Supplementary Materials

Fig. S1. An alignment of cox3 proteins of apicomplexans, dinoflagellates, and of *Chromera* and *Vitrella cox*3 ORFs. Cox3 sequences of dinoflagellates (*Durinskia, Hematodinium, Karlodinium, Kryptoperidinium, Oxyrrhis*), *Chromera*, and *Vitrella* are all highly derived. The alignment has been constructed in CLC Genomics Workbench v.6.5. and edited manually. Consensus sequence is shown at the bottom, and most conserved motifs are indicated with arrows. Genus abbreviations are the following: P., *Plasmodium*; E., *Eimeria*; B., *Babesia*; T., *Theileria*.

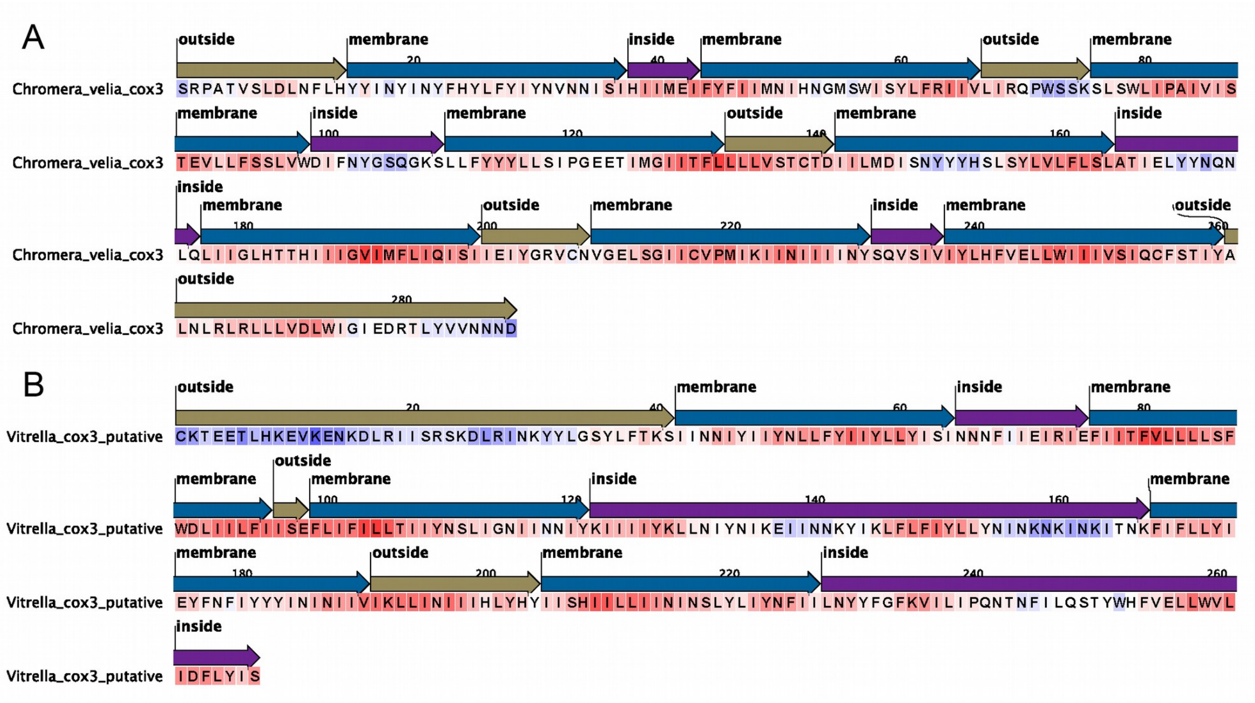


Fig. S2. *Chromera* and *Vitrella* putative cox3 proteins. Transmembrane domains predicted with CLC Genomics Workbench v.7.0.4 are shown with blue arrows, and Kyte-Doolittle hydrophobicity values are color-coded with red representing the highest values. Due to uncertainty in actual start positions, *Chromera cox*1-*cox*3 ORF sequence downstream of the conserved *cox*1 fragment is shown, along with the corresponding portion of the *Vitrella cox*3 ORF (see also supplementary fig. S5 showing a full *cox*3 ORF of *Vitrella*).

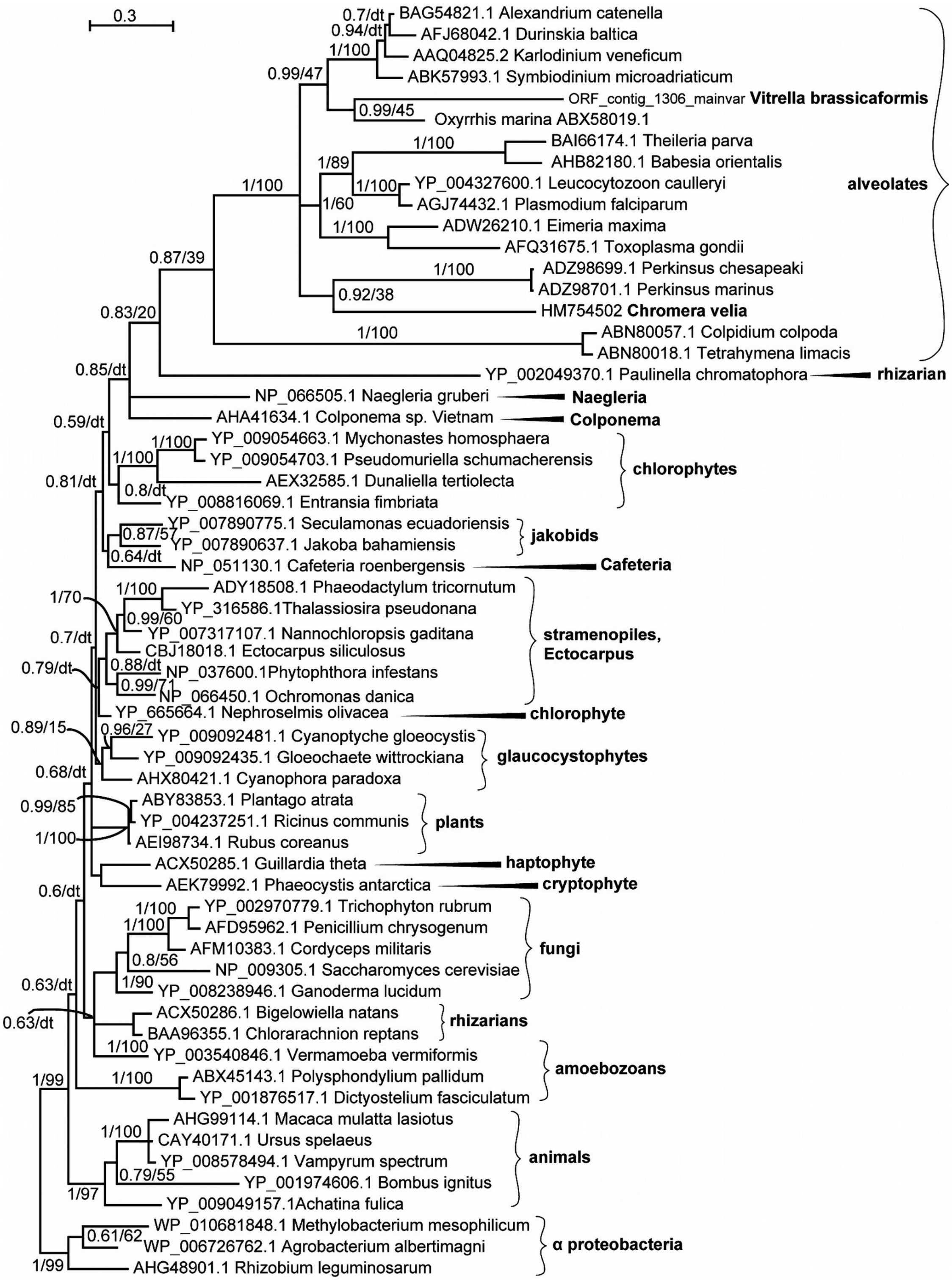


Fig. S3. Bayesian phylogenetic tree as inferred from cox1 amino acid sequences. Number above branches indicate Bayesian PP/ML bootstrap support.

