

editing genome atgactattattcctcgagagcgtgaagcaaaaaagtaaaaattgtattgattgtaat
M T I I P R E R E A K K V K I V I/V D C N

editing genome ccagttgtaacagatttttgataagtgggcacagccaggtcattttcccgtaactctggca
P V V T D F D K W A Q P G H F S R T L A

editing genome aaaggctcactacgactacttggatttgaatttgcagatgcgcatgattttgac
K G P T T T T W I W N L H A D A H D F D

editing genome attcactactctgatctacaagagatttctcgtaaagttttagtggtcattttgggtcag
I H T S D L Q E I S R K V F S V H F G Q

editing genome ttgggtattattttgatctggctcagtggtatataatttcatggagcagcattttcta
L G I I L I W L S G I Y F H G A R F S N

editing genome tacgaatcttggtgatagatccagttcatattaaagccagagctcaagtagtttggccc
Y E S W L I D P V H I K P R A Q V V W P

editing genome attgtaggtcaagaattctgaatggtagcgtaggtgggaattccaaggagttcagatt
I V G Q E I L N G D V G G E F Q G V Q I

editing genome gtctctggtttgttccaactatggcgtggtatgtattgtaagtgaactacaattgtat
V S G L F Q L W R G S G I V S E L Q L Y

editing genome agtacagctattgggtttgatactttctagagcaatatttttcaggttgggttcat
S T A I G G L I L S R A I F F A G W F H

editing genome tatcacaagcttccccacgtctagagtggtttcaaatgttgagtctattgaaatcat
Y H K A S P R L E W F Q N V E S I L N H

editing genome catcttcgggcttttgggtttgggagcttttctggcagggcaccaaattcatatt
H L S G G L L G L G S L S W A G H Q I H I

editing genome gcgattccagtaaatcgacttttgatgaagaggttgatccaaagagattcctcttct
A I P V N R L L D E G V D P K E I P L P

editing genome catgagtggtactagtcgtgaattgatccagctgtttcca ttttagaaaggga
H E W I T S R E L I S Q L F/Y P -/S F R K G

editing genome c tttgtaccatttttacacttcaatggcgtgagtatagtgacttttcaacgtttcagga
F/L V P F F T L Q W R E Y S D F L T F R G

editing genome ggattgaactcctgttactgggggttggatttgacagatgaagctcatcatcttgc
G L N P V T G G L Y L T D E A H H H L A

editing genome attgcggtattgtttattatcgctggtcatatatacgaacgaatttggattggacat
I A V L F I I A G H I Y R T N F G I G H

editing genome aatatacgtgaaattttggagccacatcggtgaccatttacgggtgctggtcataaagg
N I R E I L E A H R G P F T G A G H K G

editing genome ttgtatgaaattttgactacttcttggcatgcacaaatttctgtaatttggctatattg
L Y E I L T T S W H A Q L S L N L A I L

editing genome ggttctcttctattattgtggtcatcatatattctataccgcttatccatttgg
G S L S I I V A H H I Y S I P P Y P Y L

editing genome gctacagattatagaaccaacttctctattcactcaccatattggttgggtttt
A T D Y R T Q L S L F T H H M W I G G F

editing genome tgtgtagttggggtggtgcgcagcagcatttttatagttcagattatgatccgaga
C V V G G A G A H A A I F I V R D Y D P R

editing genome aacaattataataatcttttgatcgatttgcgcacatcgatgcggtaatctctcat
N N Y N N L L D R F V R H R D A V I S H

editing genome ttgaactgggtttgtatttttcttggctccatagttttggcctatataatcataatgat
L N W V C I F L G L/F H S F G L Y I H N D

editing genome acgataagtgcattgggtcgtcctcaagatatgttagtgatacggcaattcagctacag
T I S A L G R P Q D M F S D T A I Q L Q

editing genome cctattttgacagtggttgcagaaactcatatgttggcacctcaagtaacttctcca
P I F A Q W I Q K T H M L A P Q V T S P

editing genome gcagaggtatagctactagtgttacgtggggcggaatttgggttgcgtggagagaag
A E A I A T S V T W G G N L V S V G E K

editing genome gtggcaatgataccaatttctgttgggtactgctgactttatagttcatcatatccatgca
V A M I P I S L G T A D F I V H H I H A

editing genome tttacgattcatgtcacagttttgattcttttgaagggtgattttattctcgtagttcg
F T I H V T V L I L L K G V F/L Y S R S S

editing genome cgactaattccagacaaagtcaaccttggatttgcgatttccgtgtgatgggtcgggtcgt
R L I P D K V N L G F R F P C D G F G R

