦCorresponding author: David Roy Smith, Department of Biology, University of Western Ontario, London, Ontario, N6A 5B7, Canada. Email: dsmit242@uwo.ca. Tel: (519) 661 2111, ext.86482.

Table of Contents

| | |
|--------------------------------------------------------------------------------------------------------------------------------------------|----|
| Supplementary Table S1. <i>Dunaliella salina</i> nucleotide substitution rates for individual protein-coding loci..... | 2 |
| Part A. Mitochondrial DNA-located genes | 2 |
| Part B. Plastid DNA-located genes | 2 |
| Part C. <i>Dunaliella salina</i> CONC-001 protein-coding ptDNA genes. | 3 |
| Supplementary Table S2. Synonymous substitution rates in plastid and mitochondrial genomes of various plastid-bearing lineages..... | 11 |
| Supplementary Figure S1. Maximum-likelihood protein phylogeny of chlamydomonadalean algae..... | 12 |
| References | 13 |

Supplementary Table S1. *Dunaliella salina* nucleotide substitution rates for individual protein-coding loci.

Part A. Mitochondrial DNA-located genes

| MtDNA locus | # codons | d_N / d_S | d_N | d_S |
|---------------------------|----------|-------------|--------|--------|
| <i>cob</i> | 389 | 0.0445 | 0.0319 | 0.716 |
| <i>cox1</i> | 509 | 0.0311 | 0.027 | 0.8686 |
| <i>nad1</i> | 295 | 0.008 | 0.0179 | 2.2332 |
| <i>nad2</i> | 434 | 0.0436 | 0.0439 | 1.0067 |
| <i>nad4</i> | 482 | 0.0505 | 0.0704 | 1.3961 |
| <i>nad5</i> | 557 | 0.0444 | 0.0415 | 0.9355 |
| <i>nad6</i> | 156 | 0.0707 | 0.0703 | 0.9939 |
| Average | 403.14 | 0.042 | 0.043 | 1.164 |
| Standard deviation | 138.46 | 0.019 | 0.020 | 0.515 |

Part B. Plastid DNA-located genes

| PtDNA locus | # codons | d_N / d_S | d_N | d_S |
|-------------|----------|-------------|--------|--------|
| <i>atpA</i> | 503 | 0.001 | 0 | 0.0139 |
| <i>atpB</i> | 479 | 0.0447 | 0.0113 | 0.2531 |
| <i>atpE</i> | 134 | --- | 0.0029 | 0 |
| <i>atpF</i> | 173 | 0.001 | 0 | 0.0166 |
| <i>atpH</i> | 81 | 0.001 | 0.0001 | 0.085 |
| <i>atpI</i> | 238 | 0.0255 | 0.0143 | 0.5616 |
| <i>ccsA</i> | 396 | 0.1162 | 0.0042 | 0.0362 |
| <i>cemA</i> | 593 | 0.2312 | 0.0054 | 0.0234 |
| <i>chlL</i> | 290 | 0.025 | 0.0568 | 2.2727 |
| <i>petA</i> | 287 | 0.0617 | 0.0027 | 0.0444 |
| <i>petB</i> | 214 | 0.001 | 0 | 0.0424 |
| <i>petD</i> | 159 | 0.001 | 0 | 0.024 |
| <i>petG</i> | 36 | 0.001 | 0 | 0 |
| <i>petL</i> | 31 | 0.001 | 0 | 0.0404 |
| <i>psaA</i> | 711 | 0.0121 | 0.0006 | 0.0472 |
| <i>psaB</i> | 734 | 0.0436 | 0.0006 | 0.013 |
| <i>psaC</i> | 80 | 0.001 | 0 | 0.0349 |
| <i>psaJ</i> | 40 | 0.4914 | 0 | 0 |
| <i>psbA</i> | 352 | 0.0331 | 0.0038 | 0.1157 |
| <i>psbB</i> | 507 | 0.001 | 0 | 0.0083 |
| <i>psbC</i> | 461 | 0.0213 | 0.0018 | 0.0848 |
| <i>psbD</i> | 351 | 0.001 | 0 | 0.0276 |
| <i>psbE</i> | 81 | 0.001 | 0 | 0.0356 |

| | | | | |
|---------------------------|---------|--------|--------|--------|
| <i>psbF</i> | 43 | 0.001 | 0 | 0.0408 |
| <i>psbH</i> | 82 | 0.001 | 0 | 0 |
| <i>psbI</i> | 36 | 0.001 | 0 | 0 |
| <i>psbJ</i> | 41 | 0.001 | 0.0002 | 0.1716 |
| <i>psbK</i> | 45 | 0.0507 | 0 | 0 |
| <i>psbL</i> | 37 | 0.9961 | 0.1247 | 0.1251 |
| <i>psbM</i> | 36 | 0.001 | 0 | 0 |
| <i>psbN</i> | 43 | 0.4303 | 0 | 0 |
| <i>psbT</i> | 30 | --- | 0.0138 | 0 |
| <i>psbZ</i> | 61 | --- | 0 | 0 |
| <i>rpl2</i> | 274 | 0.053 | 0.0014 | 0.027 |
| <i>rpl5</i> | 180 | 0.001 | 0 | 0.0123 |
| <i>rpl14</i> | 118 | 0.001 | 0 | 0.0441 |
| <i>rpl16</i> | 136 | 0.41 | 0 | 0 |
| <i>rpl20</i> | 111 | 0.0627 | 0.0035 | 0.056 |
| <i>rpl23</i> | 115 | 0.0726 | 0.0034 | 0.047 |
| <i>rpl36</i> | 36 | 0.0943 | 0 | 0 |
| <i>rps4</i> | 251 | --- | 0.0015 | 0 |
| <i>rps8</i> | 143 | 0.3676 | 0 | 0 |
| <i>rps9</i> | 195 | 0.001 | 0 | 0.0159 |
| <i>rps11</i> | 129 | 0.1006 | 0 | 0 |
| <i>rps12</i> | 128 | 0.001 | 0 | 0.0427 |
| <i>rps18</i> | 139 | 0.001 | 0 | 0 |
| <i>rps19</i> | 91 | 0.0351 | 0.0045 | 0.1284 |
| <i>tufA</i> | 417 | 0.001 | 0 | 0.013 |
| <i>ycf3</i> | 170 | 0.001 | 0 | 0.0178 |
| <i>ycf4</i> | 219 | 0.144 | 0.0038 | 0.0266 |
| <i>ycf12</i> | 32 | 0.001 | 0.0001 | 0.0551 |
| Average | 201.353 | 7.842 | 0.005 | 0.090 |
| Standard deviation | 182.918 | 26.859 | 0.019 | 0.324 |

Part C. *Dunaliella salina* CONC-001 protein-coding ptDNA genes.

Locus

atpA

Sequence

ATGGCAATGCGCACACCAGAAGAATTAAGTAATCTAATTAAGATTTAATTGAAGAATATACACCAGAAGTAA
AAATGGTGGATTTTGGTATTGTTTTCCAAGTTGGAGATGGTATTGCTCGTATCTATGGATTAGAACGAGTAATGT
CTGGTGAATTATTAGAATTTGAAGATGGTACATTAGGTATTGCTCTAAACCTAGAAGCAAATAATGTTGGTGGT
GTATTATTAGGTGATGGTTTAAAAATTACAGAAGGTAGTCGTGTACGTTGTACAGGTCGTATTGCTGAAATTCC
AGTTGGTGAAGCTTATTAGGACGTGTAGTTAACTCATTAGCACAAACCAGTAGATGGTAAAGGACCTATTAACA
CAACTGATTCACGAGCAATTGAAGCAGCTGCTCCAGGTATTGTTTCACGTCGTTACAGTATATGAACCATTACAA
ACAGGATTAGTTTCTGTAGATTCAATGATTCCAGTAGGACGTGGTCAACGTGAGCTTATTATTGGTGACCGTCA
AACAGGTAACAGCAATTGGTGTGATACTATTCTAAACCAAAAAAGGTAAGGTGTTATTTGTGTTTATGTAG
CAATTGGACAAAAAGCTTCATCAGTTGCACAAGTATTAACCTACATTAATAAGAGCGTGGTCTTTAGACTATACA
ATTATAGTAATGTCAACTGCTAACGAACCAGCTACTTTACAATATTTAGCTCCTTACACTGGTGCTACTTTAGCA
GAATATTTTCATGTACACGGGTCGTCCAACCTTTATGTGTTTACGATGATTTATCAAAACAAGCAGCAGCCTACCGT

GAAATGTCTTTATTATTACGTCGTCCACCAGGTCGTGAAGCTTACCCAGGTGACGTTTTCTATCTTCACTCTCGT
 CTTTTAGAACGTGCTGCTAAATTAATACAGCATTAGGTGAAGGTAGTATGACAGCATTACCTATTGTTGAAAC
 ACAAGAAGGAGACGTATCAGCTTATATTCCAACAAACGTTATTTCTATTACTGATGGTCAAATCTTCTTATCTTC
 TGGTTTTATTTAACTCTGGATTACGTCCAGCTATTAACGTAGGTATTTCTGTATCACGTGTAGGTTTCATCAGCTCA
 ATTTAAAGCAATGAAACAAGTAGCAGGTTCTTTAAAAATAGCGTTAGCACAATTCGGTGAATTAGAAGCTTTCT
 CACAATTTGCTTCTGATTTAGATCAAGCTACACAAAACCAATTAGCTCGTGGTTACCGTCTTTCGTGAAATTTTAA
 AACAAACCACAAAGTTCACCTTATTCATTATCTGAAATGGTTATTAGTATCTATACTGGTACAAAAGGTTACTTAG
 ATACTTTACCAGTTGAAAAAGTAACATCATTCTAACAGGTCTTCGTAACCTATATTGCTTCAAATCATAAACAAT
 TCGAAACAATTATTAATGATACAAAAGTATTTACTCCTGAAGCTGAAGGTATTTTACAAAAGCTATTCCAGAA
 TATCTTACAGAAATTAAGCAAGTTAA

