

**Supplementary material to the paper:**

**A new lineage of non-photosynthetic green algae with extreme organellar genomes**

Tomáš Pánek<sup>1,2</sup>, Dovilė Barcytė<sup>1</sup>, Sebastian C. Treitli<sup>3</sup>, Kristína Záhonová<sup>1</sup>, Martin Sokol<sup>1</sup>, Tereza Ševčíková<sup>1</sup>, Eliška Zadrobílková<sup>2</sup>, Karin Jaške<sup>1</sup>, Naoji Yubuki<sup>2,4</sup>, Ivan Čepička<sup>2</sup>, Marek Eliáš<sup>1,\*</sup>

<sup>1</sup>Department of Biology and Ecology, Faculty of Science, University of Ostrava, 701 00 Ostrava, Czech Republic

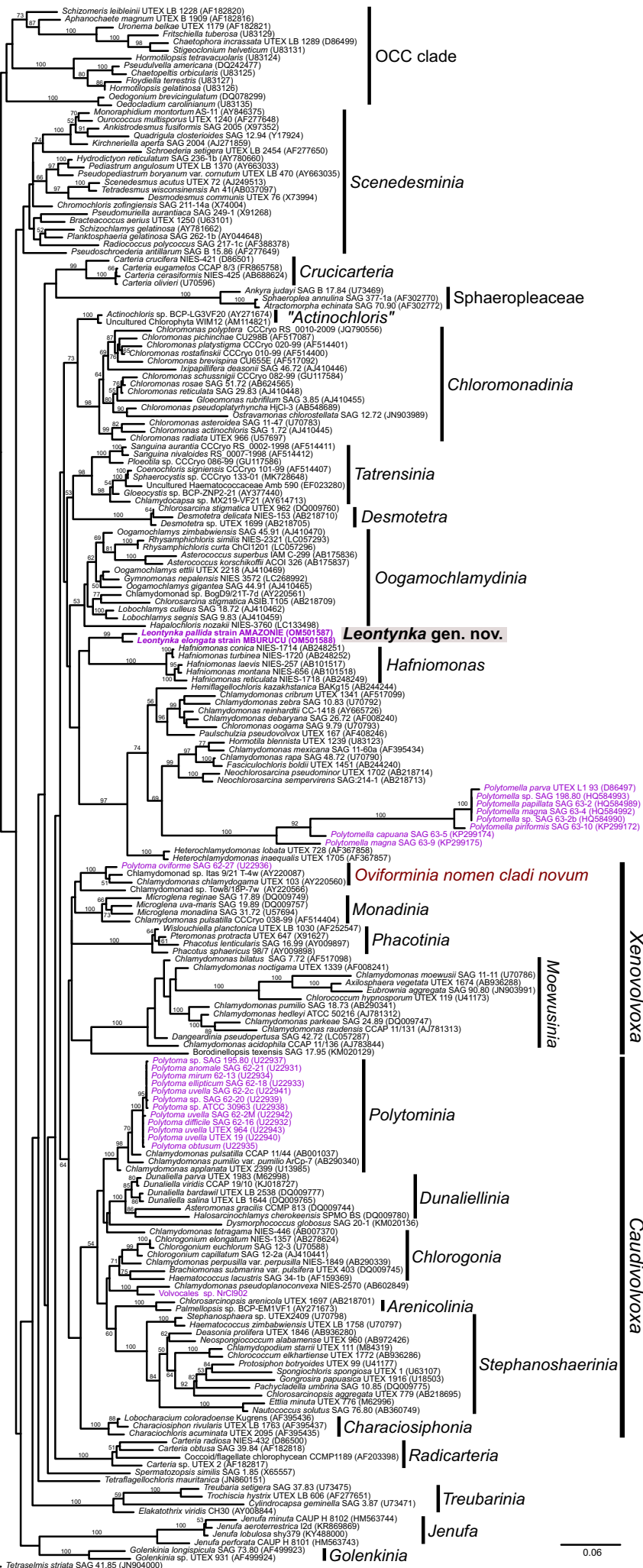
<sup>2</sup>Department of Zoology, Faculty of Science, Charles University, 128 43 Prague, Czech Republic

<sup>3</sup>Department of Parasitology, Faculty of Science, Charles University, BIOCEV, 252 42 Vestec, Czech Republic

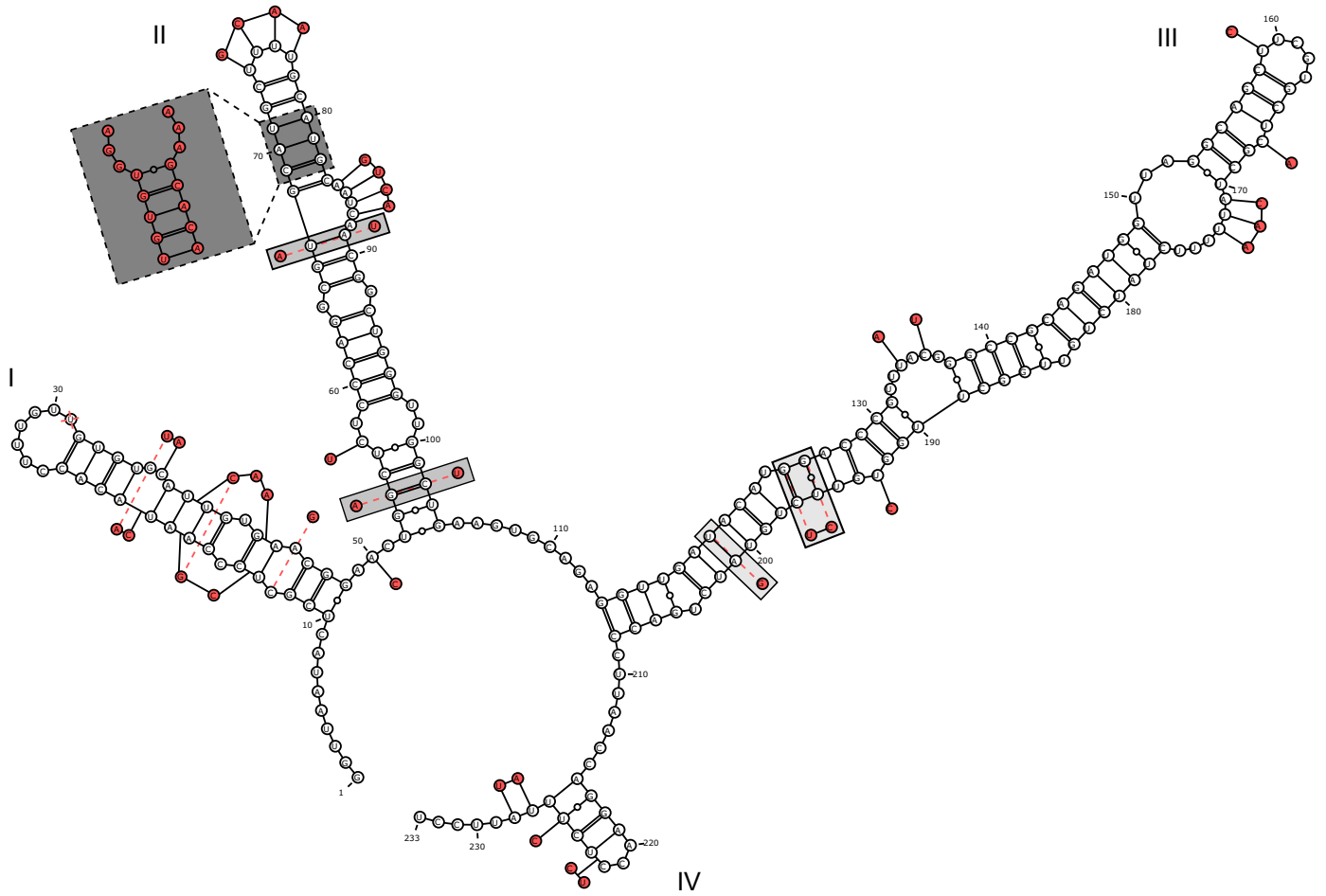
<sup>4</sup>Bioimaging Facility, University of British Columbia, V6T 1Z4 Vancouver, Canada

Correspondence: [marek.elias@osu.cz](mailto:marek.elias@osu.cz)