editing genome ggtgggacttgcagtttccggttgggatcatgttttcttgggttttctggaatgat
G G T C Q V S A W D H V F L G L F W M Y

editing genome aactctctttttagttgatttttctttagtttggaaaatacagtcgagtggttgggt
N S L S I V I F H F S W K I Q S D V W G

editing genome tcggtgactagttcagtgctctcatattaccggaggtaattttctcaaagtcgcaat
S V T S S S V S H I T G G N F S Q S A N

editing genome actattaatggttggctacgtgatttcttgggtcagtcgtctcaggttattcaatct
T I N G W L R D F L W A Q S S Q V I Q S

editing genome tacggatcgtctcttccggtatggattgatttttctgggtgctcatttttgggca
Y G S S L S A Y G L I F L G A H F V W A

editing genome tttagtttgatgtttttagttggtggttatttggcaagaattgatcgagttcatt
F S L M F L F S G R G Y W Q E L I E S I

editing genome g atttgggcacaataaaacttaaggttaataccaataattcagcctcagcgttgagatt
I/V W A H N K L K V I P I I Q P R A L S I

editing genome acacaaggtcgtgcagttggagttgcacattatttgcgtgggtgattagaacaacttg
T Q G R A V G V A H Y L L G G I R T T W

editing genome tcgttcttctagcgcgtattatttctgttggtaa
S F F L A R I I S V G -

Supplementary Figure 1. RNA editing found in the *psaA* transcript in strain TGD. The nucleotides after RNA editing are indicated above the genome sequence. The letters in red and blue are the results of base-conversion and base-insertion editing, respectively. The amino acid changes led by RNA editing are shaded in red or blue.

```

atgatctctatcctagtaaaacttcctgaagcatatagaccatttaacccaattgtagat
+1  M  I  S  I  L  V  K  L  P  E  A  Y  R  P  F  N  P  I  V  D

gtaatacctgttattcccgatTTTTTctactactagcattcgtatggcaagcgtcggta
+1  V  I  P  V  I  P  V  F  F  L  L  L  A  F  V  W  Q  A  S  V

agtttccgattagaaaaagttaaaataacgctaagtaaattTTTTTctttatattattat
+1  S  F  R  L  E  K  V  K  I  T  L  S  K  F  F  S  L  Y  Y  Y

tattatggagggttattttccaacttgcagctattttcgctggttgtaactgtaggaccagc
+1  Y  Y  G  G  Y  F  P  T  C  S  Y  F  R  C  C  N  C  R  T  S
+2      M  E  V  I  F  Q  L  A  A  I  F  A  V  V  T  V  G  P  A

agtaatcattcttttggcagcacaaaaagggaatcttttattttttttcttatgaatatt
+1  S  N  H  S  F  G  S  T  K  R  E  S  F  I  F  F  L  M  N  I
+2  V  I  I  L  L  A  A  Q  K  G  N  L  L  F  F  F  L  *

tcggataaccaagtcttttatagctctatacgttgcactccttacaggtataatagctctc
+1  S  D  N  Q  V  F  I  A  L  Y  V  A  L  L  T  G  I  I  A  L

cgtctaggcattaatttgtatgaaggtttaacattaatttttacatgctctta
+1  R  L  G  I  N  L  Y  E  G  L  T  L  I  F  T  C  S  *

```

Supplementary Figure 2. Organization of *psbK*, *psaM* and *ycf12* in the TGD plastid genome.

PsbK-PsaM fusion protein is encoded in +1 reading frame. The amino acid residues in red and purple correspond to the conserved domain for PsbK superfamily (PRK02553) and that for PsaM superfamily (CHL00190), respectively. The putative amino acid sequence of Ycf12 is most likely encoded in the same genome region but on a different frame (+2). The amino acid residues in blue correspond to the conserved domain for PSII_Ycf12 superfamily (CHL00184).