atpB

ATGACTAATGAAACAGCAAATATTGGACGTGTAGTTCAAGTTATCGGGCCAGTACTTGATATCGTTTTTGCAA
 AGGTCAAGTACCTAACATTTACAATGCTTTAACTATTAAGAAGTAAAAATGCCGACGGGTTAGAATTAGCTGTAA
 CTGTTGAAGTTCAACAATTATTAGGTGATAATGCTGTTTCGTGCAGTATCAATGTTACCTACAGAAGGTTAATGC
 GTGGTATGGAGGTACTTGATACACAAAAACCATTAAGTGTCCAGTTGGTCTGATACGTTAGGTCGTATCTTTA
 ACGTATTAGGTGAGCCGGTTGACAATTTAGGTCCTGTAATGCTGGTCAATCAAATCCAATTCACCGTCAAGCT
 CCTACTTTTCGTAGACTTAGATACTCGTTTATCTATTTTTGAAACAGGTATTAAGTAGTTGACCTTTTAGCACCT
 ACCGTCGTGGGGTAAAATTGGTTTTATTGGTGGTGTGGTGTAGGTAAAACAGTACTAATTATGGAACCTTATT
 AACAACTTGTCAAAGCACACGGTGGTGTATCCGTTTTCGCTGGTGTAGGTGAAAGAACACGTGAAGGTAATG
 ACCTTTATGGTGAATGAAAGAATCTGGTGAATTTCAAGAAAATAACTACCAGAATCTAAAGTAGCCTTAGTT
 TACGGTCAAATGAATGAACCACCTGGAGCAGTATGCGAGTAGGTTAACAGCTTTAACAAATGGCAGAATATT
 CCGTGACGTAACAACAAGACGTTTTATTCTTCAATGATAACATTTTCCGTTTTGTCAAGCTGGTGTGAAGT
 TTCTGCCTTATTAGGTCGTATGCCATCAGCGGTAGGTTACCAACCTACTCTTGCTACAGAAATGGGTACATTACA
 AGAACGTATTACAAGTACAAAAGATGGTAGTATTACTAGTATTCAAGCTGTTTATGTACCTGCTGATGACTTAA
 CAGATCCTGCTAGTGCTACAACCTTCACTACTTAGATGCTTGTACTGTACTTTCAGGAGGTTAGCATCAAAG
 GTATCTATCCAGCAGTAGATCCTTTAGATTCTACATCAACAATGTTACAACCATGGATCGTAGGGGAGAAACAT
 TATTCAGCAGCTTCGAATGTTAAAGAAACATTACAACGTTATAAAGAGCTTACTGACATTATCGCGATTTTAGG
 CCTAGATGAATTAAGTGAAGAAGATCGTCAAACAGTTGCTCGTGCACGTCGTATTGAACGATTCTTATCTCAAC
 CTTTCTTCGTAGCAGAAAGTTTCACTGGAGCACCAGGTAATACGTAACCTTTATCAGAATCTATCGATGATTTCA
 ACCGCATTATCAGCGGTGAACCTTGATAATTTACCTGAAGCAGCTTTCTATTTAGTTGGTAAACATGAACGATGTTT
 TAGCCAAAGCAGAAACACTTCAAGGCTAA

atpE

ATGAATTTACAAATTTCAATCTAACACCAGAACGTCCTTTTTGGAATGGTCAAGCTGACGAAATGTTTTACCA
 ACTGAAACAGGTGAGATGGGTGTTTTAAAAAATCATGCTCCTATTATTACTGGTTTAGATGTAGGTGCAATGCTT
 ATTCGTACAAATCAAGAATGGACTTCTTTAGCGATTATGGGTGGTTTTGCATTAGTTAAACAAAATGTTGTAAC
 ATCTTAGCAAATGAAGCACAATCTTCTGAAAACATTGACTCCTGAAGAAGCAAAAAATGATTTTGAACAGCAA
 AAAATAATGAAAAAGCTGAAGGGTTAAAGATAAAGTAGAAGCTAATTCGCTTTTAAACGTTCTAAAGC
 ACGTTTTCAAGTAGCAGTAAAATCAAAAAACTTTTTAA

atpF

ATGCTTCTTTAGGTCATGGTGGTTTTGGAATTAATACAAATATTTTTGAACTAATCTTATAAAATTTAGCTGCT
 GTTATTGGAATTGTTGTTTCATTTGTAGGTAATAATTTAAATGCTTTATTAGAAGATCGTAGAAAAACAATTCCT
 TCTAATTTAGAAGAAGCTAATCAAAGAGCAATTGAAGCACAAGAAAAATTACAAAATGCCAAAACCTCAGTTAG
 AATTATCTAAGAAAAAAGCACAAGAAATCCGTGACGAAGGTGTCTCTCGTGCAAGTAATGAAATTCAAACAGT
 AGTAAGTCAACATGAAATTCGATTAGCTAGTTTAGAAGATTTAAAAATGAAACAGTTCAATTTTATCAACAAA
 AAGCATTCAAAGAAGCTTATGTTTATGTAGTATCTCGTATTATGAGTGTGACGTGACGTGACGTGACGTTAA
 TAGATTCTACATATCATGTTATTGTTAAACAATTTCTATGTTTTCAAAATTTACAGAATTTAAACCAAAATCTATCTC
 TTAA

atpH

ATGAACCCTATTGTTGCTGCTGCTTCTGTTGTGTCTGCTGGTTTTAGCTGTTGGTTTTAGCTGCTATTGGTCTGGTA
 TGGGTCAAAGGTACAGCTGCAGGTTATGCTGTTGAAGGTATTGCACGTCAACCTGAAGCTGAAGGTAAAATCCGT
 GGTGCTTTATTATTAAGTTTCGCGTTTATGGAATCTTTAACTATTTATGGTCTAGTTGTAGCTCTTGCTCTCTTTT
 CGCTAACCCTTTTGCTGGTTAA

atpI

ATGATTAATCCTTTATTAGAAATTGCTGAAGTTTCGGTAGGACAACATTATTATTGGAATTAGGAAATTTTGAA
 TTACATGGTCAAGTTTTAATTACTACTTGGGTAGTAATTGCTATTCTTCTAGCTTTATCTTTATTAGGTAACAGTA
 ATTTAAAACCAACTCCAGACGGTTTACAAAATTTTACTGAATATATTACTGAATATATTCGTGATTTAGCAAAA
 ACACAAATTTGGTGAAGAAGATTATGTAATAATGGGTTCCTTTTATTGGTACATTATTCTTATTTATCTTTGTGTCTA
 ACTGGTCTGGTGCATTATTTCTTATCATATTATTGAAATTCCTAATGGTGAACCTGCTGCTCCTACGAATGACA
 TTAATACTACGGTAGCTTTAGCTTTATTAACATCAATTTTCATATTTCTATGCGGGTATTTCTAAAAAAGGACTTG
 GCTATTTCAAACGTTATATTTACCTGCAGCATTTTTATTACCAATTAACGTAAGTACTAGAGGATTTTCAGTAAACCCG
 TTTCAATTAAGCTTTTCGCTTTTTGGAATATCTTGGCCGATAAGCTAGTTGTGCGAGTTCTAGTATCTCTGTACC
 CCTAGTTATTTCAATCCCTGTTATGCTTCTTGGAGTTTTACAAAGTGTGTACAAGCTCTTGATTTGCGACTTTA
 GCTGGAGCCTACATTTGGCGAAAGCTTAGAAGATCACCACTAA

ccsA

ATGAATGCGCAACTTGCAAATCTTTCTTTTTAGTTTTATTATGTTTCGATGATTTGTTATTGGATTGGGCCTTTAA
 GTCGTGACCTATTTGGAATTTCTTCAAACCTATCGCCGCTCGGCGATCCGAGCGGACGAGACTTTAAAG
 TCCGCTCGTAGTACGGTAATTTCTACTGGACTAAATCAATCTGTAATACCGTCTGAATTAGTAGTAGACGGGCTT
 GGTACAAAAGTAGTAATAATTCCTCTTTGTTTTCTTCAGAAACAAACGTTTCTACCTTACAAACAAATGCTGGT

TTTTTAGGAGGAGTACTAAATAATATTAGATCACAAATTGGATATATAAAAACCTTCTTTTAATATTAGCAATAGT
 TTACAACATAATTGGACGTTTCAGGAATTATCTTATCAAATGTTTTATTAATGGTTTTATTGTGTAACCGTTGGAAA
 GAATCTGGACATTTTCCATTAAGTAATCTATATGAATCTTTAATGTCTTATCTTGGAGTTTTACTGTTATTCATT
 TATTTCTTGAAAACCTTACTAACTCACAAACAGCATCAAACCAAAGTGTAGGTAATTTATAGGTGCTATAACT
 ACACCTAGTGTCTTATTACAAAATGCTTTTGTACATTCACCTTACCTAAAAGAAATGCAAGGTTCTTCATCGTA
 GTACCTGCTTTACAATCTAATTGGTTAATGATGCATGTTACAGTAATGATTATAAGTTATGCAGCTTTAATTTTA
 GGTTCACTACTATCTATTGCTTACTTAATTATTACCATAAAAACGTGAATCAAATTTTGATACAAAATTTAGATTCA
 AATAATATAACAAATAAGGACACAGAAGAACTAATTCTGAATTCGTTAGTGCTAAAGGTACTTGTATTTC
 TAATAAAAAAATCTGCGTCAGTCTATAGCTTTAACCTTAGACAACCTTAAGTTATAGAATTTTAGGTATAGGTTT
 CCCTTTCTTAACTATTGGAATTTTGTCTGGTGTGTATGGGCTAACGAAGCATGGGGCTCTTATTGGAGTTGGGA
 TCCCAAAGAACTTGGGCTTTTTAACGTGGTTAGTTACGCAATTTATTACATACTCGACTAAATAAAGGTTG
 GGAAGGTCAAACCAGCTATTATTGCCTCTTTAGGTTTTATTGTTACTTGGATTTGCTATTTAGGAGTAAACTT
 AATAGCGAAAGGTTACATAGTATGGTTGGATTATAATAA