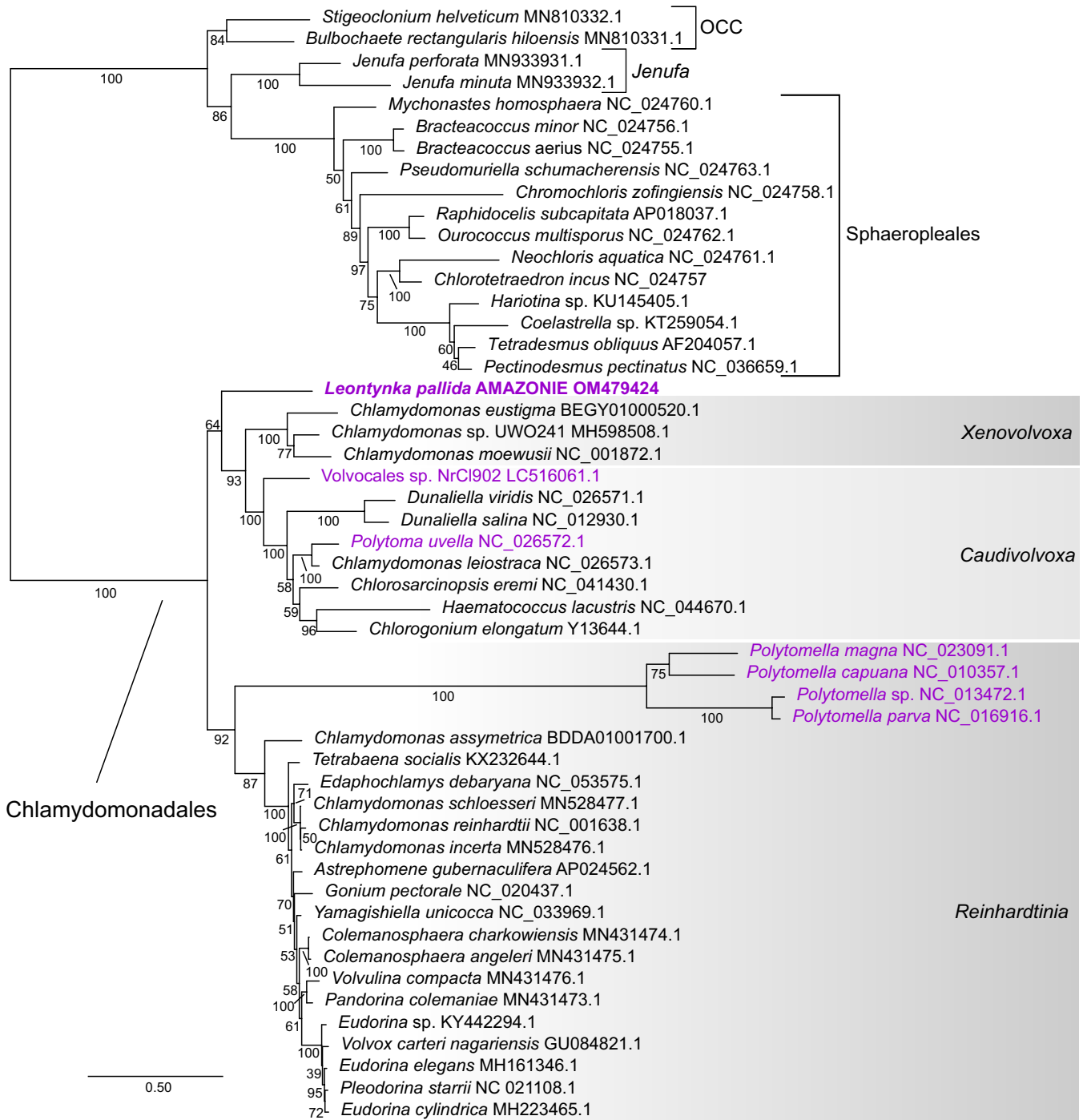
**This file contains supplementary Figs S1 to S9.**



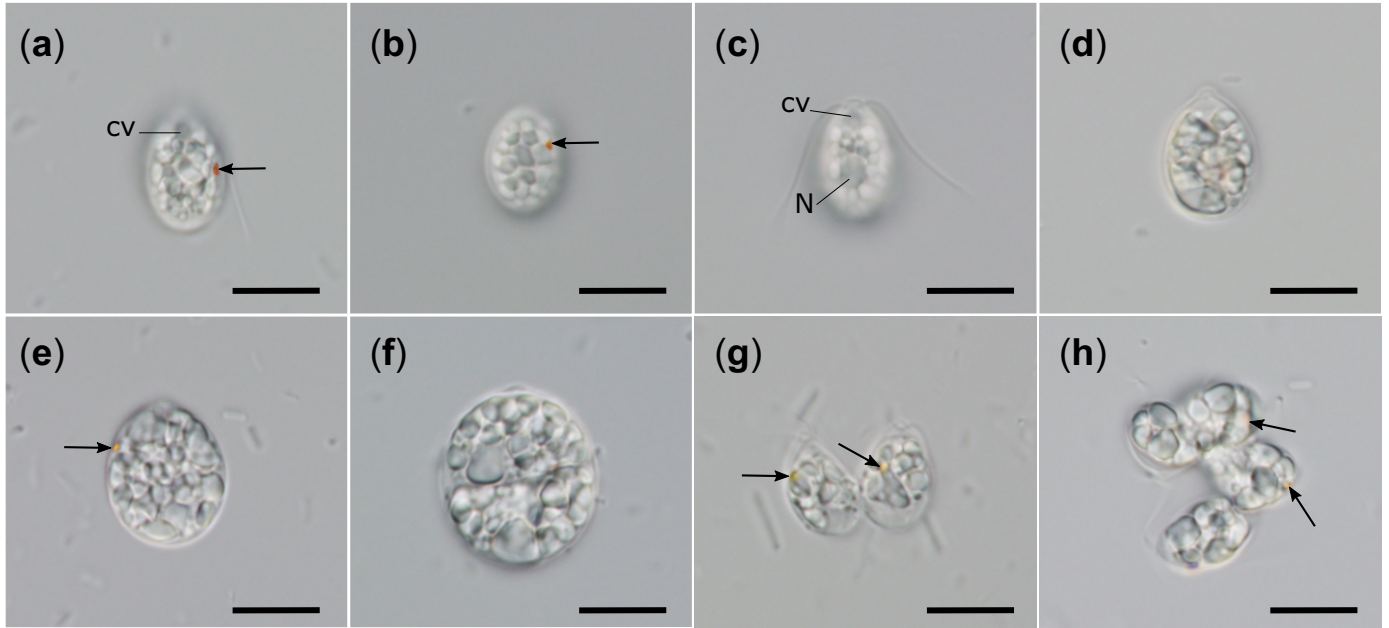
**Fig. S1** Maximum likelihood phylogenetic tree (IQ-TREE, TIM2+I+G4 substitution model) of 18S rRNA gene sequences from Chlorophyceae. The chlorodendrophycean *Tetraselmis striata* is used as an outgroup. Bootstrap support values are shown when  $\geq 50$ . Previously demarcated main clades [12] are annotated in the tree together with the newly designated clade "*Oviforminia*". Sequences from non-photosynthetic taxa are in colour. The new genus *Leontynka* (highlighted) forms a novel clade without apparent specific affinities to other particular lineages of Chlamydomonadales. The clade "*Actinochloris*" is labelled provisionally, as the *bona fide* *Actinochloris* genus belongs elsewhere.



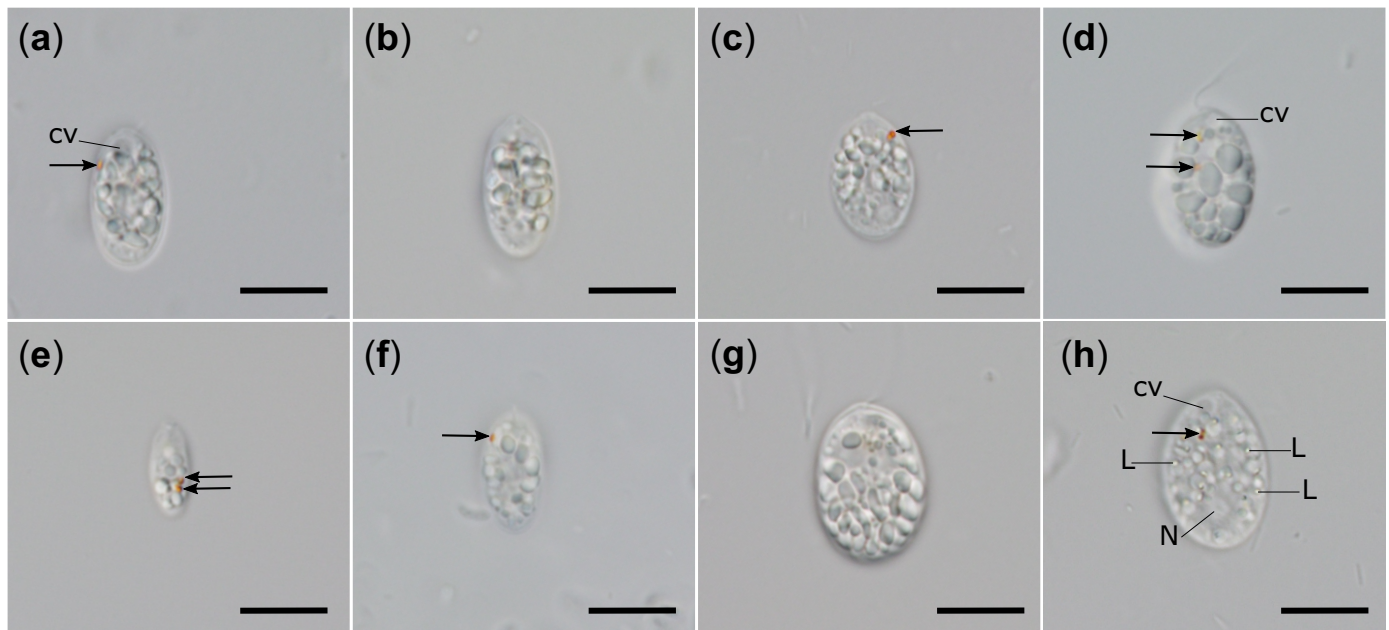
**Fig. S2** Predicted secondary structure of the ITS2 region of *Leontynka pallida*, with differences in the corresponding region of *Leontynka elongata* mapped onto it. Classical compensatory base changes in helix II are highlighted by a light grey background, and a region more substantially differing between the two species is highlighted by a dark grey background.