A

```

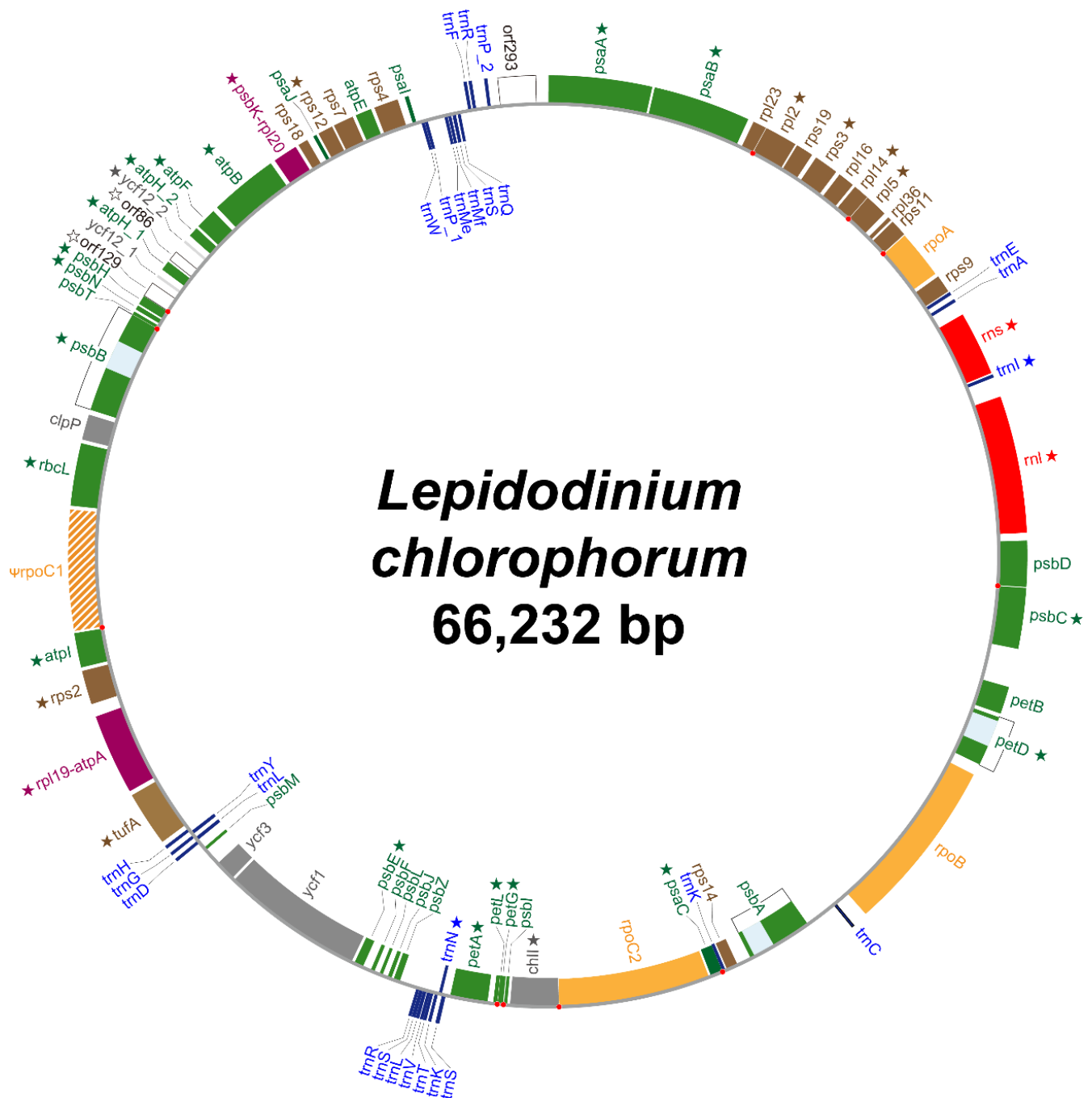
atgattcgtcgcagacatcatacttgggccaaagacaaactagtaactatTTTTTTTTTcaga
M I R R D I I L G P R Q T S N Y F F F R
aatctatTTTTTggaggaacaatctTTTtctcattgggtttccaaatTTtaggaaaaat
N L F F G G T I F F L I G F P N F R K N
ctattgcctaaagctatagaaattaaatttctaccccaaggtatcggtatatgcttttat
L L P K A M E I K F L P Q G I V I C F Y
ggtaggttggaataagattcagtttgatatgggggttattagaagatTTTtgttgattggg
G R L A M R F S L Y G V I R R F L L I G
agaggagttaatgagtataataaaaaaaagcaaaattcatattcttcggtggggatttc
R G V N E Y N K K K A
                                K K S K I H I L R W G F
ctaacaaaattcgacgaatagaatttcttatttgctttctgaacttgaatcgatcgggt
P N K I R R M E F S Y L L S E L E S I G
tggtaaatcaaaaacaattgcttcagcctattgaattgaaaatataTTTtctTTTgaaag
L V N Q K Q L L Q P I E L K M Y F L L K
atcgacgaaaaattatTTTTTtgaacctaataatagaaaatattagctcgctagaagaaa
D R R K I I F L K P N M E N I S S L E E
tagaacaatttcttctaacttggcgaagtttctacaaattcctttgaaagaagaattac
M E Q F S S N L A K F L Q I P L K E E L
ccttcgattTTTTtcttattttcattaagtga
P F D F F F I F I K -