cemA

ATGATACAAAAAACAACCTTTTTATCGTTTTTAAATCTAATTCGGACCCGCTAAAAATTTTTTATCATT
 TCTAAAGATAATGAACTACAAGCGGGCCCGCAAATTAGGTTTAAACAATGATTATGATAAAAAATCAAAAA
 GATTTTCACGTTTGACGTCTGTAAAGGCGGCACCTAAACAAGAAGTAGTTTCGATTACTTATGAAGAAATAGGA
 TTATTTCTCGTTCATTTAGTCGTGTTTTAGATCGATTTTTAAACAAATTTTTTCAGATGTAGAAAATTTAGTAA
 TACAAGAATATAGATTCTATCGTTATTTATTTTTAAACAACCTTAAAAATGCTTTTTAATTTTTTTTTGTTCTTTT
 CTCGTGAATTTTTAGCAAAATCTTATGTTGTACGACCAATTAATGAAATTTTTTGGAAATCAAAAACAACTGAA
 ATTTTTTAAATTCATATCAACAAAAACATGCATTTGCCGAATTAAGATTTTGAAGAAAATTAATTTTTGAG
 TCATTAATTTATCCAAAATCGCTTAGTCAGCCGTAATTTGGAGAAGAAAATAAGTCTCTTCTTATCAATTTACA
 CCTATAAGAAATACGATGTCACCGCCTATTAAGAAAATCAAACAACAAAAGAAAATTTACTTATTAACAAAA
 TTAACGAAGTTAATTTGTAAAGTAAATTTCTAAAAATTTGGGTGGGTTCAACTTCAGATTTTGTAAATTTTTATCA
 ATTTCCAAAACCTATCGAATTTCAAAGCCATTACTTATGCCTTTTTATAGCAAAAAGTAAATAACGTTTCGTTT
 CGAGCTCGGACGAGCTCCACATGATAATTTATCTACAGTAAATCAACATGGTAAAATCGGTTATATTGATCATT
 CTCTTCATCTGTAAAGTTAATATTTAAAAAACAACAAATGAAAGCGAAGCAAAAAGAGAAGATTTACCTCA
 GAAATATCAAAATCTTCTCATTTTTCTATTCTTAATGAAAATATTTGAACTCTCCTCCCTCGGCTCGTTCCGAG
 GCCGAATCGACCCCGCCACAAAAATAAAGATTTTTTGCCATCCGGTAATAGGACAGACAGACTCAATATT
 GTTATCGAATAATGAAAATATTCTTCAAAGAGGTAATATTTCTTATTCAAACGATATGCAAAATCTCAAGTAAAA
 AAGAGGCTGAAGCTGCCCTCACGGATGGACTTCGCACATCAATAAATCCAAAATCGGTTCTTTTTGAAACTTTA
 AATATTTCAAAGAGTTGGAACCAAAAAACAACCTCACCAACTTCAAACACTCATTTTTTAGTTTATCAAATAT
 TGAAGCTGACAAAAATATCCAACAACGTTATCAAGAAAAACTATAGAATTAGCACAAACGTTATAACGATCAA
 AGTGTGAAAGCAATTACAAATTTTTTTGGTGTATGGTTTCATTAGTAACCTTATTTATTTATTGTATGTTCTTG
 AAATCAAATTAATAACTCAATCTTTTTATTAGAAGTCTTTTTGGTTTAGATGATAGTAAAAAATCTTTGTT
 GATCTTTTAGTTACAGATTTATTAGTAGGGTATCACTTTTCAAACCTTATGGAACTATTTTTGAGTTGTTT
 AACCATTAGGTTTACCAGAAAGTCAAACCTGGTATTTTTTTATTGGTAGCAACCTTACCTGTTTTATTAGACGTTT
 TATTCAAATATTTAATTTCCGTCATTTAAATCGTTCATCACCAGCTACAGTTGCTACTTATCATGCTATTATTGA
 ATAA

chlL

ATGAAATTAGCTGTTTATGGTAAAGGTGGTATAGGAAAAACAACAGCAAGTTGTAATATTTCAATTGCATTAGC
 AAAACGTGGAAAAAAGTATTACAAATTTGGTTGTATCCTAAACATGATTCAACATTTACATTAACAGGATTTT
 TAATCTCATAATTATTGATACTTTACAAAGTAAAGATTAATCATTACGAAGATGTTGGCCAGAAGATGTAATTT
 ACCAAGGATATGGTAACGTAGACTGTGTAGAAGTCTGGTGACCACCTGCAGGTGCTGGTTGGCGATTATGTA
 GTAGGAGAAACAGTAAAACTATTAAGAATTAATGCTTTCTTTGAATATGATATTATTCTGTTTGTATGTTTAA
 GGGGATGTTGTTTGGTGGTTTTGCAGCACCCTAAATTAACGATTATTGTATTATCGTTACTGATAACGGG
 TTCGATTCAGTGTGCTGCAACAGAATTGCTACGTCGGTAAGAGAAAAAGCACGAGTTCACCCATTACGTCT
 GGCAGGTCTTATTGGTAATCGTACATCTAAACGAGACCTTATAGATAAGTATGTTGAAGCTTGCCCTATGCCTGT
 TTTAGAGGTTTTACCTTTAATTGAAGATATTAGAATTTCAAGAGTAAAAGGTAAAACCTTATTTGAAATGGCTG
 AAGGAGAGCCTTCACTTTTATACATATTAGACTTTTACTTAAACATAGCTGACCAATTAATCTGAACCAGAA
 GGGTTGTTCTCGTGAATTAGGGGATCGTGCCTTATTTACTACTACTATCAGACTTTTATCTTCTCTTCAACAT
 CTAATACGTGATACTACTAATAACAAATCTGAAAGTAATGATTTAGATTTCTTTCTGGTTTTAA

petA

GCTTATCCGATTTTTGCACAACAAGGTTATGAAAACCCACGTGAAGCTACAGGTCGTATTGTTTGTGCTAACTGT
 CACTTAGCACAAAAACCTTCAGAAATTGAAGTACCTCAAGCAGTTTTACCTGATACAGTATTTGAAGCTGCTGT
 TCAAATCCATATGACAAATCAGTTAAACAAGTAAATGCTAAAGGTAAAAAGCTGATTTAAACGTAGGTTTTG
 TTTAATTTCTCCAGAAGGATTTGAATTAGCACCAGCTGATCGTATCCCAGAAGAAAATGAAGAAAAAAGTAGGA
 AACGTTTATTACCAACCTTATAGTCTTGACAAAAAATATTCTTGTAGTTGGTCTTTACCAGGAAAAACATAT
 AGTGATTTAACTGTACCTATCTTTCTCTGATCCAGCAAAAAAATAAATGTTTTCATATTTAAATACCCTATC
 TATTTAGGTGCTAACCGTGGACGTGGTCAACTTTATCCAGATGGTTCTAAATCAAACAACACTGTTTATAACACT
 CCAGCTCCTGGTAAAATTTGTTTCTATTACAACCTGCTAATGAGAAAAAAGGTGGTTATTTAGTAAATATTGAAAA
 AGCTAATGGTGAACAGTAACTGAAAAAATCTCCAGGACCTGAATTAATTGTTCAAGAAGGACAAGTATTA
 ACAGCTGATCAGCCTTAACTACAAACCCTAACGTAGGTGGTTTTGGTCAAAAAGATGTAGAAGTTGTTCTTCA
 AAACCCTGCAGTATTCAAGGATTATTAGTTTTCTTTGCTTTTTGATTATTAGCACAAAGTATTATTAGTTCTTAA
 AAAAAACAATTCGAGAAAGTTCAACTAGCAGAGATGAACCTTAA

petB ATGAGTAAAATTTACGATTGGTTCGAAGAACGTTTAGAAAATTCAAGCAATTGCTGATGATATTACAAGTAAATA
 TGTTCCGCCTCATGTAAACATTTTCTACTGTCTTGGTGGTATTACATTTACTTGTTCCTTAGTTCAAGTAGCTACT
 GGTTTTGCAATGACTTTCTATTACCGACCAACTGTAGCAGAAGCTTTCGCTTCAGTACAGTATTTAATGACTGAT
 GTAACTTTGGATGGTTAATTCGTTCAATTCACCGTTGGTCAGCTAGTATGATGGTTTTAATGATGATTTTACAC
 GTTTTCCGTGTTTACTTAAACAGGTGGTTTCAAACGACCAAGAGAGCTGACATGGTTAAGTGGTGTATTATGGCC
 GTTTGACTGTTTCATTCCGGTGAACAGGTTATTTACCTTGGGACCAAATCGGTTACTGGGCTGTTAAAATT
 GTAACAGGTGTTCCCTGAAGCTATCCCTGTTATTGGTGGTCCCTTAGTAGAATTATTACGTGGTGGTGTAGGTGTT
 GGTCATCAACATTAACACGTTTCTACAGTTTACACACATTTGTTTTACCATTATTAACAGCTGTTTTCATGTTAA
 TGCCTTCTTAATGATTAGAAAACAAGGTATTTTCAGGTCCTCTATAA

petD ATGTCTGTAACAAAAAACAGATTTAAGTGATCCAGTATTAAGGCTAAAATTAGCGAAAGGAATGGGACATA
 ACACATATGGTGAGCCAGCTTGGCCAAATGACTTACTTTACATGTTCCCTGTTGTAATTTTTGGTACATTTGCAT
 GTGTTATTGGTTAGCTGTATTAGAACCAGGTTCTATTGGAGAACCCAAACCTTTGCTACACCTCTTGAAA
 TTTTACCAGAATGGTATTTCTACCCAGTTTCCAAATTTACGTGTTGTACCTAACAAATTTAGGTGTTGTTT
 AATGGCAGCTGTACCGGCTGGTTAATGACAGTACTTTTTATTGAAAATATTAATAAATTTCCAAAACCTTACC
 GTCGTCCTATCGCTACTATTTTATTCTTAGTAGGTAAGTGTAGCAGCTGTTTGGTTAGGTATTGGTGCTACTTTCCC
 TATCGATATTTCTTAACTTTTGGATTATTCTAA