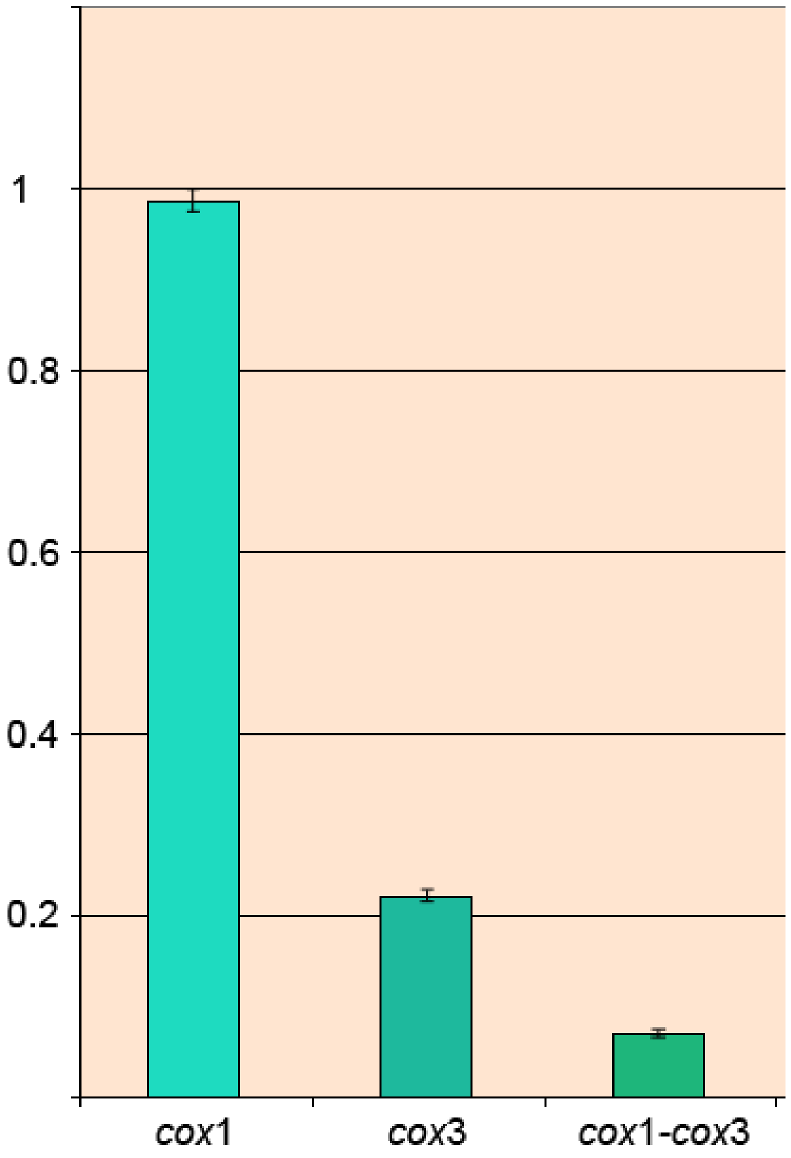


Fig. S4. Relative abundance of *cox*1, *cox*3 transcripts, and the junction region (*cox*1-*cox*3), as determined with quantitative PCR using the respective oligos. Standard deviation was calculated based on two PCR replicates. Relative concentration was determined based on a calibration curve made with a serially diluted sample.