```

B

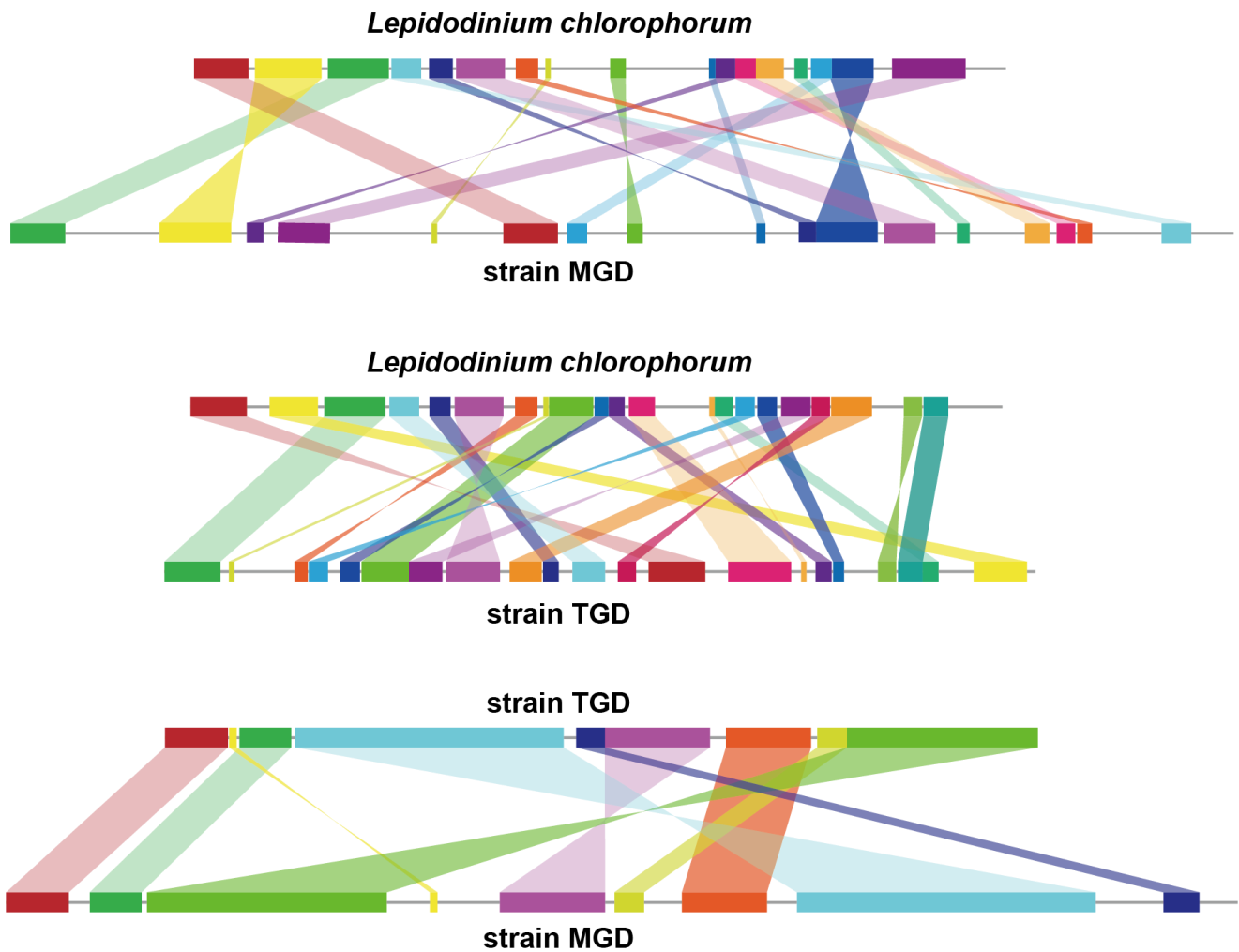
MGD ψ Ycf4	M-----IRRDIILGPRQTSNYFFFRNLFFGGTIFFLIGFP-NFRKNLLP-
<i>P. minor</i> Ycf4	MMSKENFQTNNTIDNIRRDLVIGSRRFSNYWWACVLSFGGIGFLLTGISSKVQTNLLPF
MGD ψ Ycf4	-KAMEIKFLPQGIVMCFYGRLAMRFSLYGVIRRFLIGRGVNEYNNKKKSKIHILRWGFPN
<i>P. minor</i> Ycf4	INYQDIQFFPQGLVMSFYGIIALVLVSLYLWACIAWSVGGGFNEFNKKDGIVRIFRWGFPG
MGD ψ Ycf4	KIRRIEFSYLLSELESIGLVNQQLLQPIELKMYFLLKDRRKIIIFLKPNMENISSLEEME
<i>P. minor</i> Ycf4	KNRRIELVYSLNEIDCIK-VDLQEGLNPRRSIYLRLLKGRDILLTR--IGQPLTLEEIE
MGD ψ Ycf4	QFSSNLAKFLQIPLKEELPFDFFFIFIK
<i>P. minor</i> Ycf4	KQAADLARFLQVGLEG-----IA