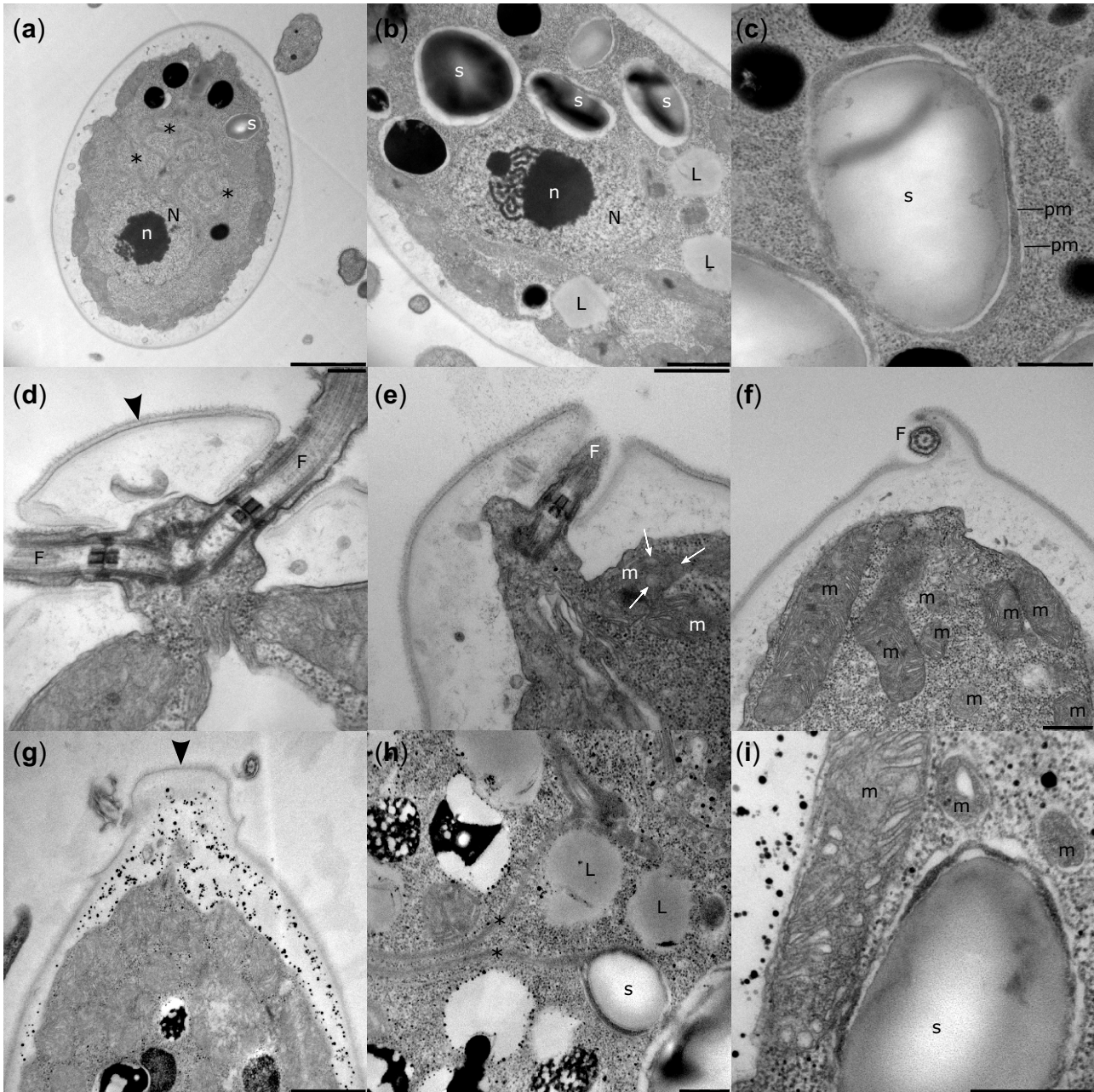
**Fig. S3** Maximum likelihood phylogenetic tree of Chlorophyceae, including *Leontynka pallida*, inferred from a concatenated set of seven conserved mitogenome-encoded proteins (2,608 amino acid positions). The tree topology was inferred using maximum likelihood analysis (IQ-TREE, LG+C60+F+G4 substitution model, 100 non-parametric bootstrap replicates). For easy display, the tree is arbitrarily rooted between Chlamydomonadales and other Chlorophyceae included in the analysis.



**Fig. S4** Light micrographs of *Leontynka pallida*. Scale bars = 10  $\mu\text{m}$ .  
 Abbreviations: arrows – eyespot; cv – contractile vacuole; N – nucleus.



**Fig. S5** Light micrographs of *Leontynka elongata*. Scale bars = 10  $\mu\text{m}$ .  
 Abbreviations: arrows – eyespot; cv – contractile vacuole; L – lipid droplet; N – nucleus.