petG ATGGTTGAACCTTATTATCAGGAATTGTATTAGGATTAGTCCAGTACTATCGCTGGATTATTTGTAACAGCA
 TACTTACAATATCGTAGAGGTGATTTAGCTAGTTTCTAA

petL ATGGTAACTATTTCAAGTTATATTGTATTATTAATCGGTGCTTTAGGTTTACATTAACACTTTACTTAGGTTTTT
 TAAAAGGAATCAAGCTTATTTAA

psaA AAACCTGGACATTTTTCTCGTGCATTAGCAAAAGGTCCTAATACAACACTTGGATTGGAAATCTTCATGCTGAT
 GCTCACGACTTTGACAACCATACTTCAGATTTAGAAGAAATCTCAAGAAAAGTATTTAGTGCTCATTTCGGGCA
 ATTAGGTATTATTTAATTTGGTTAAGTGGGATGTATTTCCATGGTGTCTGTTTTTCAAACCTATGAAGTTGGTTA
 TCTGATCCAACACATATTAACCTAGTGCACAAGTAGTATGGCCAATTGTTGGTCAAGAAATTTAAACGGTGA
 CGTAGGTGGTGGATTCCAAGGTATTCAAATTACTTCTGGTTTCTCCAACCTATGGCGAGCTAGTGGTATCACTAG
 TGAATTACAATTATATCAACTGCTATTGGTGGTTAGTGTAGCAGCAGCTTGTTCCTTTGCTGGTTGGTCCAC
 TACCATAAAGCAGCACCAAAATTAGAATGGTTCCAAAATGTGAATCTATGTAAACCACCACCTAGCTGGTTTT
 ATTAGTTTACGTAGTTAGCTTGGGCAGGTCATGAAATTCATGTATCATTACCAGTTAATAAATTTAGTAGTC
 TGGGGTTGATCCTAAGGAAATACCATTACCTCATGAATTTTTATTAATCAATCTATTATTGCTGATTTATACCC
 AAGTTTCTCAAAGGATTAGCACCATTCTTACATTAATTTGGGCTGAATATAGTGATTTCTTAAACATTTAAAGG
 TGGTTTAAACCTGTAAACAGGTGGTTTATGGTTATCTGATACTGCTACCACCCTTAGCTATTGCTGTATTATT
 CTTAGTAGCTGGTACCAATACCGTACAACTGGGGTATTGGACACAGTATTAAGATATTTTAGAATCACACA
 AAGTCCATTTACTGGTAATGGTCAATGCTGGTTTATATGAAATTTAACTACTTCTTGGCATGCTCAATTAGCTA
 TTAACCTAGCTTTATTTGGTTCCGTTTCTATTATTGTAGCTCACCACATGTATGCAATGCCTCCTTATCCTTACT
 AGCAACAGACTCGGTACTCAACTTCTTTATTTACACACCATGTATGGATTGGTGGTTTCTGTGTTGTTGGTGC
 AGGTGCACACGCAGCTATTTTCATGGTTCGTGACTATGATCCAACAAACAACCTACAACAACCTATTAGACCGTG
 TAATTCGTCACCGTGATGCTATTATTTACACTTAACTGGGTATCTATTTTCTTAGGTTTCCACAGTTTCCGGATT
 ATACATTCATAATGATACAATGAGTGCTTTAGGTGCTCCACAAGATATGTTCTCTGATACAGCAATTCATTAACA
 ACCTGTTTTTGCACAATGGATCCAAAACACGCATTTTACAGCACCTCAACTAACAGCTCCTAATGCGTTAGCTGC
 TACAAGTTAACTTGGGGTGGTGTAGTGTGACGTTGGTGGTAAAGTAGCGATGATGCCTATCGCTTAGGTA
 CATCTGACTTCTTTGTTCAACCACCTCCAGCTTTTCAAACTCAGTAACTGTTTTTATTAAAAGGTGTTTT
 ATTTGCTCGTATCTCGTTAATTTCTGATAAAGCTCAATTAGTTTCCGTTTCCCTTTGATGGACAGGACG
 TGGTGGTACATGTCAGGTTTTCAGCTTGGGACCACGTTTTCTTAGGTCTTTTCTGGATGTATAACTCATTATCAATT
 GTTATTTTCCACTTTAGCTGGAAAATGCAATCTGATGTTTGGGGTACTGTAAGTATTCTGGTGTATCTCACATT
 ACTGGTGGTAACTTCGCTCAAAGTGCTAACACAATTAATGGTTGGTTACGTGATTTCTTTGGGCACAATCATCT
 CAAGTTATCCAATCATACGGTTCAGCACTATCAGCTTACGGTCTTATGTTCTTAGGTGCTCACTTCGTATGGGCT
 TTTTCATTAATGTTCTATTCTCTGGTCTGGTTACTGGCAAGAGTTAATCGAGTCAATTTGATGGGCTCACAA
 AAATTACGTGTAGCACCATCAATTCACCACGAGCTCTTAGTATTACACAAGGTCGTGCTGTTGGTGTAGCTCA
 CTACCTTTTAGTGGTATCGCTACTACATGGTCTGTTCTTAGCAAGAATTATTGCTGAGGTTAA

psaB ATGGCAACAAAGCTGTTTCCGAAATTTAGCCAAGGCTAGCACAAGATCCAAGTACTCGTCTGATTTGGTATGG
 ACTTGCTACAGCACATGATTTTGAAGTCAATGATGGTATGACTGAAGAAAATCTTTATCAAAAAGATTTTGGCTTC
 TCATTTTGGTCAATTAGCAATTTTCTCTATGGACATCTGGAAAATTTATTCATGTGGCTTGGCAAGGTAATTTT
 GAACAATGGGTAAGTATCCTATTTCATGTACGTCCTATTGCTCATGCTATTTGGGATCCACACTTTGGTCAACCA
 GCTGTTGAAGCATTACACGTTGGTGGTGCCTGCTGGTCCAGTAAATATTGCTACATCTGGTGTATACCAATGGTG
 GTATACAATTGGTTACGTTCTAACCAAGAATTATATGTTAGTTTCAAGTATTTTATAGCTTTAGTTTTCAGCAGTTTTT
 TTATTTGCTGGTTGGCTACATTTACAACCTAATTTCAACCATATTATCATGGTTAAAGATGCTGAATCAGT
 TTAACACCATTAGCTGGTTTATTTGGTGAAGTCTTTAGCATGGACAGGAAATTTAGTTTACAGTACTATT
 CCAGAATCACGTGGACAACATGTAGGCTGGGACAACCTTCTATCAGTTTTTACCACACCCTCAAGGTTTAACTCC
 GTTCTGGTCAAGTAACTGGGCTGCTTATGCACAAAACCCAGATACAGCTTCTCATGCTTTTGGTACTGCTGATGG
 TTCAGGTAAGTCAATTTAACTTTTTTAGGTGGTTTCCATCCACAACTCAAAGTTTATGGTTATCTGATATGGCT
 CATCACCATTAGCTATTGCAGTCTTTTTCATTGTTGCAGGTCACATGTACCGTACAACTTTGGTATTGGACAC
 CGATTAGAAGCTATCTTAGAAGCTCATACACCGCTGCTGGTGGTTTAGGAGCTGGTCATAAAGGATTATTTCA

CACAGTTAACAATTCTTTACATTTCCAATTAGGTTTAGCTTTAGCTTCAGTAGGTACGATTACATCATTAGTAGC
 ACAACACATGTATTACATTACCACCTTACGCTTATTTAGCAGTTGATTTTACAACACAAGCGTCTCTTTACACACA
 CCATCAATATATTGCTGGTTTTATCATGTGTGGTGCATTTGCTCATGGTGCTATTTTCTTTATTTCGTGATTATGAT
 CCTGAACAAAACAAAGGTAACGTTTTAGCAGCTGTTTTAGATCATAAAGAAGCTATTTTACACACTTAAGTTG
 GGTATCTTTATTCTTAGGTTCCATACTTTAGGTTTACGTTTCATAATGACGTAGTTCAAGCTTTCCGGTACACCA
 GAAAAACAAATTTAATTGAACCAGTTTTTGCACAATGGATTCAAGCTGCTCAAGGTAAGTCATTATATGGATT
 TGACTTATTATTAGCTTCTCCTTCTAGCCCAGCTTACTCAGCTGGTCAAAGCTTATGGCTTCCAGGTTGGTTAGA
 AGCTATTAATAATAACCAAAACTCTTTATTCTTAAACAATTGGGCCAGGTGACTTCTTAGTACACCATGCTATTGC
 TTTAGGTCTTCATACAACAACACTTATTTAGTTAAAGGTGCCTTAGACGCACGTGGTCTAAATTAATGCCAGA
 TAAAAAAGATTTCCGGTACAGTTTCCCTTGTGACGGTCTGGTCTGGTGGTACTTGTGATATTTAGCTTATGA
 TGCTTTCTATTTAGCTGTTTTCTGGATGCTAAACACTATTGGTTGGGTTACTTTCTATTGGCATTGGAAACATTTA
 ACATTATGGCAAGGTAATGTATCACAATTCGATGAATCTTCTACTTATTTAATGGGTTGGTTAAGAGATTACTTA
 TGGTTAAACTCACAACCTTATTAATGGTTACAACCCATTTGGTATGAACAGTCTTTCAGTTTGGGCATGGACT
 TTCCTTTTTGGACACCTTGTATGCAACTGGTTTCATGTTCCCTATTTCTTGGCGTGGTACTGGCAAGAGTTAA
 TTGAAACTCTGTTTGGGCTCACGAAAAACTCCATTAGCTAACTTAGTTTACTGGAAAGATAAACCCAGTAGCT
 TTATCTATTGTTCAAGCTCGTTTAGTAGGTTTAGCTCACTTCTCAGTGGGTTATATCTTTACTTATGCTGCTTCTT
 AATTGCATCTACTGCAGGACGTTTTGGTTAA

psaC

ATGGCTCATGTTGTAATAAATTTATGATACTTGTATTGGTTGTACACAATGTGTACGTGCTTGTCTTTAGACGTTT
 TAGAAATGGTACCTTGGGATGGTTGTAAAGCGGCACAAATGGCTTCATCTCCTCGAACAGAGGACTGTGTTGGT
 TGCAAACTTGTGAAACAGCATGCCAACTGATTTTTTAAAGTGTACGTGTTTATTTAGGTAATGAAAGCACAAAG
 AAGTTTAGGTTTAGCTTATTA

psaJ

ATGAAAGATTTTACAACATATTTATCAACAGCTCCTGTTGTTGGTTTAGGCTGGGCTATTTTTACATCTGGTTTAT
 TAATTGAAATTAATAGATTTTTCCAGATCCATTAGTTTTTCTTTTTAA