Supplementary Figure 3. Ycf4 has been pseudogenized in the MGD plastid genome. (A) Ycf4 amino acid sequence is not encoded by a single continuous reading frame. The N-terminus (red) is encoded in +1 reading frame, while the C-terminus (blue) is encoded in +2 reading frame. We found no *ycf4* transcript in the RNA-seq data. (B) Similarity between the Ycf4 amino acid sequences between strain MGD and *P. minor*. The identical amino acid residues between the two sequences are shaded.



Supplementary Figure 4. Circular map of the plastid genome of *Lepidodinium chlorophorum*.

Open reading frames (ORFs) encoding proteins involved in photosynthesis, translation, transcription, and other function are colored in green, brown, orange, and grey, respectively. Functionally unassigned ORFs are shown in white. Fused ORFs are highlighted in purple. *rpoC1* is considered as a pseudogene. Ribosomal RNA genes and transfer RNA genes are colored in red and blue, respectively. The ORFs/genes, of which transcripts received RNA editing, are marked by stars. Red dots indicate the overlap of two neighboring ORFs/genes.



Supplementary Figure 6. Pair-wise comparison of the organization of the peDinoflagellate plastid genomes. Top, synteny between the *Lepdodinium chlorophorum* and MGD plastid genomes. Middle, synteny between the *L. chlorophorum* and TGD plastid genomes. Bottom, synteny between strains TGD and MGD.

Supplementary table 1. Number of the RNA-seq short reads aligned with each of the peDinoflagellate plastid genomes.

	<i>Lepidodinium chlorophorum</i>	Strain TGD	Strain MGD	<i>Pedinomonas minor</i>
Number of the read pair subjected to the mapping ($\times 10^6$)	190	202	138	8
Number of the read pair aligned with the plastid genome ($\times 10^3$)	95	385	386	373

Supplementary table 2. ORF/gene fusion and overlapping found in the three green peDinoflagellate plastid genomes.

	<i>Lepidodinium chlorophorum</i>	Strain MGD	Strain TGD	<i>Pedinomonas minor</i>
Shared at least two genomes	<i>psbD/psbC</i> <i>petL/petG</i> <i>rpl14/rpl5</i> <i>rpl23/rpl2</i>	<i>psbD/psbC</i> <i>petL/petG</i> <i>rpl14/rpl5-rps8</i> <i>rpl23/rpl2</i> <i>tufA/rpl19</i> <i>psbI/psaI</i> <i>trnW/rps18</i> <i>rpl5-rps8/rpl36</i> <i>rpl36/rpl11</i> <i>rpoA-rps9</i>	<i>psbD/psbC</i> <i>petL-petG</i> <i>tufA/rpl19</i> <i>psbI/psaI</i> <i>trnW/rps18</i> <i>rps8-rpl36</i> <i>rps8-rpl36/rps11</i> <i>rpoA/rps9</i>	<i>psbD/psbC</i>
Not shared	<i>rps11/rpoA</i> <i>psbB/psbT</i> <i>petG/psbI</i> <i>chlI/rpoC2</i> <i>atpI/rpoC1</i> <i>psbH/orf129</i> <i>psaC/trnK</i> <i>rpl19-atpA</i> <i>psbK-rpl20</i>	<i>rpl16/rpl14</i> <i>psbJ/psbZ</i> <i>rpoC1/rpoC2</i> <i>rpl5-rps8</i> <i>rps19-rps3</i> <i>rps12-rps7</i>	<i>psbK-psaM/secG</i> <i>psbK-psaM/ycf12</i> <i>psbL/psbF</i> <i>rpl20/rps18</i> <i>rpl2/rps19</i> <i>rps19/rps3</i> <i>psbK-psaM</i>	<i>cysA/trnG</i>

Note: ORF/gene names connected by a dash and a slash indicate ORF/gene fusion and overlapping, respectively.