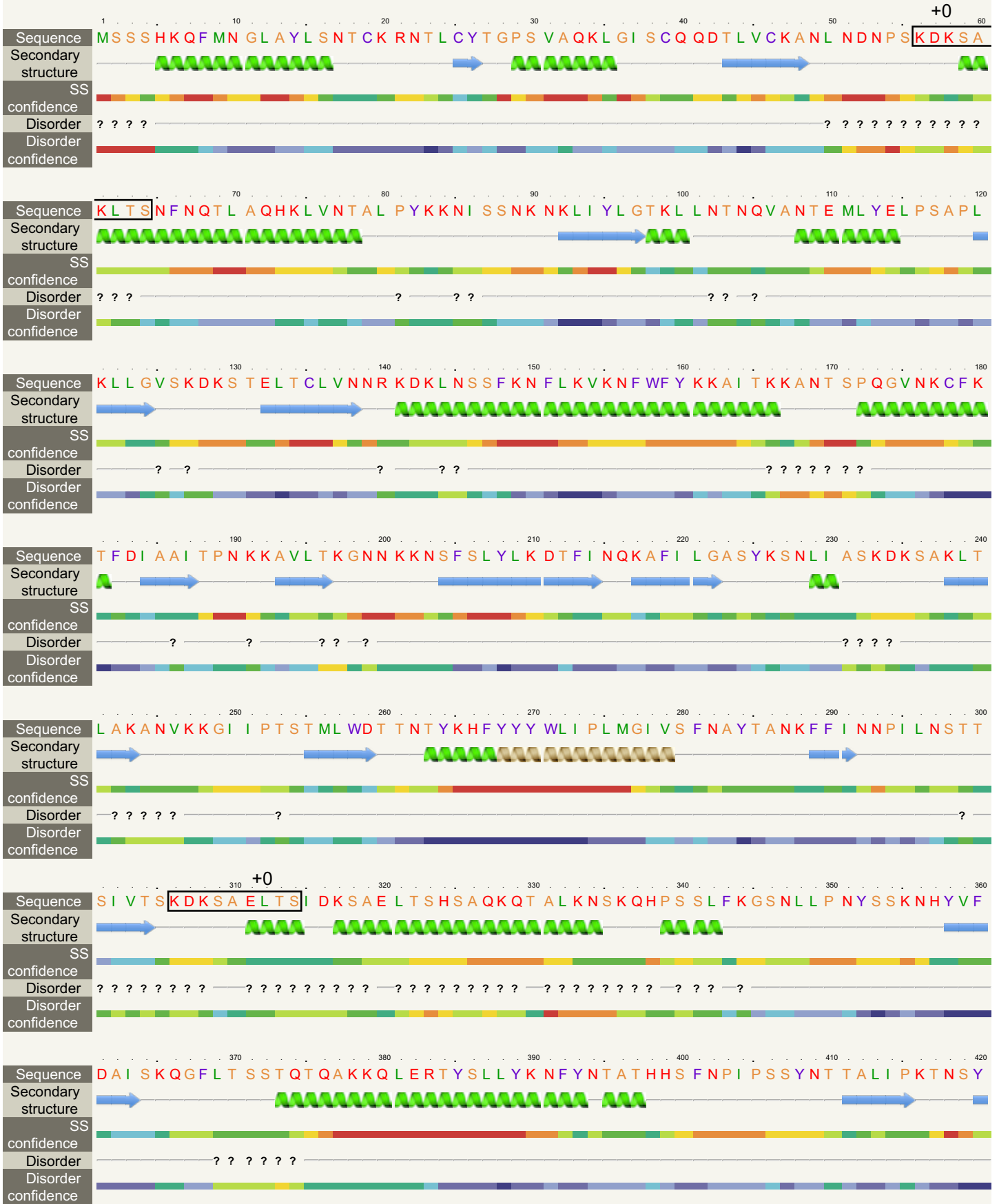
**Fig. S6** Ultrastructure of *Leontynka elongata* (a–f) and *Leontynka pallida* (g–i). (a) Cell with a slightly posterior nucleus and highly convoluted leucoplast. (b) Cell with a central nucleus and multiple starch blocks. (c) Two membranes surround the plastid. (d) Prominent keel-shaped papilla and two flagella. (e) Cross section through mitochondria with discoidal cristae. (f) Longitudinal section through mitochondria with discoidal cristae. (g) Ovoid cell with a keel-shaped papilla. (h) Presence of lipid droplets in an older cell. (i) Mitochondria containing putative tubulo-vesicular cristae (longitudinal section through the organelle). Abbreviations: F – flagellum; L – lipid droplet; m – mitochondrion; N – nucleus; n – nucleolus; pm – plastid membrane; s – starch. Asterisks mark “bridges” between plastid compartments; black arrowheads indicate papillae; white arrows indicate discoidal cristae. Scale bars: a = 2  $\mu\text{m}$ ; b, g = 1  $\mu\text{m}$ ; c, f, h, i = 0.5  $\mu\text{m}$ ; d, e = 0.2  $\mu\text{m}$ .

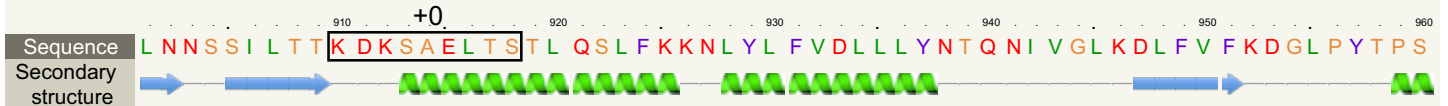
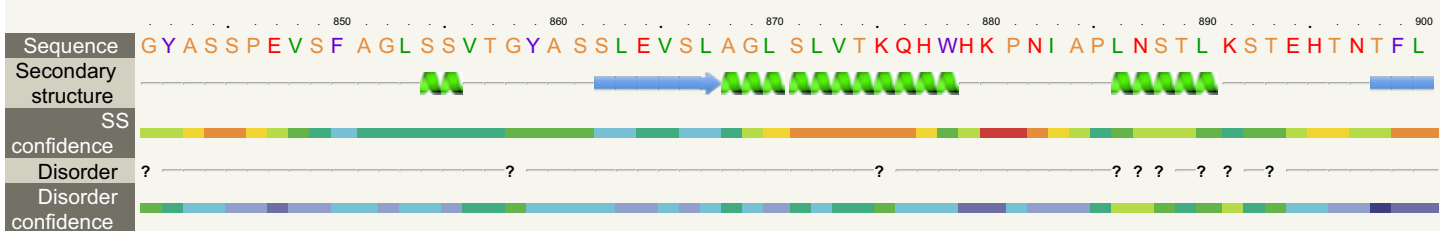
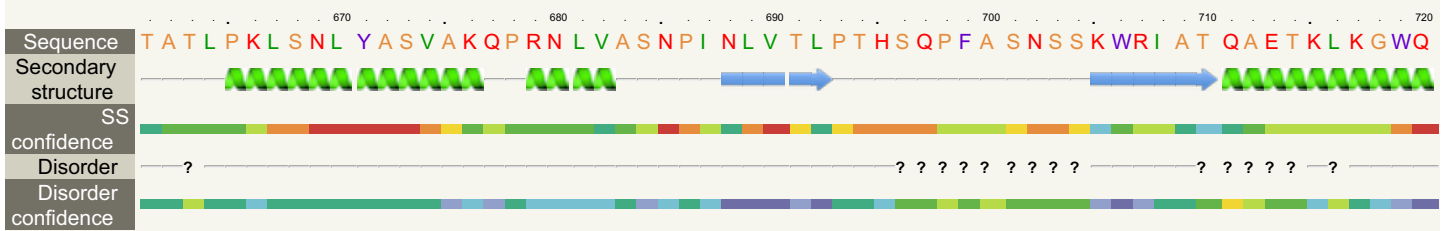
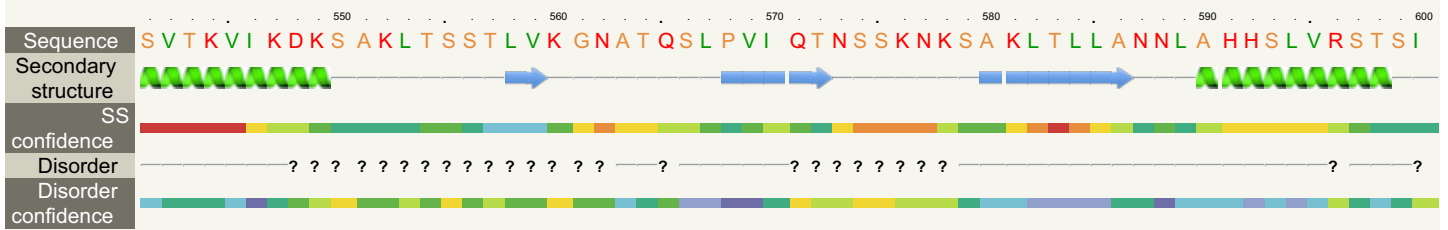
# Phyre2

Date Fri Nov 13 17:28:33 GMT 2020

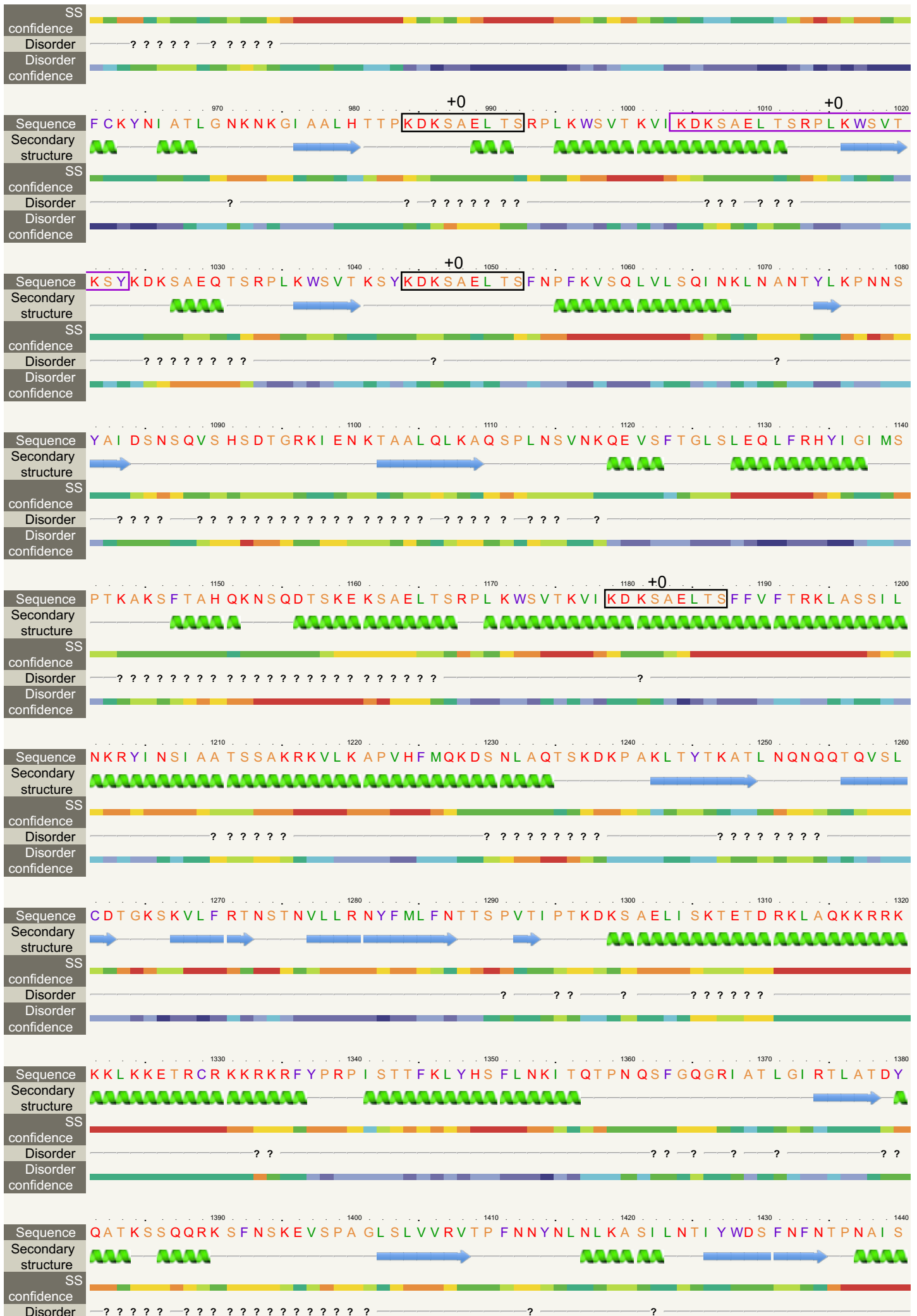
Confidence Key  
 High(9) Low (0)  
 ? Disordered ( 30%)  
 Alpha helix ( 54%)  
 Beta strand ( 11%)  
 TM helix ( 2%)