psbA

ATGACAGCAATTTTAGAACGTCGTGAAAATACTAGCCTTTGGGCACGTTTTTGTGAGTGGATCACTTCAACTGA
 AAACCGTTTATACATCGGTTGGTTGGTGTATCATGATCCCAACATTATTAACCTGCGATCTCAGTATACATCAT
 CGCATTATCGCAGCACCTCCAGTTGACATCGATGGTATCCGTGAACCAGTTTCTGGTTCATTATTATACGGTAA
 CAACATCATCACTGGTGCAGTTGTTCCAACCTCAAACGCTATCGGTTTACATTTCTACCAATCTGGGAAGCTGC
 TTCTTTAGATGAGTGGTTATACAACGGTGGTCCCTTACCAATTAGTTGTATGTCACCTTCTTCTTAGGTGATGTTGT
 TACATGGGTCGTGAGTGGGAATTATCTTACCCTTAGGTATGCGTCCATGGATCGCTGTTGCTTACTCAGCTCCA
 GTAGCTGCTGCAACTGCTGTATTCATCATCTACCCTATCGGTCAAGGTTCACTTCTGATGGTATGCCTTTAGGT
 ATTTCAAGTACTTTAACTTCATGATCGTATTCCAAGCTGAGCATAACATCTTAATGCACCCATTCCACATGTTT
 GGTTGTGCTGGTGTATTTGGTGGTTCATTATTCTCAGCAATGCACGGTTCATTAGTTACTTTCATCTTTAATCCGTG
 AAACACTGAAAACGAATCAGCTAACGCTGGTTACAAATTCGGTCAAGAAGAAGAACTTACAACATCGTAGC
 TGCTCACGGTACTTTGGTTCGTTAATCTTCCAATATGCCTCATTCAATAATTCTCGTTCATTACACTTCTTCTTA
 GCTGTATGGCCAGTTGTATGATCTGGTTAAGCTTTAGGTTATTTCAACTATGGCATTGCACTTCAAGAACTGTTT
 AACTTCAACCAATCAGTAGTTGATTCAAACGGTCTGTATTAACACTTGGGCTGACATTATCAACCGTGTCTAA
 CTTAGGTATTGAAGTAATGCACGAACGTAACGCGCATAACTTCCCATTAGACTTAGCTTCAGTAGAAGCTCCTT
 CAGTTAACGGTTAA

psbB

ATGGGATTACCGTGGTATCGTGTTCACACAGTAGTAATTAATGACCCAGGTCGACTAATTTCTGTGCATTTAATG
 CACACAGCTCTAGTAGCTGGTTGGGCTGGTGTATGACATTTTGAATTTGCAGTTTTTGATCCATCAGATCCA
 GTATTAACCCATATGTGGCGTCAAGGGATGTTCTGTTCTTCTTACACGTTTAGGTTGACACAATCATCGG
 GGTGGTTGCAAAATTAGTGGTGAACATCTTCAAACCCAGTATGACACAAGTACATCTGAGGTTAAGGTCAGCTTCGCA
 CATTGTTCTTTCAGGTTTATTATCTTGGTTCAGTTTGGCACTGGGTTTACTGGGATTTAGAATTATCCGTGAT
 CCACGTACAGGTAACACTGCATTAGATTTACCAAAAATTTTTGGTATTATCTTTTCTTAGCAGGACTTCTTTGT
 TTTGGTTTTGGTGTCTTCCACGTAACAGGTGTTTTTGGACCTGGTATTTGGGTATCAGATCCATATGGATTAACA
 GGTAGGTACAACCAGTAGCTCCTTCTGGGGTGTGAAGGGTTTACCCTTACAACCCAGGTGGTGTACCAGC
 TCACCATATTGCTGCTGGTATTTTAGGTGTATTAGCAGGTTTATTCCACCTTTGTGTTCTGTCATCAATTCGTTA
 TATTTTGGTTTATCAATGGGTTCTATTGAATCAGTTTTATCAAGTAGTATTGCTGCCGATTTCTGGGACGTTTTCG
 TAGTAGCAGGTACTATGTGGTATGGTTCTGCAGCAACTCCAATTGAATTTTGGTCCAACACGTTACCAACTTGG
 GACCAAGGTTTCTTCCAACAAGAAATTCAAAAACGAGTAGCACAAAGTACATCTGAAGGTTTATCTTTTCAGA
 AGCTTGGGCAAAAATTCCTGAAAAATTAGCTTTCTATGATTACATTGGTAATAACCCAGCTAAAGGTGGATTAT
 TCCGTACAGGTGCTATGAACAGTGGTGTGGTATCGCAGTAGGTTGGTTAGGACACGCTAGTTTTAAAGATCAA
 GAAGGTGCTGAACCTTTTGTTCGTGCTATGCCTACTTTCTTTGAAACTTCCCTGTTGTTTTAATTGATAAAGACG
 GTGTTGTTCTGTGCTGACGTACCATTCCGTAAAGCTGAATCAAATACTCAATTGAACAAGTTGGTGTTCAGTTA
 CATTCTATGGTGGTGAATTAATGGTTTAACTTTACTGATCCCTTAACTGTTAAAAAATATGCACGTAAAGCTC
 AATTAGGTGAAACTCTTGAATTTGACCGTTCGACTTACAATTGACGGTGTATTCGGTAGTACGCCACGTTGGT
 GGTTCACTTTCGGACACTTATCTTTGCCTTATTATCTTCTTGGTCATATTTGGTCATATTTGGTCAAGAACTATTTT
 CGTGACGTTTTCTGCTGGTATTGATGAAGACATTAATGATCAATTAGAATTCGGTAAATATAAGAACTTGGTGA
 TACTTCATCTGTTCTGTAAGCTTTCTAA

psbC

GTGGAAACACTTTTTAACGGAACATTAACCTGTAGGTGGTCTGTGACCAAGAACTACTGGTTTCGCATGGTGGGC
 CGGAAATGCTCGTTTAAATCAACCTATCAGGTAACTATTAGGTGCACACGTTGCACACGCTGGTCTTATCGTTTT
 CTGGGCTGGTGTATGAACCTTATTCGAAGTTTCTCACTTCGTTCCAGAAAAACCTATGTACGAACAAGGTTAAT

TCTTTTACCTCACATTGCTACTTTAGGTTATGGTGTAGGTCCAGGTGGTGAAGTACTTGATACTTTCCATACTTC
GTATCAGGTGTTTTACACCTTATCTCTTCAGCTGTTTTAGGTTTTGGTGGTGTACCCTCTTTAATTGGACCTG
AACTTTAGAAGAATCTTACCCATTCTTTGGTTATGTTTGGAAAAGATAAAAAACAAAATGACTAACATTTTAGGTT
ACCCTTAATCATTTTAGGTTGTGGTGTCTGGCTTCTGTGTAAAAGCTTTATACTTTGGTGGTGTATGATAC
TTGGGCGCCTGGCGGTGGTACGTTTCGTATCATTCAAACCCTACAACCTAACGCTGCTATCATCTTTGGTTACAT
CGTTAAGTCACCTTTCGGTGGTATGTTGGATTGTTAGTGTGGATAACTTAGAAGACATTATTGGCGGTACAT
TTGGATTGGTACTTTATGTATTTTAGGTTGGTATTTGGCACATTTACACAACCTCCATGGCCTTGGGCACGTCGTC
TTTTGTTGGTCTGGTGAAGCTTACCTTTCTTACAGTTTAGCTGCTATTTCTCTTATGGGTTTCACAGCTTGTGTT
TCGGTTGGTTCAACAACACTGCATACCCTAGTGAGTTCTACGGCCCGACAGGTCTGAAGCGAGTCAAGCTCAA
GCATTTACTTTCTAGTACGTGACCAACGTTTAGGTGCGAACGTTGCATCAGCACAAAGTCTACTGGTTTAGGT
AAATACTTAATGCGTTCTCCAACAGGTGAAATATTTTCGGTGGTGAACCTATGCGTTTCTGGGACTTCAGAGG
ACCTTGGGTAGAGCCTTTACGTGGCCCTTCAGGTTTAGACTTAGTTAAACTAAAAATGATATTCAACCTTGGC
AGAAGCTGTCAGCTGAATATGACTCATGCTCCATTAGGTTTAACTTAACTCTGATGGTGGTGTGGCAACA
GAAATTAACGCTGTTAACTTCGTTTCTCCTCGTTCTGGTTAGCTACATCACACTTCTGTTTAGGTTTCTTCTCT
CGTAGGTCACCTATGGCATGCTGGTCGTGCTCGTGCAGGTTTCGAAAAAGGTATTGACCGTGTG
ACGAACCTGTTCTTCAATGCGTCTTTAGACTAA

psbD ATGACTATTGCGATTGGTACATATCAAGAAAAACGTACTTGGTTTCGATGATGCAGATGACTGGCTTCGTC AAGA
CCGATTTGTAATCGTAGGTTGGTCAGGTTTATTATTATTACCATTGTGCTTACTTTGCTGTAGGTGGTTGGTTAACT
GGCTGTACTTTTCGTTACTTTCATGGTACACTCATGGATTAGCAAGTTTACATACATTGAAGGTTGTAACCTTCTAACT
GCAGCTGTTTCTACACCAGTAAACAGTTTAGGTCACTCATTATTATTTCGTTTGGGCTGCTGAAGCACAAGTGAT
TTAACTCGTTGGTTCCAACTTGGTGGTTTATGGGCATTCGTAGCATTACACGGTGTCTTTGGTTAATTGGTTTCA
TGTTACGTCAGTTTGAATGTCACGTTTCAGTAACTTACGTCCATAACAACGCTATCGCGTCTCAGCACCAATTG
CTGTATTGCTGTCTGTATTCTTAATTTACCCTTAGGTCATCAGGTTGGTCTTTGCTCCAAGCTTTGGTGTGTC
ATCAATCTCCGTTTCATTTTATTCTTCCAAGTTTTTACATACTGGACTCTAAACCCATTCCACATGATGGGTGTA
GCTGGTGTTTTAGGTGCAGCATTATTATGTGCTATTACGGTGTACTGTAGAAAACACTTTATTTCGAAGATGGT
GACGGTGTAAACACTTTCCGTGCATTCAACCCAACCTCAAGCTGAAGAACTTACTCAATGGTAACTGCTAACCC
TTTCTGGTCACAAATTTTCGGTGTAGCTTTCTCTAACAAACGTTGGTTACTTCTTCATGTTATTTCGTACCTGTA
ACTGGTCTTTGGATGAGTGCTTTAGGTGAGTAGGTTTAGCTTTAAACTTACGTGCTTACGACTTCGTATCACAA
GAGATCCGTGCTGCTGAAGATCCTGAGTTCGAAACATTCTACACTAAAAACCTTCTACTTAAACGAAGGTATTG
TGCTTGATGGTGTCTCAAGACCAACCACATGAAAAATTAACATTACCTGAAGAAGTTTTACCTCGTGGTAAACG
CTCTATAA