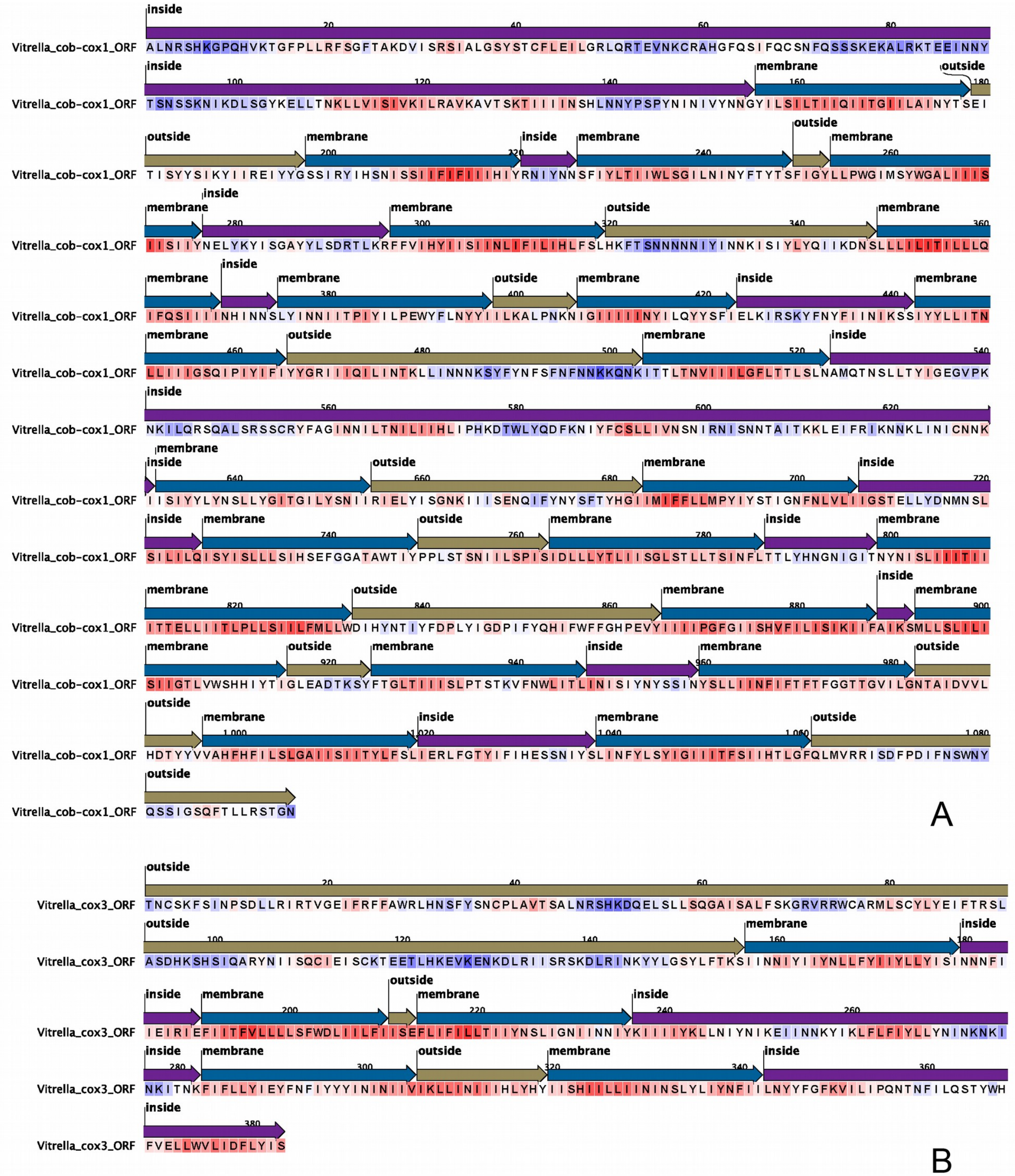


Fig. S5. Coding sequences from two *Vitrella* mitochondrial contigs (assembled from Illumina reads with Velvet): *cob*-*cox*1 fusion ORF, and *cox*3 ORF. Transmembrane domains predicted with CLC Genomics Workbench v.7.0.4 are shown with blue arrows, and Kyte-Doolittle hydrophobicity values are color-coded with red representing the highest values.

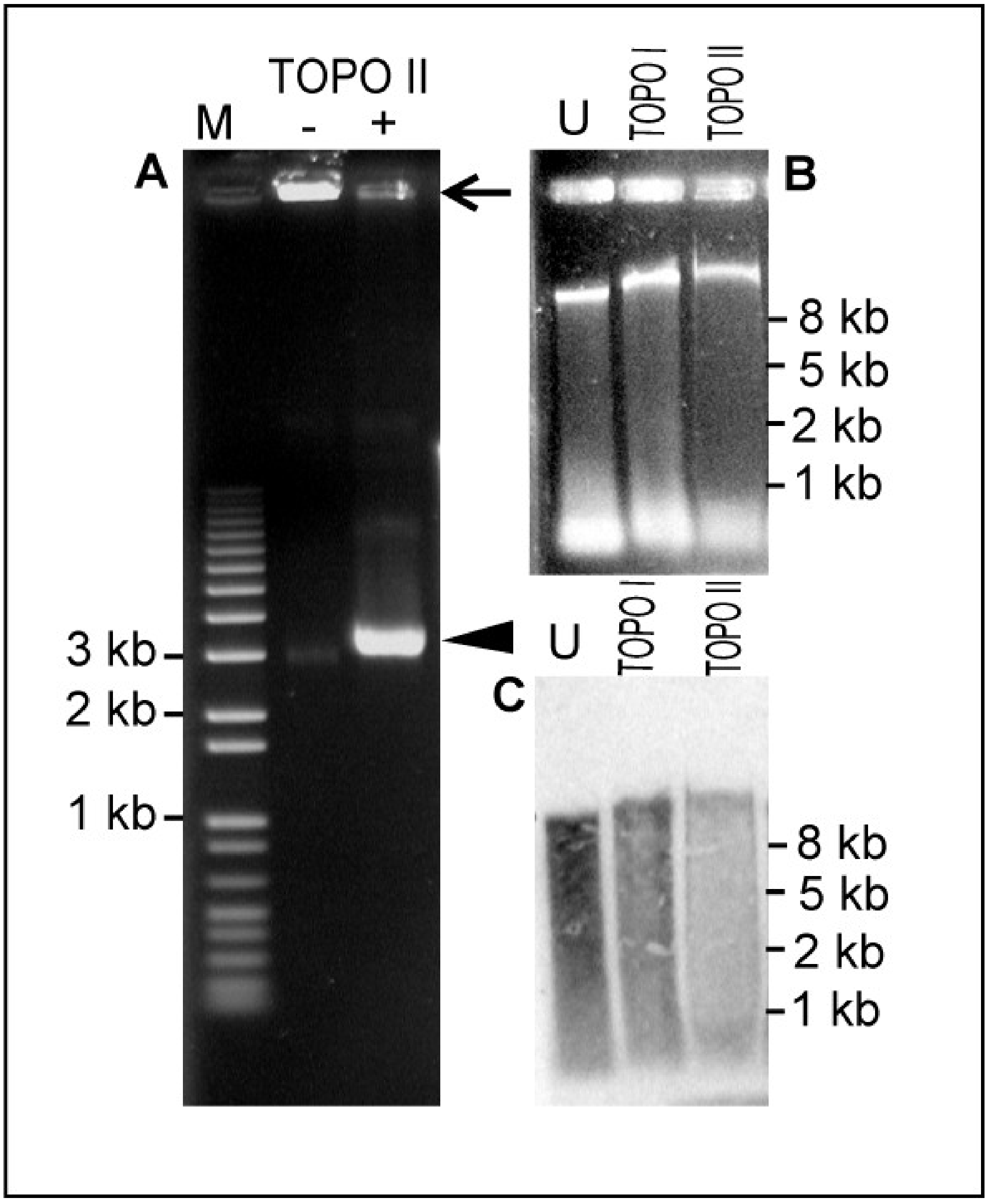
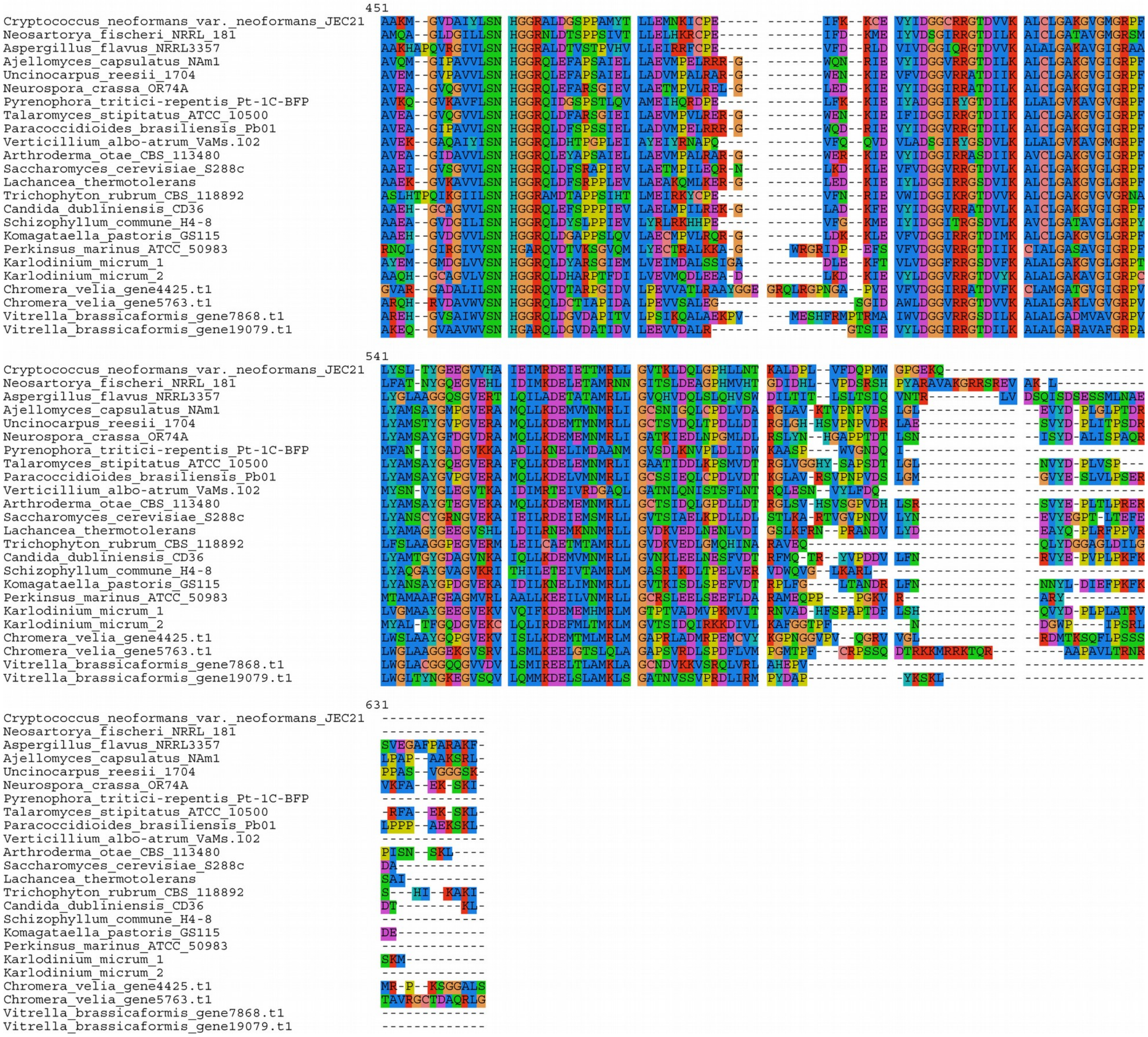


Fig. S6. Linear topology of mitochondrial DNA molecules in *C. velia*, hybridizing with the *cox*1 probe. (**A**) The mitochondrial DNA network composed of catenated DNA circles from *Crithidia fasciculata*, isolated as described previously (Jirků et al. 1995), was used as a control. While non-treated mitochondrial DNA network stays in the slot (-), upon topoisomerase II treatment, decatenated circular molecules enter the gel (+). (**B**) Ethidium bromide-stained gel of non-treated total *C. velia* DNA (U), and treated with topoisomerase I (Topo I) and topoisomerase II (Topo II). 1 μg DNA per lane was resolved in 0.75% agarose gel and stained with ethidium bromide after the run. The DNA was treated either with topoisomerase I or topoisomerase II following protocols described elsewhere (Lukeš et al. 1998). (**C**) Southern blot analysis of the gel in (B) hybridized with the *C. velia cox*1 gene as a probe labeled with [α-^{32}P]dATP. Hybridization was carried out as described elsewhere (Jirků et al. 1995). Molecular size markers are indicated.

```
Alignment: C:\Users\obornik\Desktop\MBE resubmission\presubmission\fig_S7_L-LDH_vob.fas
Seaview [blocks=15 fontsize=6 A4] on Fri Dec 19 14:17:35 2014

                                              1
Cryptococcus_neoformans_var._neoformans_JEC21 LI---PYSEVNKH-- ----------NKPDD CWVVIDGKVYDLTEF AES-HPGGSSPIYRA AGRDATAIFQPIHPP GTVEDGLDPKAMVGL
Neosartorya_fischeri_NRRL_181                 LL---STKEVSKH-- ----------KSPDD CWIVVDNKVWDMTDF LEE-HPGGSTIILKY AGRDATKAYSEVHSP GVIKSNLPPEKYKGI
Aspergillus_flavus_NRRL3357                   MI---EPSEVAKH-- ----------NSADS CWVVLYGKVYDVTNF LEN-HPGGSAAILAL AGKDATEEYDTIHPS GLLEEYLDPKACLGV
Ajellomyces_capsulatus_NAm1                   LL---TGPEIQKH-- ----------NSKNS CWIIVHGKAFDVTEF LPE-HPGGQDIILKY AGKDATDEFEPIHPP DTLDKYLDSSKHLGP
Uncinocarpus_reesii_1704                      PL---TGAEIAKH-- ----------NSAKS CWVIVHGKAYDVTDF LPE-HPGGQKIILKY AGKDATEEFDPIHPP DTLDKYLDPSKHLGP
Neurospora_crassa_OR74A                       PL---TGVEVAKH-- ----------NTQDD CWVIVHGKAYDITEF LPE-HPGGMKIILKY AGKDATEEFDPIHPP DTLEKYLAKDKHLGP
Pyrenophora_tritici-repentis_Pt-1C-BFP        --------------- --------------- -----------MAPF -------ENAIPFRS FLLSYFPTFTQLKMR T-SVLLFASC-----
Talaromyces_stipitatus_ATCC_10500             KL---TGQDVAQH-- ----------NNKNS CWVIVHGKAYDVTEF LPE-HPGGQKIILKY AGKDATEEFEPIHPP DTLDKYLDASKHLGE
Paracoccidioides_brasiliensis_Pb01            LL---TGPEIQKH-- ----------NSKDS CWVIVHGKAYDVTDF LPE-HPGGQKIILKY AGKDATEEFEPIHPS DTLDKYLDPSKHLGP
Verticillium_albo-atrum_VaMs.102              --------------- --------------- --------------- --------------- -----------MHSS LLLITQLAAV-----
Arthroderma_otae_CBS_113480                   KL---TGEEIAKH-- ----------SSPSS CWVVVHGHAYDVTDF LPEEHPGGQQIILKH AGKDATAEFDPVHPP DTLDKYLDRSKHLGP
Saccharomyces_cerevisiae_S288c                KI---SPAEVAKH-- ----------NKPDD CWVVINGYVYDLTRF LPN-HPGGQDVIKFN AGKDVTAIFEPLHAP NVIDKYIAPEKKLGP