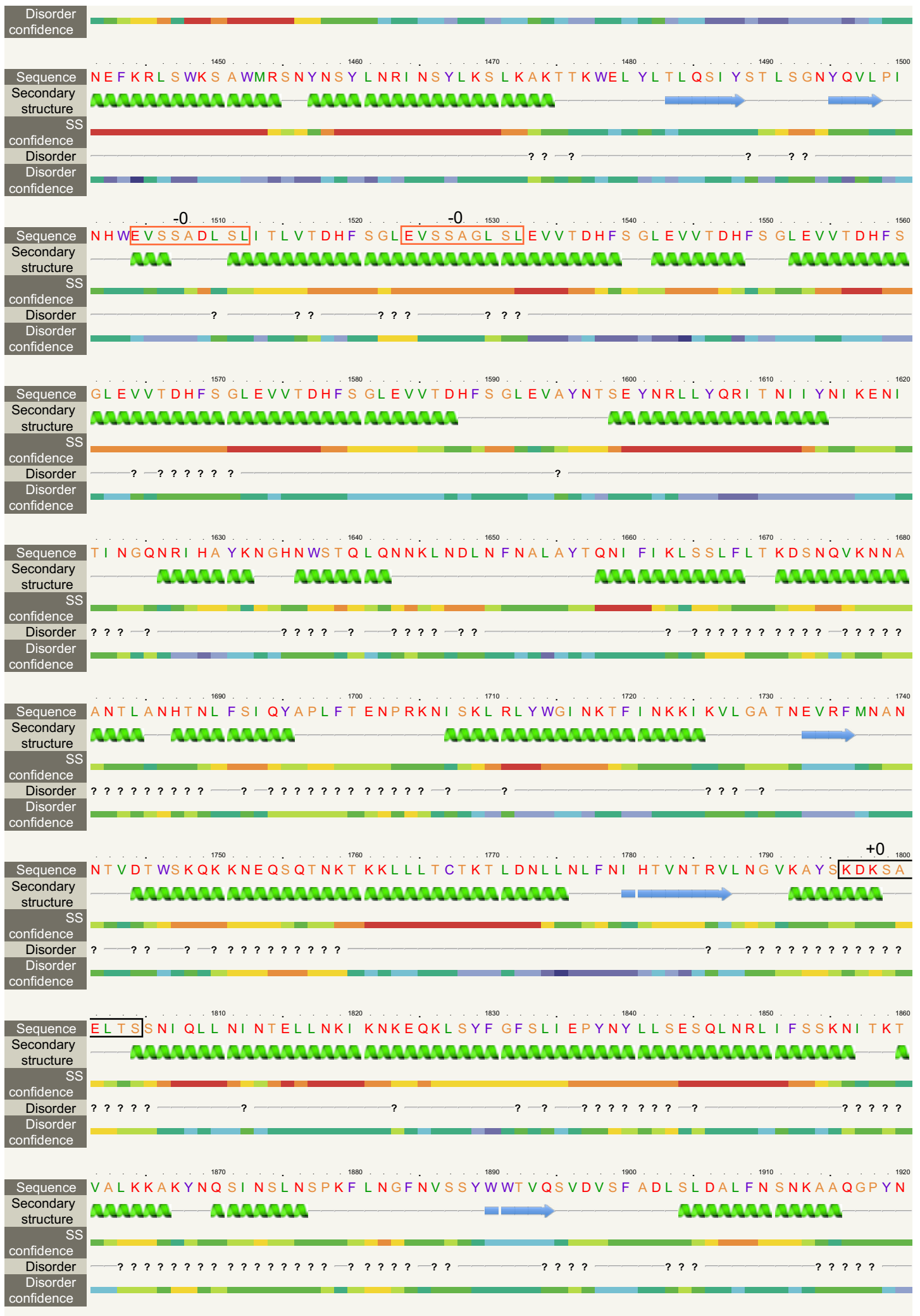
Secondary structure and disorder prediction **FtsH - *Leontynka pallida***