psbE ATGGCTGGAAAACCAGTAGAACGTCCTTTCTCAGATATTTTAAACAAGTATTTCGTTATTGGGTTATTACAGTATT
ACTGTACCTTCATTATTTATTGCAGGTTGGCTTTTGTAGTACAGGATTAGCTTATGACGTATTTGGCAGTCCAC
GTCCAAATGAATATTTTACAGAAGATCGTCAAGATGCTCCTTTAATTACTGACCGTTTTAATGCTTTAGAACAAG
TTAAAAAATTATCAGCACATAA

psbF ATGACTACAAGAAAATCAGCTGAAGCTGTTACATATCCAATTTTTACTGTACGTTGGTTAGCTATTCACGCAATT
GCAGTTCCTACTATTTCTTTTTAGGTGCAATTACAGCTATGCAATTTATAACAACGTTAA

psbH ATGGCTACAGGAAAAATAGTACACAACTTCAACATCACAGAACCAGGAATTGTTACGCCATTAGGAACTT
TATTACGTCTCTAAACTCTGAAGCAGTAAAGTTTTACCTGGATGGGGTACTACAGTCTTATGGCTGTAGCAA
TTCTTTTATTTGCTGTTTTTTTACTTATTATTCTAGAAATTTATAACAGTTCATTAATTTTAGATGGTGTACTAAT
AGCTGGGAAAGTTTATCCGCTAAATAA

psbI ATGTTAACATTA AAAAATCTTTGTTTATACTGTGTAACTTTCTTCGTTGGATTATTCATTTTTGGATTCTTATCAA
ATGATCCATCTCGTAATCCTGGTAAAGGTTTAGACTGTGCGACACGTTTCCTATAA

psbJ ATGTCAAACATTGGAACACTGGACGTATTCCTTTATGGCTGTGGGTACAGTTGTTGGTCTTTTAGCTATTGGT
TTATTAGGAATCTTTTCTATGTTTCTACGTTGGTTTAGGTTCTTCTCAATAA

psbK ATGGCAAATTTAGTAATGGTTATTGCAAACTTCTGAAAGCTTATGCTCCTTTTCACCAATTGTTGATGTTTTAC
CTATTATTCCAGTTTTATTTATTTATTAGCTTTTGTGGCAAGCTTCAAGTTAGTTTTAGATAA

psbL ATGGCTAGACCAAATCCAAATAAACAAGTAGTTGAATTAACCGTTCAAGTCTATATTGGGGACTTTTTATTAAT
TTTTCGTTTTAGCTGTTTTATTCTGTGCGACCCGTTTTCCCTATAA

psbM ATGGAAGTAAATATTCTAGGATTAATTGCTACTGCTTTTATTATTATTTCAACTTCTTTTTTATTAATTCTTTA
CGTTAAAAACAGCCGATCTGAACAATCTGAATCTTAA

psbN ATGGAAGTCCGGCTTTTTTCTTTACCTTCTTTTTATGGTTTCTTTTATTAAGTGCTACAGGTTATTAGTTTATGT
TAGTTTTGGACCTCCTTCAAGAAAATTGAGAGATCCTTTTGAAGAACATGAAGATTAA

psbT ATGGAAGCTTAGTTTACACTTCTTATTAATTGGAACGTTAGGTATTATCTTTTCGCTATCTTTTTTAGAGAGC
CTCCACGTATTGCAAAAATAG

psbZ ATGACATCAATCTTCAAATTCGCTTACTTGGTTAGTTTTAGTTTCTTTTGCATTAGTAGTAGGTGTACCAGTAG
TGTTGCTCTCTCAATGGTTGGACTGAAAATAAAGGAGTAGTTTTCTCTGGTTAAGCGTATGGTTTTTATTAG
TTTTTGCTGTAGGTGATTTAATTCTTTTGCAGTTTAA

rpl2 ATGGGAATTCATTTATTAATCCGTGTACACCAGGTACACGTAATCGATCCGATCAGATTTTAGTGAATTGACT
CGTACAAAACCAGAAAAGTCATTAACATCTTATAGACAAAGATCAAACGGACGTAATAACCGTGGTATTATTA
CATGTGCTCATCGTGGAGGGGGTCACAAACGTTTATATCGTCAAATGATTTTAGACGTGATAAAATGGAGTA
CCTGCCAAAGTTATATCGATGGAATATGATCCAATAGAAATGCAAGAATTGCTCTTTTACATTATGAAGATGG

AGTAAAACGTTATATTTTACATCCACGTGGTTTAAATCCAGGTGAAACGATTATTTCTGATTTAAATGCACCTAT
 TTTAGTAGGTAATCTCTACCTTTACGTAATATTCCTTTAGGTGCCGAAATTCATAACGTTGAATTTCAACCCGG
 TTCTGGAGGTCAATTAGCACGAGCTGCTGGCTCTTTAGTTGAAATTTAGCTAAAGAAGGTGATTTTGTAACAGT
 TCGTTTACCATCAAAAAGAAATTCGTTTTATTTCAAAAAATAGTTGGGCTACTCTTGGTCAAGTTGGGAATGTAGA
 AGCTTATAACTTAATCGGTAAAGCTGGTCTCTCGTGGTATAGGTATTCGACCAACTGTGCGTGGTTCAGC
 CATGAATCCAGTAGATCACCCGCATGGTGGTGAAGGACGTGCACCTATTGGGCATAGTCGTCCATTAACAC
 CTTGGGGACGCCCTGCTTAGGTAACCTAACGCGTAAACCAAAAAATTTAGTAATGTTTTTATTATCCGTAAA
 CGTAAATAG
rpl5 ATGGCACAAACACAAAGATAGAAAAACAATACGTAGAAACAATTGTTCTTACACTTATGAAAAAATCCAAT
 AAAAAATATTCACCAAGTACCTAAAATTGAAAAAATTGTAATTAATCGTGGTATTGGTAGTGCATCACAAAAC
 CAAAAATGGTGGATTCCGCTTTAAAAGAATTAGGTATGATTGCTGGACAAAAGGTGTTATTACACGTTCTAA
 AAAAGCAATTGCTAGTTTTAAATTAAGAGAACAATGCCAGTAGGTGTTGCTGTGACTTTACGCGGTAATCGTA
 TGTATGGATTTTTAGACCGATTAAATCCATTTAGCATTACCGATTTCGCGATTTCAGGTTAATTAACCAAGAA
 GTTTTGACAAAAAAGGTAATTATAGTTTAGGTCTAGAAGAACAATTAATGTTCCCTGAAATTGAATATGATAAA
 ATTGATCAAATTCGTGGTATGGATATTTCTATTGTAACAACAGCAACAAAACAAGAAGAAGGATTAAGCCTTTT
 ACAAGAATTTGGATTACCTTTTAAATAA
rpl14 ATGATCAAACCGCAATCTTTTTTAAATGTTGCTGACAATAGTGGTGCACGTAATAATGTTGATTTCGTGTTTTA
 GGATCAGGTTTACGTAATATTGGTGACACTATTATTGCAGTCGTAAAAGATGCTATTCCATAATATGCCAACAAA
 AAAATCTGATGTTGTACGAGGAGTTATTGTACGTACAAAAAAGGCATTTCGTAGACCAAAATGGTATGAATATTA
 AATTTGATGATAACGCTGCTGTTATTGTTAATAAAGAAGAAATTTACGAGTACTCGTGTTTTTGGCCTATTG
 CTCGTGAATTAAGAGATCGTAATTTTACTAAAATTATCTTTAGCACAGAGGTTGCTATAA
rpl16 ATGTTAAGTCCAAAAAGAACAATAATTCGTAACAACATCGTGGTAGATTAAGGAAAAAGCGTATCGTGGAA
 ACACAATAGCTTTTGGTGATTTTGGCTTTACAAGCTTTAGAACCATGTTGGATTACTTCTCGTCAAATTGAATCAG
 GAAGACGTGTTCTTACACGTTATGTACGTCGTGGTGGTAAATTATGGATTTCGATTTTTCCAGATAAACCTATTA
 CTGCACGTCCAGCTGGTACACGTATGGGATCTGGTAAAGGTGCTCCTGAATATTGGGTAGCTGTTGTTAAACCA
 GGTAATAATTTATGAAATTCGCGGTGTTAATGAAATTTGAGCGCGTCACGCTTTACGAATTGCTTCGTATAAAA
 ATGCCAATCAAAACTAAATTTTTAGGCCGTACTTCTGTTTAA
rpl20 ATGACTCGTGTAAAACGTGGTAATGTTTCTAGAAAAACGCCATAAAAAAATATTAAGTTAAATAAGGGTTCCG
 TGGTTCAGCATCGGCTCTTTATCGTACTGCTAAAACAACAAAATATGAAAGCTTTGCGTTATTCTTATGCTAATCG
 TCGTCAAAAAAACGTGAATTTAGAAGTATTTGGATTACACGTTTAAATGCTGCGATTAGACGTTATGGATTGA
 ATTATTCTGAATTTATTTATTTAAAAAATCGTTCAATCCAATTAATCGTAAAGTTTTAGCTCAATTATCAAT
 TTGTGATCCTGAAGCGTTACACAATTGTTACTTTTTTAA
rpl23 ATGACTGAAACTACAGAAAATGTAACATAAAGATGCTAATATGTCGCCTCAATCAATTTCTTTTTAATATTGATTTA
 GTCAAATCCTATTATTACAGAAAAAACCTATTTAGCCCTATATAAACATCGTCAATATACGTTTGATATTGAT
 AAACGTTTAACTAAAACGCAAAATTAATAAATTTCTCTAATTTATTTAATGTGGATGTAATTGCAGTTAATACA
 CACATACCTCCTCGACAACAACCTCGTGTGGTATGGCACAAGGTTTCCGTCCTCTTTTTAAAAGAGCTATTGTA
 ACTTTAAAAGAAGGACAATCAATTAATTTCAATTTAAAACCAAATTCATAA
rpl36 ATGAAAGTACGTTCTTCAGTTAAAAAATCTGTGACAAATGTCGTGTTATTCGCCGAAAAGGTACTGTAATGGT
 TATTTGTTCTAACCTAAACATAAACACGTCAGGTTAA
rps4 ATGGCTAGATATATTGGACCCAAATTAATAATTTCTGTCGAATTGGTAAATTACGTGGTTTTAACACGAAAAAA
 ACCTTTTCTGTCGTGTTTATCGTGGAAAGAGGTCCTTTAAGAGGTAAGTAATTCCTCCAGGACAACATGGTCTTAT
 TAAACTATTTAAAACAAGACCTTATGATTTCTGTGAATCAGATTACTAATTCGTTTAAAAGTAAAACAACGTTT
 ACGTTATAACTATGGTTTAAACAGAACGTCAATTAGTAAATTATGTACGTAAAAGCCAAAAAATTAAGAAGCA
 ACAGGTCAAGTTCTTTTACAATTATTAGAAATGCGTTTAGATAATATTGTTTTCTGTTTAAATATGGCTCCGACT
 ATTGTAGCTGCACGTCAATACATTAGTCATGGACATATTCGTGTAATAATAAAAAAGTTAATATTCCAAGTTA
 TATGTGTAACCTAAAGACGTAATATCAGTTGCTATGAAAGAAAAATCTCTTATTCTTATTAATAGAACTTAA
 ATGAATATTATCAACGTATGCAGTTTTATAAAAAACGTTTAGAAAAAACACTTGCGTTTATTTTATTTTACGCTAA
 AACTAGTACCTAATATGGGTCTGCTTTACAATTAATTAATGGTCCAGGAGCTGTTGTAATAATAATAATAGA
 CGTGTACGTAATCCTAATCATATTTGTAATTTCTAAAGACGTTCTGTCTATTACAACACGTGAAGGAACACGTC
 AATAAACTTTCTTAA
rps8 ATGATTACATCATTAACCTCGAGGTTTTATTTCATGACAGTATTAGTGACATGCTTACACGAATTAGAAATGCTTGT
 TTAGCAAAAAAATCTAATGTTATTGTACCACACACTAAGCTGAATCTAGAAATAGCACATATTTTAGAAAAAGA
 GGGTTTTATCCAAGGTTTCAAATGTCACTAGAAATCGAATGATATAATTATTTCGTTTAAAATATCGTTCAAAACA
 AGTTTATCGTGGTAAAACAAAAGAATCTTGTATTACTAATCTTAAACGAATTAGTAAACCTGGCCTTCGTATTTA
 TGCAAACTAAAGAAATTCACGTAATTTAGGTGGTACAGGTAATTGTAATTGATCAACTCTACTGGTATTAT
 GACTGATCGAGAAGCTCGTTTTAAAAAATGGTGGGAGATAATTTGCTCAGTTTGGTAA
rps9 ATGGCGAGATTGGTAAAAAATCAATTTATGAATTTTTAGCGCGTAGTGTAAGAACGTAAGGAAGCTGTAG
 CACAGGTGCAAATTGTTAAAGGTACAAATGAAGCACCGGGTGGTCAATTTATTATTAATAACATACCAGCAGA
 AAATTTATTACACAATAACACATGTTCAATTTCTTTCAGTTAAAGCACCTATTAGTATTTTACAAAGCTTACATGA
 TGCTAGTGTAGAAGGTGAAAATACTGATTTTTAAAAATGACTTAGGTAATTTAATACTGTTGTAAGGTTAAAG
 GTGGTGGCTTAGTTGGTCAAGCAGAAGCTATTAATTAGGATTAACACGAGCTATTTGTCATTTATATGCTATTA
 CACCGAATAATCAAAATAAATCTATTAGAGAATTTCCAATTGAAAATTCATGGAAGAATCTTCAAAAAATTC