Lachancea_thermotolerans                      GI---EVSEVMQH-- ----------NKPDD CWIVLNGEVYDVTSF LQM-HPGGAARIMEV AGNDATRKFYSIHSD STLEKMKEQLVYIGK
Trichophyton_rubrum_CBS_118892                MI---DAEEVAKH-- ----------NRTDS CWVVLYGKVYDVTNF LPN-HPGGANIILQL AGRDATEEYDPIHPT GTLEDSLPPEALVGR
Candida_dubliniensis_CD36                     VL---SLQEVSQH-- ----------NKKDD CWVIIHDKAYDLSDF IDE-HPGGSAIIMKY AGKDATKAFDPIHPG DTLTKYLQPKYHKGE
Schizophyllum_commune_H4-8                    --MNWTLEEVKKHNT RECPPRLAPAPLPRS RYTLLCPLSFDPSHH RQD-HPGGAAIILKY AGRDATAAYEPIHPA DALEKNLPASAHLGP
Komagataella_pastoris_GS115                   KV---TPEELAKH-- ----------NTGTD CWVAINGKVYDLTEF LPQ-HPGGRNVILKR AGKDASKVFNPIHPP DAISKFLPADKFVGV
Perkinsus_marinus_ATCC_50983                  --------------- --------------- --------------- --------------- --------------- ---------------
Karlodinium_micrum_1                          --------------- --------------- --------------- --------------- --------------- ---------------
Karlodinium_micrum_2                          --------------- --------------- --------------- --------------- --------------- ---------------
Chromera_velia_gene4425.t1                    MI---SFEEIAKH-- ----------NNRNS CWVVLHGVVYDMTKF LDE-HPGGAAIILKQ AGKDGTKAFDAIHPR DIV-KMLPPEAVLGK
Chromera_velia_gene5763.t1                    --------------- --------------- --------------- --------------- --------------- ---------------
Vitrella_brassicaformis_gene7868.t1           --------------- --------------- --------------- --------------- --------------- ---------------
Vitrella_brassicaformis_gene19079.t1          --------------- --------------- --------------- --------------- --------------- ---------------

                                              91
Cryptococcus_neoformans_var._neoformans_JEC21 VDPATLPKV------ --------------- V-----DKKKEDGEQ RRVDLAEII--GLPD FDAAAKANLTSKAWA YMSSGATDQY-----
Neosartorya_fischeri_NRRL_181                 LDTSAIDEE------ --------------W LKQPPSENPQVVLDN EKPPLHTLI--NSHD FELVASKTASKKTWA FYSSASTDLI-----
Aspergillus_flavus_NRRL3357                   LSASAPEAAEPTR-- ------------I-- -----SSQSASKEAS EETPLSSLL--NLAE IEQAAKRKLSPKGWA YYSSATDDSI-----
Ajellomyces_capsulatus_NAm1                   VAMDTV--------- ----------MHEDK AVDPEEEERQFR-IQ HMPPLEQCY--NLLD FEAVARRIMKKTAWA YYSSGADDEM-----
Uncinocarpus_reesii_1704                      VDMNTV--------- ----------MQEDK GVDPEEEARQER-VK RMPLLEQCY--NLMD FEAVARRVMKKTAWG YYSSGADDEIVGQSH
Neurospora_crassa_OR74A                       VDMSTV--------- ----------VTEKA EVDPEEQARLKR-IE EMPLLEQCY--NLLD FEAVAKRVMKKNAWA YYSSAADDEI-----
Pyrenophora_tritici-repentis_Pt-1C-BFP        -----AL-------- -----ASPFLDEPDT GID-AQ-FGAFVANK TLPPLNSLV--GLPD FQWVGRNFMNASSFT YYRNGAAGEW-----
Talaromyces_stipitatus_ATCC_10500             VDMSTV--------- ----------EHDEK VKDPLEIEREER-IK RMPALAQCY--NLLD FESVAREVMKNTAWA YYSSGADDEI-----
Paracoccidioides_brasiliensis_Pb01            VAMDTV--------- ----------MQEDK AVDPEEEERQLR-IQ RMPLLEQCY--NLLD FEAVARRILKKSAWA YYSSGADDEI-----
Verticillium_albo-atrum_VaMs.102              -----VL-------- -----GEPWINEVDT GLE-TYLLGTNYTEG TLPLLKDIR--ALPD FDWAARQHLNDQMYS FYRTGTAGEQ-----
Arthroderma_otae_CBS_113480                   VDMASM--------- ----------AQEKK AVDPEEEQRLER-IR RMPSLDQCY--NLMD FEAVASRVMKKTAWG YYSSGTEDEM-----