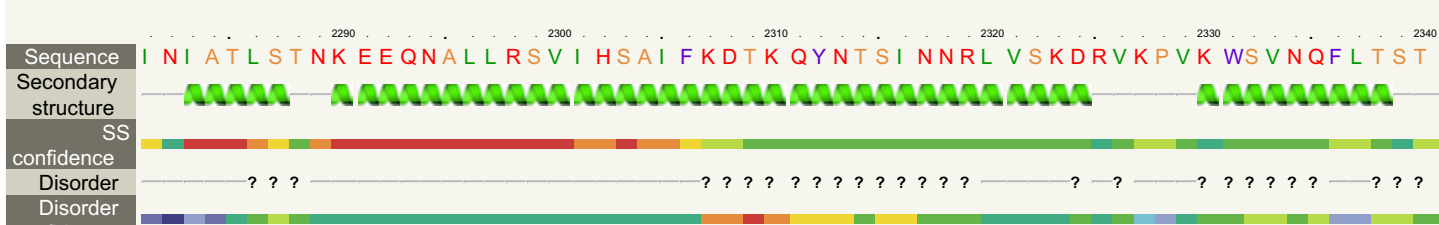
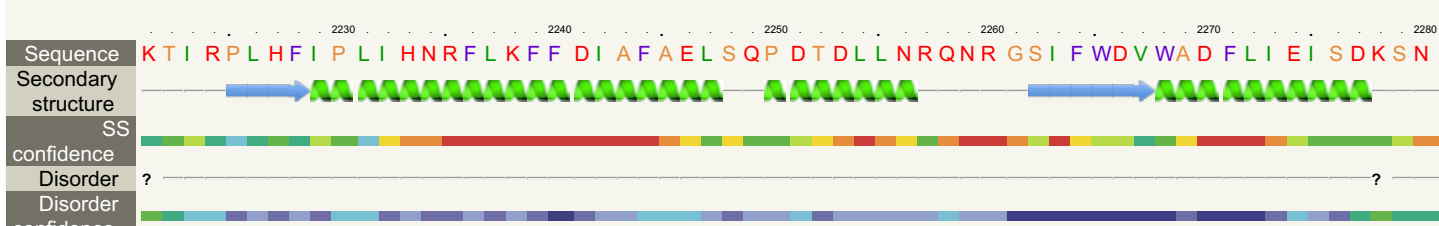
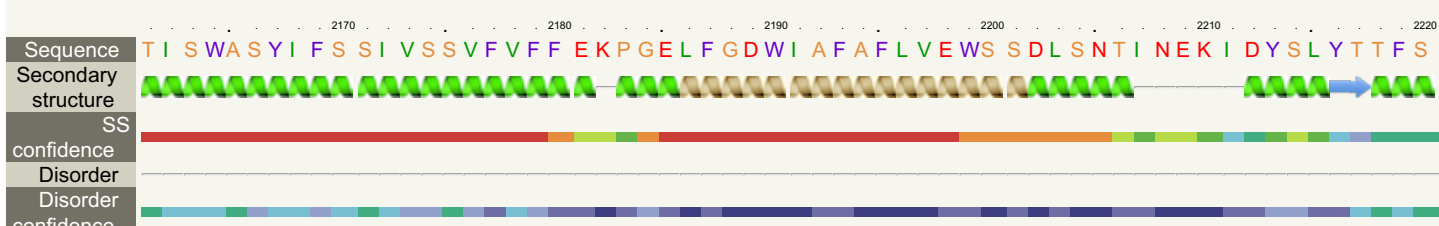
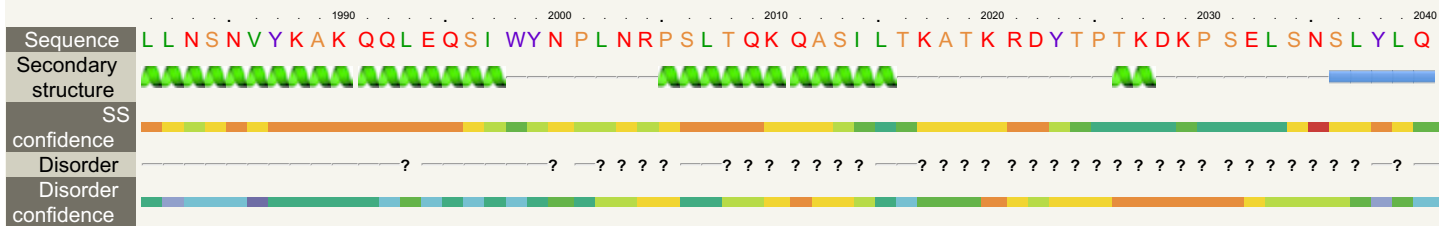
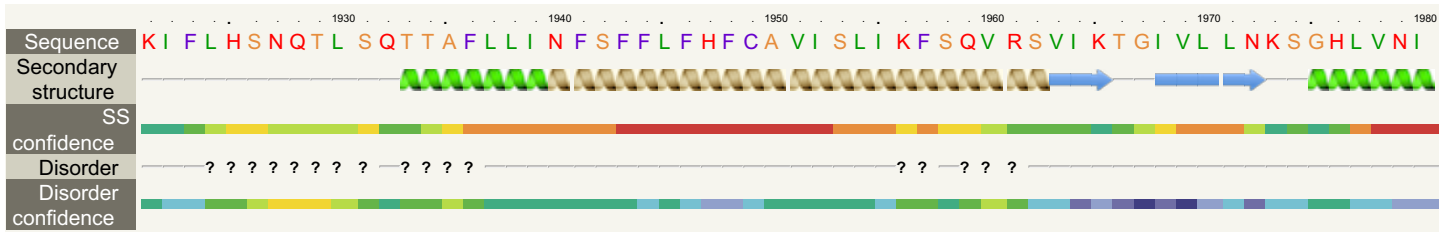