TCAAGTGAATGAGCTATGAAACTACAATAGAACTACGTAAACAATTCAAAGATAAAGGATTTTTAACACAAG
 ATTCACGTGTTAAAGAACGTAGAAAATATGGTTTAAAGAAAGCAGCAAAGCTTACAATACCACAAACGTTA
 A
rps11 ATGGCTAGACAAAACAAGAAAACCTTACATCAGCTAAAACAAGAAAACGTGTTTATAGAGGAATTGTTACATAC
 AAGCTGGGCATCATAATACTATTGTAACCTATTACTAACATTAGAGGAGAAGTTTTATGTTGGAGCTCATCTGGT
 GCTTGTGGCTTTAAAGGAAAACGTAAATCAACTGGATTTGCAGCAAAAAAGCTGCAGAAAACAGCTGCTAAAA
 GATCACGTGATTTTTTAATGCGAGAAGCTAAAGTCTTAGTAAATGGTCCGGGACAAGGACGTGAAAAGTCCATT
 CGTGAAATTTTTAAAGCAGACATAAAAGTAAATGTAATTCGTGAAAAATCAGGAATTCCTCATAACGGATGCCG
 CCCACAAAAGCTCGTAGTGATAA
rps12 ATGCCTACTATTCAACAATTAATACGTTTCAGCAGCAAAAAACAACCACAAAAAGTAAAGCACCGGCTTAA
 AATCATGTCCACAACGACGAGGATTTGCTTACGTGTTTATACTGTAACGCCTAAAAAACCAACTCAGCGTTA
 CGTAAAGTAGCTCGTATTTCGATTAACCTCAGGTTTCGAAGTTACAGCATATATTCCAGGTATTGGACATAATCTT
 CAAGAGCATGCGGTGGTCTTGTTCGTGGTGAAGGGTAAAAGATTTACCAGGTATCGTTATCATATTGTACG
 TGGTACTTTAGATACTGCTGGTGTAAAAAACCGTACTCAAAGTCGTAGTAAATATGGTGCTAAAAATGGCATCAA
 AACAGCAGCTAAAAATAA
rps18 ATGGCAGCTCCAAATCTTCAAAAAAACTACAAAAAATGTCCGTGGTAAATTTAACACATCAAATACAAATCC
 ATCTCGTGTTTATCTTTATCGCAATTATTAGGTAGTGTTCAAAAAACAACGTACTCAACGTGCTAAACGATCAAC
 AGGTAGAACTCCTATTAACCATTAATTCACCAAAGTCTTTAATTATTATTTAAAAGAAAAACCTGAAAAAG
 CTGTATATACTCGACGTATTATTGATTATAACATTGTGGCTTATTACAACGTTACATCGGTTTAGGTGGTAAAA
 TTTTACTAGAAGACAAACACGTTTAAACAGCTAAACAACAACAGCTATATTGCTAAAAACAATTTAAAGTGCACGT
 GTAATGGGTTTATTACCATTGTAAAGTAAAGAAAAAGGTTTTTTTAGATAA
rps19 ATGCCACGTTCAATTACAAAAGGTCCATTTGTAGCAGGACATTTATTAAAAAAAGTAGAAAAATTTAAATAGTAT
 AGGTAAAAAATTGTTATGAAAACCTGGTCTCGTTCATCTGTGATTATTCCTCCAATGATAGGTACACAAATTGC
 TTGTTATAATGGTAAAGAACATATTCCAATTTATATAACAGACCAATGGTGGGTCACAAACTAGGGGAATTTG
 CCTAACACGTAATTATCGTGGTCACATTAAGAACGATAAAAAATCAAAACGATAA
tufA ATGGCACGTGCTAAATTTGAACGTAAAAAACCATGTAAATATTGGTACTATTGGACATGTAGACCATGGTAA
 AACTACATTAACAGCTGCTATTACAATGACATTAGCAGCTAAAGGTAGTGGATTAGCAAAAAAATATGATGAA
 ATTGATTCTGCACCTGAAGAAAAAGCTCGTGGTATTACAATTAATACAGCACACGTAGAATATGAAACAGATA
 ACCGCCATTATGCACACGTAGATTGTCCAGGTATGCTGATTATGTTAAAAACATGATTACAGGTGCTGCACAA
 ATGGATGGTGTATTTTAGTTGTATCTGGTGCAGATGGTCCAATGCCACAAACAAAAGAACATATTTTACTTGTCT
 AAACAAGTTGGTGTTCCTAATATTGTTGTATTCTTAAACAAGAAGATCAAGTAGATGATAAAGAATTATTAGA
 ATTAGTAGAATTAGAGGTTTCGTGAAACTTTAGATAAATACGAATTCCTGGTGATGAAATTCCTATCGTTTCAG
 GTTCTGCTTTATTAGCTTTAGAAGCACTTGTGAAAAACCTAAAATTTAAACGTGGTGAACCAATGGATTGAT
 AAAATTTATGCTTAATGGATAAAGTAGATTCTTATTCCAACACCAGAACGTCAAACAGAAAAACCTTTCTT
 ATTAGTTGTGAAGATGTTTTATCTATTACAGGTCGTGGTACAGTTGCTACAGGTCGTGTTGAACGTGGTACTTT
 AAAACTTGGTGAACCGTAGAAGTAGTAGGTTAAAAGAAACAAAAAGTACAGTTGTTACTGGTTTAGAAATG
 TTTAAAAAACATTAGACGAAACAATGGCTGGTGATAACGTAGGTGATTATTACGTGGTATCCAAAAAAG
 ATATCGAACGTGGTATGGTTTTAGCTAAACCTGGTAGTATTACACCACATACAAAATTTGAATCTCAAGTTTATA
 TTTTAAACAAAAGAAGAGGGTGGTCTGCTCACTCAGCTTTCTTAACTGGGTACACACCACAATTTATGTACGAACA
 ACAGACGTAACCTGGTAAAGTAGCTGGTTTTTCACATATTCAAATGCGTAATCCTTCATCTGTTGCTGAAGAACAT
 TCTAATAAAATGGCAATGCCTGGTGACCGTATTAGTATGGTTGTTGAATTAATTAACGCAATTTGCTATTGAAAA
 AGGTATGCGTTTTGCTATTTCGTGAAGGTGGACGTACTGTAGGTGCAGGTGTTGTTACATCTATTATTGAATAA
yef3 ATGCCACGTACACAACGAAATGATAATTTTATTGATAAAACATTTACAGTAATTGCTGATATTATTTAAAAAGTT
 TTACCTACTTCACAACGTGAAAAACAAGCTTTTTCTTATTATAGAAATGGTATGTCAGCTCAAGCCGAGGGTGA
 GTATGCTGAAGCACTTCAAATTTATTATGAAGCTATGCGTTTAGAAACAGATGCATATGACAGAAGTTATATTC
 TATATAACATAGGACTAATTCATACTAGTAATGGTGAACATGGAAGAGCACTTGAATATTATTATCAGTCTTTA
 GAAAGAAATCCTTCTTACCTAGTGCTCTAAATAATATTGCGGTAATTTATCATTACCGAGGAGAACAAGCTGC
 TTTCCGGTGGTCAATCAGAAATTCACAAATTTATTTGAAAAAGCTGCAGATTTTGGAAAGAAGCTATTTCGCT
 AGCTCCAACAATTTATTGAAGCACAAAACCTGGTTACAAATGACAGGTCGCTGGACTGGTGGTGCATAA
yef4 ATGAAGGAGAAAAATATGACAAAAAATGAACAAAAAACATCGCTAGAAACGAGTGTCTTAACAGGTACGCAA
 AATAAACTGCCAATAAAAAGCAGCGCGTAACCTTATAAGACGTTATATTATTATAGGAGAAAGACGTATAAGTA
 ATTATTGGTGGGCTTTCATTGTATTATTAGGCTCTTCTGGATTTTTATTAAGTGGGATTTGTTCTTATAATTATAA
 TGCAGGTGATTCTCAGGAATTCTATTTAATTTGCTTCCGTTCTTGATTTTTAAAGACATTGCTTTTTTCCCACAG
 GGACTTTAATGCTTTTTTATGGAGGTTTAGGATTTCTTTAAGTTTCTATTGGTGGTTTCTTATCTTTTGGAAATGT
 AGGAGGTGGTTTTAACGAATTTAATAAAAAAGAAGGCTTTATTAGAATTTTCAGATGGGGTTATCCAGGAAAAA
 ATCGTCTGATCGATTTTACATATCCTTTAAAAGATTTAAGAAGCCATCAAAGTTGAATTTAAAACAAGGATTAGAT
 CCTCAACGAACTTTTTTTAAAACCTAAAAGGAAAAAAGAAATTCCTATTACTGGAAATAGGTGACCTTTAAC
 TTTAAAAGAAATTTGAAAAACAAGCCTCTGAATTAGCTAATTTTCTGGATGTAAGTATTCTTACATAA
yef12 ATGAGTCTAACTCTTATTCTACAGTTAGTAGCCTTATTTGCGGTTGTAGCAGCTGGGCCACTTGTGTTGTTTTAT
 TATCTGTTTCGTGGTGGTAATCTTTAA