Saccharomyces_cerevisiae_S288c                LQGSMPPELVC---- ------------YAP GETKEDIARKEQLKS LLPPLDNII--NLYD FEYLASQTLTKQAWA YYSSGANDEV-----
Lachancea_thermotolerans                      LKGKFAE-------- --------------- TIS-EEDLRIRESRK NIPSLGLIF--NLSD FEAVAKEVLPKSTYA YFATGSSDEF-----
Trichophyton_rubrum_CBS_118892                VDEKSLAKLAPKP-- ------------A-- --------KAEKTTE GPPPLSTLL--NMDD FEAAASKQLTHKAWA YYFSGADDLI-----
Candida_dubliniensis_CD36                     VEKKKKKKKQPSTTT NNDADEYD--DDDDD PPTEEELKRRQL-VK NKPDISQIY--NLYD FEFVARHTMDPIGWA YYSSSADGEA-----
Schizophyllum_commune_H4-8                    VNTDAASTLEAAV-- ------------RNR RKTK-DEIRIEEAHK YKPALSHVL--SLRD MEDVAKKVLPHKAYA YYSSSTEDTI-----
Komagataella_pastoris_GS115                   LDGELP--------- -----------EEEE VLTDAEIERQER-IA NKPPLSSMF--NVYD FEYVAQNILDEAAWA YYSSAADDEI-----
Perkinsus_marinus_ATCC_50983                  --------------- --------------- --------------- -------ML--NAMD FKAVAEKKVSAEGWA YLYSAAGDEF-----
Karlodinium_micrum_1                          --------------- -----EEPLEPV--- EDEEEQVEEEKA-AD DTPDISQMV--NVWD FEVIAKRNVTKEAWA YLMSGADDEI-----
Karlodinium_micrum_2                          --------------- --------------- --------------- -----AHML--NLND FERAAAKSMDSQGYG YYASAANDEV-----
Chromera_velia_gene4425.t1                    ADMSTAPAPSAEA-- ------------ET- -A--PPSGPASA-AA EKPPIDHML--NIFD FESLAQQSMQKPGWD YYCSGADDEK-----
Chromera_velia_gene5763.t1                    --------------- --------------- --------------- --------------- --------------- ---------------
Vitrella_brassicaformis_gene7868.t1           --------------- --------------- ------------MAV SANGGFHPI--NLDD YEATAKERLTTQAHG YFSSGANDQW-----
Vitrella_brassicaformis_gene19079.t1          --------------- --------------- -MSM-------SRSK ASAEESVPV--NLDD FERIADKRLTKAVHG YYSSGANDEV-----

                                             181
Cryptococcus_neoformans_var._neoformans_JEC21 ------------TLD LNRKAFNSILFRPRV LVDVE-IA--DTRTQ MLGQDTSLPIFISPA GMAKLAHP-EGECLL AKAAGQ-----SNII
Neosartorya_fischeri_NRRL_181                 ------------TRD ANKSCFDRIWFRPRV LRNVR-SV--DSRTR ILGVDCSMPLFVSPA AMAKLIHP-DGECAI ARACER-----KGIM
Aspergillus_flavus_NRRL3357                   ------------TKI YNNLIYRSILLRPRV FIDCR-EC--DLSTR FLGLKLGLPIYISPA AMARLAHP-QGEAGI AAACRK-----FGAM
Ajellomyces_capsulatus_NAm1                   ------------TLR ENHSAFHKVWFRPRI LVDVQ-NV--DISTT MLGSPTSVPFYVTAT ALGKLGHP-EGEVCL TRAANT-----HNVI
Uncinocarpus_reesii_1704                      ARETPTADNGKQTMR ENHSAFHKIWFRPRI LVDVE-NV--DISTT MLGTPVSVPFYVTAT ALGKLGHA-DGEVCL TKAAAS-----HDVV
Neurospora_crassa_OR74A                       ------------TLR ENHAAFHRIWFRPKV LVDVE-KV--DFSTT MLGTKVDIPFYVTAT ALGKLGHV-EGEVLL TRAAKK-----HNVV
Pyrenophora_tritici-repentis_Pt-1C-BFP        ------------SYR NNLEIFQRFRLRPRF LTDVT-NVPNTMPTT ILGHNFSSPIFISPC ARGGYANDA-GEVGL AKGAGE-----AGIL
Talaromyces_stipitatus_ATCC_10500             ------------TMR ENHTAFHKVWFRPRI LVDVE-KV--DFSTT MLGSKTSVPFYITAT ALGKLGHP-EGEVVL TKAAHN-----HEVI
Paracoccidioides_brasiliensis_Pb01            ------------SLR ENHSAFHKIWFRPRV LVDVQ-NV--DITST MLGTPVSAPFYVTAA ALGKLGHP-EGEVCL TRAANT-----HNII
Verticillium_albo-atrum_VaMs.102              ------------SYR HNLDIWKSVQFRSRH LSDVT-KLNETLATT ILGYNFSAPFFIAPA ARGVYGDPERAELNF VEAAGK-----ENIL
Arthroderma_otae_CBS_113480                   ------------TMR ENHTAFHKIWFRPRI LVDVE-QV--SISTT MLGTPVSVPFYVTAT ALGKLGHP-DGEVCL TRASAT-----HDVI