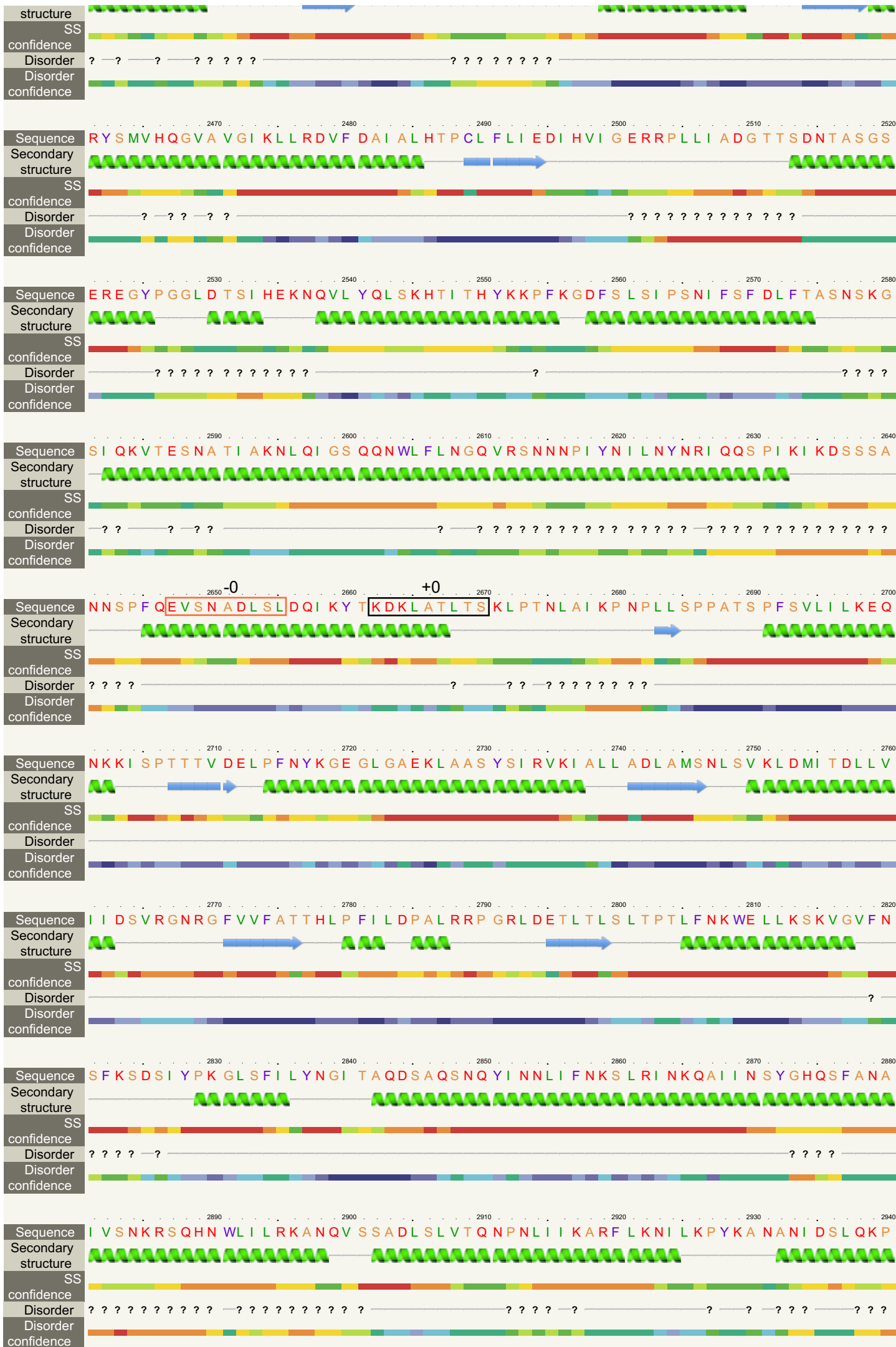


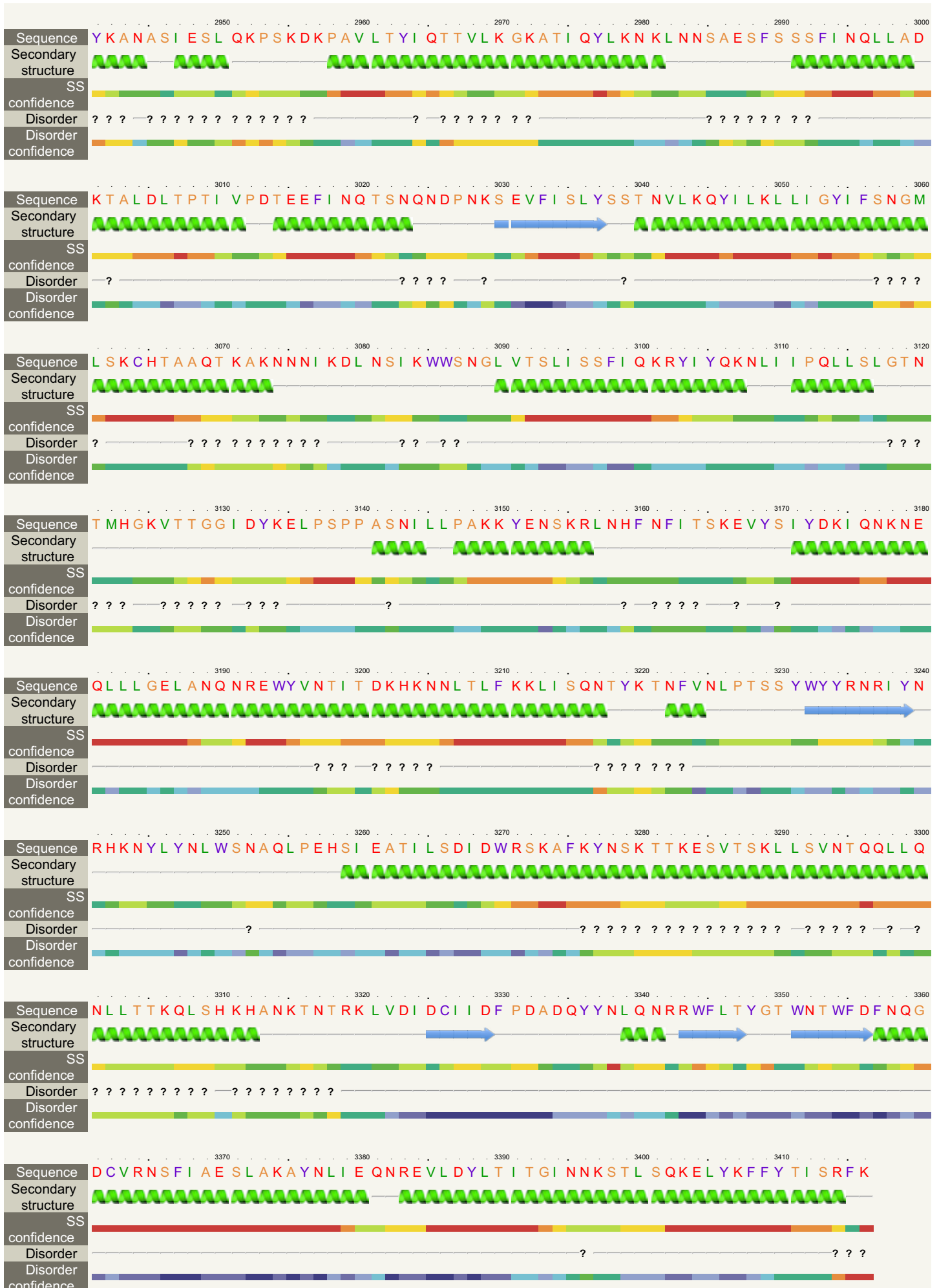












**Fig. S7** Occurrence of the “variant 8” repeat (see Fig. 4) in the FtsH protein of *Leontynka pallida* mapped onto its predicted structure. Protein model was constructed *in silico* using Phyre2. The “variant 8” repeat in RF +0 and -0 as well as a larger repeat containing the “variant 8” repeat in RF +0 are highlighted in coloured boxes.

10 20 30 40 50 60 70 80 90 100  
realigned-end .....  
mtDNA\_final AACATGTTCTTTGGCTCGTCCGAGTCCGTCGGCTTCGCCCTCGCTCCGTCGGAGCCCCATGTTCTTCCGACTTTGTGACGTAGCCCAACATGTT

110 120 130 140 150 160 170 180 190 200  
realigned-end .....  
mtDNA\_final CTCTCAACATGTTCTTAAACATGTTCTCCTAACATGTTCCAGAGAACATGTTCTCTTGGCTCGTCCGAGTCGCTCGGCTTCGCCCTCGCTCCTCG

210 220 230 240 250 260 270 280 290 300  
realigned-end .....  
mtDNA\_final TCCGAGCCCCATGTTCTTGCACTTTGTGAGATGCTTTCCCAAGCCCCAACATGTTCTCTTAAACATGTTCTCTTAAACATGTTCTCTTGGCTCGTC

310 320 330 340 350 360 370 380 390 400  
realigned-end .....  
mtDNA\_final CGAGTCGCTCGGCTTCGCCCTCGCTCCTCGTCCGAGCCCCATGGAATGCTTACAGGCCAGGACTGCCCTTCGGGGCAGATACTACATGTTCTCTTG

410 420 430 440 450 460 470 480 490 500  
realigned-end .....  
mtDNA\_final TGGCTCGTCCGAGTCGCTCGGCTTCGCCCTCGCTCCTCGTACGAGCCAGGTCATGTCCTAACCCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGA

510 520 530 540 550 560 570 580 590 600  
realigned-end .....  
mtDNA\_final ACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGA

610 620 630 640 650 660 670 680 690 700  
realigned-end .....  
mtDNA\_final CGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGA

710 720 730 740 750 760 770 780 790 800  
realigned-end .....  
mtDNA\_final CCCCGCCGTC AACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGA

810 820 830 840 850 860 870 880 890 900  
realigned-end .....  
mtDNA\_final TAGTTCACATGAGTCTGTAAGTATTATGATTCTGCTAAGTCTTCCCGAACATGTTCTCTTGTGGCTCGTCCGAGTCGCTCGGCTTCGCCCTCGCTC

910 920 930 940 950 960 970 980 990 1000  
realigned-end .....  
mtDNA\_final CTGTCCTTCGCCCTCATGTTCTTGTGGCTCGTCCGAGTCGCTCGGCTTCGCCCTCGCTCCTCGCTCCGCTCATGTTCTCTTAAACATGTTCTCTTGT

1010 1020 1030 1040 1050 1060 1070 1080 1090 1100  
realigned-end .....  
mtDNA\_final GGCTCGCCGAGTCGCTCGGCTTCGCCCTCGCTCCTCGCTCATGTTCTCTTAAACATGTTCTCTTGTGGCTCGTCCGAGTCGCTCGGCTTCGCC

1110 1120 1130 1140 1150 1160 1170 1180 1190 1200  
realigned-end .....  
mtDNA\_final TCGCTCCTCGTACGAGCCAGGTCATGTCCTTAAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAAC

1210 1220 1230 1240 1250 1260 1270 1280 1290 1300  
realigned-end .....  
mtDNA\_final CGTTCGCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGA

1310 1320 1330 1340 1350 1360 1370 1380 1390 1400  
realigned-end .....  
mtDNA\_final CCCCCGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCG

1410 1420 1430 1440 1450 1460 1470 1480 1490 1500  
realigned-end .....  
mtDNA\_final TGCTAAGTATTATGATATTCCGATGCTTTCCCAAGCCCCAACATGTTCTCTTGTGGCTCGTCCGAGTCGCTCGGCTTCGCCCTCGCTCCTCGCTC

1510 1520 1530 1540 1550 1560 1570 1580 1590 1600  
realigned-end .....  
mtDNA\_final CCTCATGTTCTCTTAAACATGTTCTCTTGTGGCTCGTCCGAGTCGCTCGGCTTCGCCCTCGCTCCTCGCTCCGAGCCCCATGTTCTTGCACTTTGTGACG

1610 1620 1630 1640 1650 1660 1670 1680 1690 1700  
realigned-end .....  
mtDNA\_final TAGCCCAACATGTTCTATCAACATGTTCTCTTGTGGCTCGTCCGAGTCGCTCGGCTTCGCCCTCGCTCCTCGCTCCGCTCATGTTCTCTTGTGGCTCG

1710 1720 1730 1740 1750 1760 1770 1780 1790 1800  
realigned-end .....  
mtDNA\_final TCCGAGTCGCTCGGCTTCGCCCTCGCTCCTCGCTCCTCGCTCATGTTCTCTTAAACATGTTCTCTTGTGGCTCGTCCGAGTCGCTCGGCTTCGCCCT

1810 1820 1830 1840 1850 1860 1870 1880 1890 1900  
realigned-end .....  
mtDNA\_final CTGTCGCCGAGCCCCATGTTCTTGCACATGTTCTCTTGTGGCTCGTCCGAGTCGCTCGGCTTCGCCCTCGCTCCTCGCTCCGCTCATGTTCTCTTAA

1910 1920 1930 1940 1950 1960 1970 1980 1990 2000  
realigned-end .....  
mtDNA\_final CATGTTCTCTTGTGGCTCGTCCGAGTCGCTCGGCTTCGCCCTCGCTCCTCGCTCATGTTCTCTTAAACATGTTCTCTTGTGGCTCGTCCGAGTC