Supplementary Table S2. Synonymous substitution rates in plastid and mitochondrial genomes of various plastid-bearing lineages.

| | Substitutions per synonymous site (SD) | | Substitution rate ratios (pt : mt) |
|---------------------------------|-------------------------------------------|-------------|---------------------------------------|
| | ptDNA | mtDNA | |
| ARCHAEPLASTIDA | | | |
| Glaucophytes^a | | | |
| <i>Cyanophora</i> | 1.01 (1.22) | 5.29 (3.17) | 1 : 5.3 |
| Green algae^b | | | |
| <i>Dunaliella</i> | 0.09 (0.32) | 1.16 (0.52) | 1 : 12.9 |
| <i>Chlamydomonas</i> | 0.30 (0.11) | 0.29 (0.05) | 1 : 1.0 |
| <i>Mesostigma</i> | 0.11 (0.06) | 0.17 (0.11) | 1 : 1.5 |
| Seed plants^c | | | |
| Angiosperms | 0.39 (0.01) | 0.13 (0.01) | 1 : 0.3 |
| Gymnosperms | 0.61 (0.03) | 0.28 (0.02) | 1 : 0.5 |
| Red algae^d | | | |
| <i>Porphyra</i> | 0.47 (0.22) | 1.76 (0.58) | 1 : 3.7 |
| HAPTOPHYTA^e | | | |
| <i>Phaeocystis</i> | 0.25 (0.16) | 2.41 (0.97) | 1 : 9.6 |

SD: standard deviation. Synonymous-site substitution rates are based on averages among loci, not concatenations.

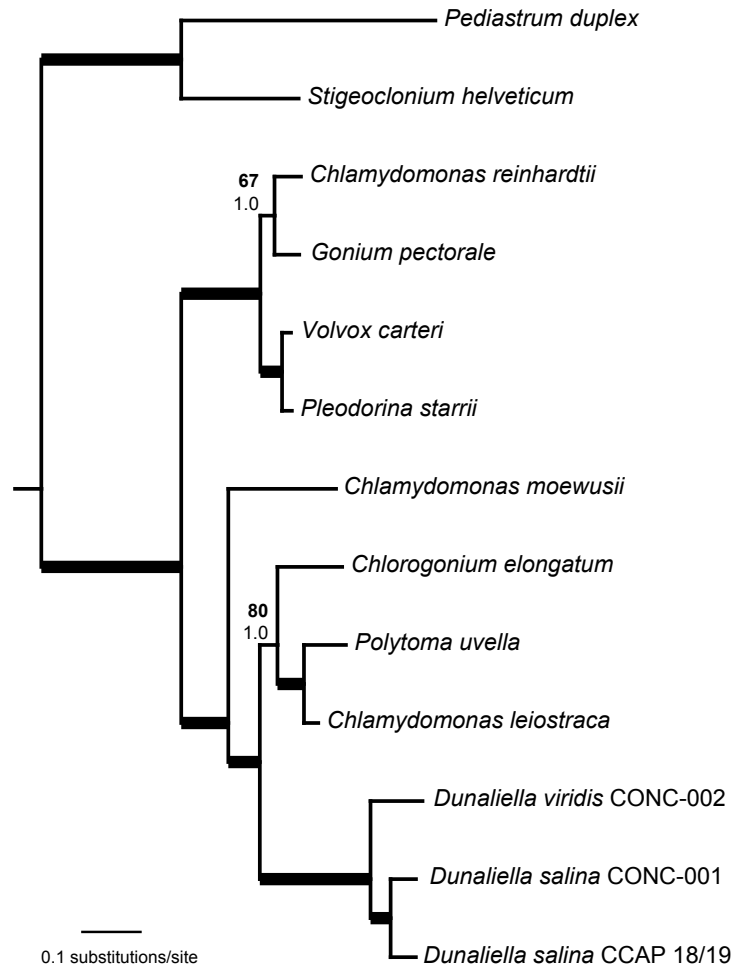
^aData from Smith, Jackson, et al. (2014).

^bData from present study, Popescu and Lee (2007) and Hua et al. (2012).

^cData from Drouin et al. (2008).

^dData from Smith et al. (2012).

^eData from Smith, Arrigo, et al. (2014).



Supplementary Figure S1. Maximum-likelihood protein phylogeny of chlamydomonadalean algae. Conceptual protein translations of seven mitochondrial genes (*cob*, *cox1*, *nad1*, *nad2*, *nad4*, *nad5*, and *nad6*) were aligned with MAFFT v7 (Katoh and Standley 2013) and manually refined using Se-AL v2.0a11 (<http://tree.bio.ed.ac.uk/software/seal/>). Multiple alignments are available upon request. The best-fit amino acid substitution model for the seven-protein concatenated multiple alignment (1931 amino acids long) was selected with ProtTest3 using the AIC selection criterion (Darriba et al. 2011). Maximum likelihood trees were estimated with RAxML v7.2.6 (Stamatakis 2006) using the selected WAG+I substitution model. Branch support was assessed with 500 non-parametric bootstrap replicates. Bayesian posterior probabilities were calculated with MrBayes 3.2.1 (Ronquist et al., 2012) using the WAG+I+Γ substitution model running a Metropolis-coupled Markov Chains Monte Carlo (MC3) for 1.5 million generations. Two independent MC3 runs were performed simultaneously starting from different random trees. Trees were sampled every 100th generation. Final posterior probabilities were calculated after discarding trees sampled from the first 500,000 generations. **Note:** Branching order of Figure 1 from the main text is based on phylogenetic analyses of Nakada et al. (2008), González et al. (2009), Smith, Hua, et al. (2013), and Figueroa-Martinez et al. (2015).

References

- Darriba D, Taboada GL, Doallo R, Posada D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics*. 27:1164–1165.
- Drouin G, Daoud H, Xia J. 2008. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Mol Phylogenet Evol*. 49:827–831.
- Figuroa-Martinez F, Nedelcu AM, Smith DR, Reyes-Prieto A. 2015. When the lights go out: the evolutionary fate of free-living colorless green algae. *New Phytol*. In press.
- González MA, Gómez PI, Polle JEW. 2009. Taxonomy and phylogeny of the genus *Dunaliella*. In: Ben-Amotz A, Polle JEW, Subba Rao DV, editors. *The Alga Dunaliella: Biodiversity, Physiology, Genomics and Biotechnology*. Enfield (NH): Science Publishers. p. 15-44.
- Hua J, Smith DR, Borza T, Lee RW. 2012. Similar relative mutation rates in the three genetic compartments of *Mesostigma* and *Chlamydomonas*. *Protist*. 163:105–115.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30:772–780.
- Nakada T, Misawa K, Nozaki H. Molecular systematics of Volvocales (Chlorophyceae, Chlorophyta) based on exhaustive 18S rRNA phylogenetic analyses. 2008. *Mol Phylogenet Evol*. 48:281–291.
- Popescu CE, Lee RW. 2007. Mitochondrial genome sequence evolution in *Chlamydomonas*. *Genetics*. 175:819–826.
- Ronquist F, et al. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 61:539–542.
- Smith DR, Arrigo KR, Alderkamp AC, Allen AE. 2014. Massive difference in synonymous substitution rates among mitochondrial, plastid, and nuclear genes of *Phaeocystis* algae. *Mol Phylogenet Evol*. 71:36–40.
- Smith DR, Hua J, Archibald JM, Lee RW. 2013. Palindromic genes in the linear mitochondrial genome of the nonphotosynthetic green alga *Polytomella magna*. *Genome Biol Evol*. 5:1661–1667.
- Smith DR, Hua J, Lee RW, Keeling PJ. 2012. Relative rates of evolution among the three genetic compartments of the red alga *Porphyra* differ from those of green plants and do not correlate with genome architecture. *Mol Phylogenet Evol*. 65:339–344.
- Smith DR, Jackson CJ, Reyes-Prieto A. 2014. Nucleotide substitution analyses of the glaucophyte *Cyanophora* suggest an ancestrally lower mutation rate in plastid vs mitochondrial DNA for the Archaeplastida. *Mol Phylogenet Evol*. 79:380–384.
- Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 22:2688–2690.