Saccharomyces_cerevisiae_S288c                ------------THR ENHNAYHRIFFKPKI LVDVR-KV--DISTD MLGSHVDVPFYVSAT ALCKLGNPLEGEKDV ARGCGQGV---TKVP
Lachancea_thermotolerans                      ------------SIR ENHYAYSRVFFKPMI LQENEYDV--DTSTE FLGSKVSLPVYISAF AGSKWAHP-LAELNL QSAAYE-----ADIM
Trichophyton_rubrum_CBS_118892                ------------SKS LNNSIYRSILLRPRV FVDCK-NC--SLATS MLGYKLNTPIIASPT AMARLAHP-SGEAGI AAACAK-----FGAM
Candida_dubliniensis_CD36                     ------------TFR LNTGSYQRILFKPRV MIDVT-EV--DTSTT MLGTNVSAPFYITAT ALGKLGHP-DGEKVL TRGAYK-----HDII
Schizophyllum_commune_H4-8                    ------------SKN ENAHAFSRFFFHARV MRPVS-RC--DPSTT ILGFKSSIPVFVSGA AMAKLAHP-DGELNI TRGCAA-----QGII
Komagataella_pastoris_GS115                   ------------TLR ENHFAYHKVFFRPRI LVDVT-NI--ELETE MLGIKTSAPFYISAT ALAKLGHP-EGEVGI AKGAGR-----GDII
Perkinsus_marinus_ATCC_50983                  ------------SYA ENEDAFSRIALVPRV LVDVS-KV--DCSSS VLGRHFDVPFYMTAV AMAKLYNV-DGEKCV ARGIGKTKEAGIDMA
Karlodinium_micrum_1                          ------------GLR ENHAAFHRVMLKPRV LVDVD-NI--DMTST ILGTKVSIPLYVTSC ALGRLYHE-DGECCL ARGAAL-----AGIP
Karlodinium_micrum_2                          ------------TKR DNCAAFSRAWLKPRV MRNVL-SI--NTRCT LLGTEFAFPIFISPA AMAGLAHE-DAEPAL ARAAGK-----LGAL
Chromera_velia_gene4425.t1                    ------------TLR ENKETFDNIWLVPRV LRDVS-SI--DFATT MLGCSSSIPVYLTAT AVAKLAHP-DGELGI IRAAHA-----HGLV
Chromera_velia_gene5763.t1                    --------------- --------------- --------------- --------------- --------------- ---------------
Vitrella_brassicaformis_gene7868.t1           ------------TLR ENREAFNRIRLRPRV LNDVS-KV--NLNTS ILGHTSTMPVLVAPT GMQKLAHP-EGETAT ARAARE-----VGVP
Vitrella_brassicaformis_gene19079.t1          ------------SLR ENRVAFERIRLRPRF LIDVD-NI--QLQTS ILGYNVSMPLIVSPT AMQKMAHP-DGETGV AIACKE-----MGIP

                                             271
Cryptococcus_neoformans_var._neoformans_JEC21 QMISTNASAPLPSII SSA----TSPSQPFF MQLYVDRNRPKTES- LLGKIN-SLGLKAIF VTVDAPAPGKREADE RSRAEV---EVASGI
Neosartorya_fischeri_NRRL_181                 QGVSNNSSYTLDQLK EAA------PSANFF FQLYVNRDRSKSAA- LLHQCSANPNVKAIF VTVDAAWPGKREADE RVKADE---NLSVPM
Aspergillus_flavus_NRRL3357                   QLISHNASMTTQQIV ANAH-----PDQIFG WQLYCLKDVKRSEK- RIAEINSIKEIKFIC LTLDAPFPGKREIEE RQKMEEL--------
Ajellomyces_capsulatus_NAm1                   QMIPTLASCSFDEIV DARG-----PDQVQW LQLYVNKDRTITKR- IVQHAQ-QRGCKALF ITVDAPQLGRREKDM RSKFSDRGS-AVQAA
Uncinocarpus_reesii_1704                      QMIPTLASCSFDEIV DAAI-----DKQTQW LQLYVNKDRDITRK- IVNHAE-KRGCKGLF ITVDAPQLGRREKDM RSKFSDPGS-DVQQT
Neurospora_crassa_OR74A                       QMIPTLASCAFDEIM DAAE-----GDQVQW LQLYVNKDRAITER- IIKHAE-KRGCKALF ITVDAPQLGRREKDM RVKFTDDGS-NVQK-
Pyrenophora_tritici-repentis_Pt-1C-BFP        YMPSLYSSIPMEDIY AARA----SKDQVMF QQIYLNGNLSSTKA- LFDKAK-SLGAKGLV ITVDSAGSAIRHRAA RYGVGSANT------
Talaromyces_stipitatus_ATCC_10500             QMIPTLASCSFDEIV DARK-----GEQVQW LQLYVNKDRAITKR- IVQHAE-KRGCKALF ITVDAPQLGRREKDM RVKFSDTGS-NVQAS
Paracoccidioides_brasiliensis_Pb01            QMIPTLASCSFDEIV DARG-----PNQIQW LQLYVNKDRGITKR- IVQHAE-KRGCKALF ITVDAPQLGRREKDM RTKFSDRGS-DVQAS
Verticillium_albo-atrum_VaMs.102              YIPSMYASKTIEEIA AGKSNSTLNGPQVIF QQIYTNANLSVTWD- NIRRAE-RTGAKAIV FTIDAPGNSVRHRAA RYDTTNANS------
Arthroderma_otae_CBS_113480                   QMIPTLASCSFDQIV DAKT-----PRQTQW LQLYVNKDRDITRR- IVEHAE-ARGCKGLF ITVDAPQLGRREKDM RSKFAEQGS-NVQAS