2010 2020 2030 2040 2050 2060 2070 2080 2090 2100  
realigned-end .....  
mtDNA\_final .....



```

      4110      4120      4130      4140      4150      4160      4170      4180      4190      4200
realigned-end  CTTGGGTTGTGCCCC - CTTACCCGTTCCCGTAACCCGTTCCCGTAACCCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGG
mtDNA_final    CTTGGGTTGTGCCCCCGTTACCCGTTCCCGTAACCCGTTCCCGTAACCCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGG

      4210      4220      4230      4240      4250      4260      4270      4280      4290      4300
realigned-end  AACCGTTCCCGTAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGTAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCG
mtDNA_final    AACCGTTCCCGTAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGTAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCG

      4310      4320      4330      4340      4350      4360      4370      4380      4390      4400
realigned-end  CCAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGATCCAAACCCCGCCCGTCAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAGG
mtDNA_final    CCAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGATCCAAACCCCGCCCGTCAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAGG

      4410      4420      4430      4440      4450      4460      4470      4480      4490      4500
realigned-end  GGGCCATGAGTCCGAGATTGGTTGGCAAGACATGATTCGCAACCCCGTAGTTCCACATGATTCGCAACCCCGTAGTTCCACATGATTCGCAACCC
mtDNA_final    GGGCCATGAGTCCGAGATTGGTTGGCAAGACATGATTCGCAACCCCGTAGTTCCACATGATTCGCAACCCCGTAGTTCCACATGATTCGCAACCC

      4510      4520      4530      4540      4550      4560      4570      4580      4590      4600
realigned-end  GCTAGTCCCA--TGATTCGCAAGTATTGTTACGCTAACTGTTCTGAGCTGCCGGTAGCTCCACGTAATTCCGCTAAGTATTAATGATTATGCAC
mtDNA_final    GCTAGTCCCAATGATTCGCAAGTATTGTTACGCTAACTGTTCTGAGCTGCCGGTAGCTCCACGTAATTCCGCTAAGTATTAATGATTATGCAC

      4610      4620      4630      4640      4650      4660      4670      4680      4690      4700
realigned-end  CCCGCTAGTCCCAATGAGTCTGCTAAGTATTTATGAGCTGCTAAGTCTTGCCCAACATGTTCTCTTGGCTCGCCGAGTCGCTCGGCTCGCT
mtDNA_final    CCCGCTAGTCCCAATGAGTCTGCTAAGTATTTATGAGCTGCTAAGTCTTGCCCAACATGTTCTCTTGGCTCGCCGAGTCGCTCGGCTCGCT

      4710      4720      4730      4740      4750      4760      4770      4780      4790      4800
realigned-end  CGCTCCGCTGCTCCGCTAGCTCCCTGATACAAATGTTCTCTTGGCTCGTCCGAGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT
mtDNA_final    CGCTCCGCTGCTCCGCTAGCTCCCTGATACAAATGTTCTCTTGGCTCGTCCGAGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT

      4810      4820      4830      4840      4850      4860      4870      4880      4890      4900
realigned-end  TTCCTGCACTTTGTGACGTAGCCCAACATGTTCTATCAACATGTTCTCTTGGCTCGTCCGAGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCT
mtDNA_final    TTCCTGCACTTTGTGACGTAGCCCAACATGTTCTATCAACATGTTCTCTTGGCTCGTCCGAGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCT

      4910      4920      4930      4940      4950      4960      4970      4980      4990      5000
realigned-end  CATGTTCTCTTGGCTCGCCGAGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT
mtDNA_final    CATGTTCTCTTGGCTCGCCGAGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT

      5010      5020      5030      5040      5050      5060      5070      5080      5090      5100
realigned-end  GCTCGGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT
mtDNA_final    GCTCGGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT

      5110      5120      5130      5140      5150      5160      5170      5180      5190      5200
realigned-end  GTCCGAGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT
mtDNA_final    GTCCGAGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT

      5210      5220      5230      5240      5250      5260      5270      5280      5290      5300
realigned-end  CCACCTGATCCAAAGCGCTGATCTGGGCCATCCGGGAGCTTGGCTTTGCCCCCGTTACCCGTTCCCGTAACCCGTTCCCGTAACCCGTTCCCGGAA
mtDNA_final    CCACCTGATCCAAAGCGCTGATCTGGGCCATCCGGGAGCTTGGCTTTGCCCCCGTTACCCGTTCCCGTAACCCGTTCCCGTAACCCGTTCCCGGAA

      5310      5320      5330      5340      5350      5360      5370      5380      5390      5400
realigned-end  CCGTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCC - CTGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGA
mtDNA_final    CCGTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGTGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGGA

      5410      5420      5430      5440      5450      5460      5470      5480      5490      5500
realigned-end  ACCGTTCCCTCGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGTAACCGTTCCCGTAACCGTTCCCGTAACCGTTCCCGTAACCGTTCCCG
mtDNA_final    ACCGTTCCCTCGAACCGTTCCCGGAACCGTTCCCGGAACCGTTCCCGTAACCGTTCCCGTAACCGTTCCCGTAACCGTTCCCGTAACCGTTCCCG

      5510      5520      5530      5540      5550      5560      5570      5580      5590      5600
realigned-end  CTCGGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT
mtDNA_final    CTCGGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT

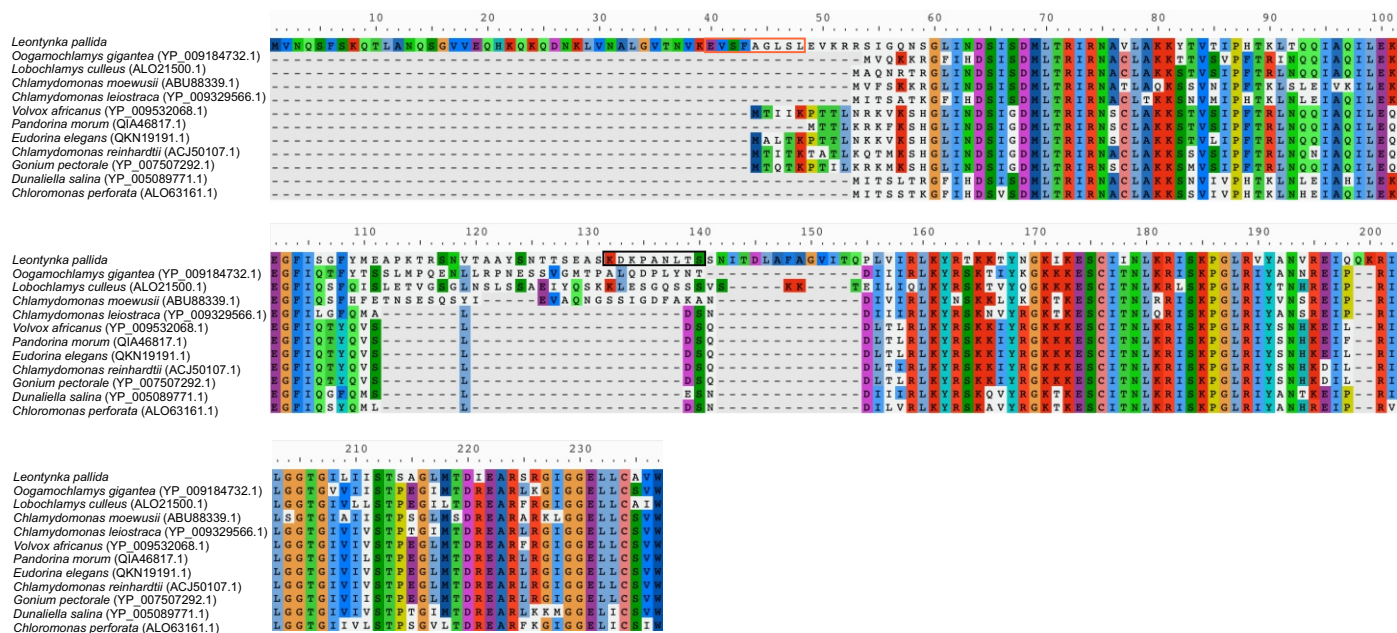
      5610      5620      5630      5640      5650      5660      5670      5680      5690      5700
realigned-end  ACGGTAAGAGGGCAACGGTAAGAGGGCAAACGGTAAGAGGGCAAACGGTAAGAGGGCAAACGGTAAGAGGGCAAACGGTAAGAGGGCAAACGGTAAGAGGGCAA
mtDNA_final    ACGGTAAGAGGGCAACGGTAAGAGGGCAAACGGTAAGAGGGCAAACGGTAAGAGGGCAAACGGTAAGAGGGCAAACGGTAAGAGGGCAAACGGTAAGAGGGCAA

      5710      5720      5730      5740      5750      5760      5770      5780      5790      5800
realigned-end  CAGAGACAACATGTTCTCTTGTGGCTCGTCCGAGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT
mtDNA_final    CAGAGACAACATGTTCTCTTGTGGCTCGTCCGAGTCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT

```

**Fig. S8** Alignment of the highly similar terminal regions of the originally assembled linear mitogenome contig. The sequence similarity of the termini is 97.7 % (along the alignment containing 5,771 nucleotide positions). The sequence marked here as “mtDNA\_final” represents the variant of the terminus that was retained in the final putatively circular-mapping full mitogenome sequence.





**Fig. S9** Occurrence of the “variant 8” repeat (translated in reading frame +0 as KDKPANLTS and -0 as KEVSFAGLSL; both boxed in colour) in a variable region of protein sequence of the ribosomal protein Rps8 from *Leontynka pallida* (full protein alignment together with representatives of other chlamydomonadalean algae).