Saccharomyces_cerevisiae_S288c                QMISTLASCSPEEII EAAP----SDKQIQW YQLYVNSDRKITDD- LVKNVE-KLGVKALF VTVDAPSLGQREKDM KLKFSNTKAGP-KAM
Lachancea_thermotolerans                      QMVPKQNSYSIEEFY ENVP-----EDQKHW SQYHFDSREEFNETL IKKLEA-QPSVKALF LNVDLRDIGNREKDS RQRALDVESSKSLSA
Trichophyton_rubrum_CBS_118892                QIISNNASMTPEEIV KGAT-----PDQVFG WQLYVQIERKKSEA- MLARINKIKAIKFIC LTLDAPVPGKRELDE RTKAIASTPAVADIV
Candida_dubliniensis_CD36                     QMIPTLASCSFDEIV DESK-----PNQTQW FQLYVNSDREITKK- IVQHAE-ARGMKGLF ITVDAPQLGRREKDM KTKSIVDLS-FVQGE
Schizophyllum_commune_H4-8                    QMVSSNASYSYAEIA HAAT-----PTQPLF FQLYKHKDDALALQ- RIREVE-ALGYKALF LTVDAVVPSKREFDI RAPWYLEELERGGPM
Komagataella_pastoris_GS115                   QMISTLASCSLDETV AAAK-----EGQSQW FQLYVNSDREVAYN- MIKHCE-ELGIKGIF VTVDAPSLGNREKDR RMKFTEDTDVDLSGD
Perkinsus_marinus_ATCC_50983                  YMIPTLASCGNGEFY GGLK----------- --------------- --DKAD-AR------ --------GRRMD-V ----------FLKGI
Karlodinium_micrum_1                          QLCPTLASCTMDEMH AARS-----PGQTQW WQLYVNKDRELTKT- VVQKAE-SLGFKALF ITVDAPQLGRRERDM RNKAKMSAN--VQTK
Karlodinium_micrum_2                          HVVANMASRELEEIT DARV-----PGQTQW YQIYVNPERSKTEA- IIKRAV-QAGVKALL VTVDTPQLGRRERDM RNKVIDSSNLSLVQK
Chromera_velia_gene4425.t1                    YMLPTLSSYAFDDML DAIQ-----PGQPFW MQLYVNRDRNMTAE- HVRKAA-KRGCRALC VTVDAPQLGRRERDM RNKLTKTEA-SIQKD
Chromera_velia_gene5763.t1                    --------------- --------------- --------------- --------------- --VDLPPLGHREADI RNRFHLPAPLRLENF
Vitrella_brassicaformis_gene7868.t1           MTLSTMSTTPLEEVA -ESS-----GKGLRF FQLYVFQQRELTVQ- LIKRAE-AAGYSAIV LTVDVPAMGRREADL RNHFTLPPHLTLANF
Vitrella_brassicaformis_gene19079.t1          MTLSTLATTSVTDVS ANCG-----GEGLRF FQLYTFRDRGITYD- LVKRAE-AAGFKALV LTVDTPTLGRRNADI RNQFSLPTGLSLANF

                                             361
Cryptococcus_neoformans_var._neoformans_JEC21 SG---------G--- --------KIGSDNK GG---------GIGR SVGGFIDPKLSWK-D IEWLR-----QHTKL PIGLKGVQTAEDAMK
Neosartorya_fischeri_NRRL_181                 AP---------A--- --------QAKNDKK GG---------GLGR VMAGFIDPGLTWD-D LVWVR-----KHTHL PVCLKGVMSADDAIL
Aspergillus_flavus_NRRL3357                   --------------- -------------RA AG---------AVLS PQVWGTDASLTWERT LNWLR-----MHTSL PIVLKGIQTYEDAIL
Ajellomyces_capsulatus_NAm1                   DG---------K--- --------SESSMDR SQ---------GAAR AISSFIDPSLSWK-D IPWFQ-----SLTDM PIVLKGVQRVDDVLR
Uncinocarpus_reesii_1704                      --------------- ---------DNSVDR SQ---------GAAR AISSFIDPSLSWK-D IPWFQ-----SITKM PIALKGVQRVDDVLR
Neurospora_crassa_OR74A                       -G------------- ----------HETNR NE---------GAAR AISSFIDPALSWK-D IPWFQ-----SVTKM PIILKGVQRVEDVIK
Pyrenophora_tritici-repentis_Pt-1C-BFP        --------------- --------------- QL------------- -------TKLTWE-V FQQLQ-----NLTDL PLIPKGIQTVEDTQE
Talaromyces_stipitatus_ATCC_10500             GG------------- ----------DSIDR SQ---------GAAR AISSFIDPSLSWK-D IPWFL-----SITKM PIILKGVQCVEDVLR
Paracoccidioides_brasiliensis_Pb01            DA---------N--- --------SESSVDR SQ---------GAAR AISSFIDPSLSWA-D IPWFQ-----SITTM PIVIKGVQRVDDVLR
Verticillium_albo-atrum_VaMs.102              --------------- --------------- VS------------- -------SALTWD-V RL------------- -------------CW
Arthroderma_otae_CBS_113480                   T-------------- ---------SGTVDR SQ---------GAAR AISSFIDPSLSWK-D LPYFR-----SLTSM PIALKGVQRVDDVLR
Saccharomyces_cerevisiae_S288c                --------------- --------KKTNVEE SQ---------GASR ALSKFIDPSLTWK-D IEELK-----KKTKL PIVIKGVQRTEDVIK
Lachancea_thermotolerans                      I-------------- ----------VTSDK SY------------- -------AKFTWK-D IDQIM-----SSTKL PIGLKGVQRGEDVVL
Trichophyton_rubrum_CBS_118892                K-------SSGG--- ----------HEITG GG---------GLGQ QLFAGTDPSLTWKDT LPWLL-----KHTDL PIVLKGIQTHEDAYI
Candida_dubliniensis_CD36                     D-------------- ----------DEADR SQ---------GSAR AISSFIDTSLSWK-D LEWFK-----SITKM PIILKGVQRVEDAII
Schizophyllum_commune_H4-8                    EFV------EEQ--- --------ADALQGQ SF---------GTAG GLIVNDDRDMTWERT IPWLR-----SVTRL PIVLKGIQCVEDALL
Komagataella_pastoris_GS115                   G-------------- ---------KTEVNR SN---------GAAA ALSSFIDTAVTWK-D IAEFK-----RRTNL PIVIKGIQRTEDVIL
Perkinsus_marinus_ATCC_50983                  --------------- --------------- --------------- -------HSRKWL-K YRAYWNACGL----- -----AVDAVRAYGE
```



```
                                              451
Cryptococcus_neoformans_var._neoformans_JEC21 AAKM--GVDAIYLSN HGGRALDGSPPAMYT LLEMNKICPE----- -------IFK--KCE VYIDGGCRRGTDVVK ALCLGAKGVGMGRPF
Neosartorya_fischeri_NRRL_181                 AMQA--GLDGILLSN HGGRNLDTSPPSIVT LLELHKRCPE----- -------IFD--KME IYVDSGIRRGTDILK AICLGATAVGMGRSM
Aspergillus_flavus_NRRL3357                   AAKHAPQVRGIVLSN HGGRALDTVSTPVHV LLEIRRFCPE----- -------VFD--RLD VIVDGGIQRGTDVVK ALALGAKAVGIGRAA
Ajellomyces_capsulatus_NAm1                   AVQM--GIPAVVLSN HGGRQLEFAPSAIEL LAEVMPELRRR-G-- -------WQN--RIE VYIDGGVRRGTDILK ALCLGAKGVGIGRPF
Uncinocarpus_reesii_1704                      AVEM--GVPAVVLSN HGGRQLEFAPSAIEL LADVMPALRAR-G-- -------WEN--KIE VFVDGGVRRATDIIK ALCLGAKGVGIGRPF
Neurospora_crassa_OR74A                       AVEA--GVQGVVLSN HGGRQLEFARSGIEV LAETMPVLREL-G-- -------LED--KIE VYIDGGIRRATDILK ALCLGAKGVGIGRPF
Pyrenophora_tritici-repentis_Pt-1C-BFP        AVKQ--GVKAVFLSN HGGRQIDGSPSTLQV AMEIHQRDPE----- -------LFK--KIE IYADGGIRYGTDILK LLALGVKAVGVGRPF
Talaromyces_stipitatus_ATCC_10500             AVEA--GVQGVVLSN HGGRQLDFARSGIEI LAEVMPVLRER-G-- -------WEN--KIE IFIDGGIRRGTDIIK ALCLGAKGVGIGRPF
Paracoccidioides_brasiliensis_Pb01            AVEA--GIPAVVLSN HGGRQLDFSPSSIEL LADVMPELRRR-G-- -------WQD--RIE VYIDGGVRRGTDILK ALCLGAKGVGIGRPF
Verticillium_albo-atrum_VaMs.102              AVEK--GAQAIYISN HGGRQLDHTPGPLEI AYEIYRNAPQ----- -------VFQ--QVD VLADSGIRYGSDVLK LLALGVKAVGMGRPF
Arthroderma_otae_CBS_113480                   AVEA--GIDAVVLSN HGGRQLEYAPSAIEL LAEVMPALRAR-G-- -------WER--KIE VYIDGGIRRASDIIK AVCLGAKGVGIGRPF
Saccharomyces_cerevisiae_S288c                AAEI--GVSGVVLSN HGGRQLDFSRAPIEV LAETMPILEQR-N-- -------LKD--KLE VFVDGGVRRGTDVLK ALCLGAKGVGLGRPF
Lachancea_thermotolerans                      AAEK--GVKAVVLSN HGGRQLDFSRPPLEV LAEAKQMLKER-G-- -------LED--KIE IYLDGGIRRGSDVIK ALCLGAKGVGLGRPF
Trichophyton_rubrum_CBS_118892                ASLHTPQIKGIILSN HGGRAMDTAPPSIHT LMEIRKYCPE----- -------VFN--RIE VWIDGGIKRGTDVVK ALCLGAKGVGVGRNA
Candida_dubliniensis_CD36                     AAEH--GCAGVVLSN HGGRQLEFSPPPIEV LAELMPILREK-G-- -------LAD--KFE VYIDGGVRRATDVLK AICLGAKGVGIGRPF
Schizophyllum_commune_H4-8                    AAEA--GVDGILISN HGGRQLDYSLPPIEV LYRLRKHHPE----- -------VFG--KME IYIDGGITRGSDVLK AVCLGATAVGLGRPY
Komagataella_pastoris_GS115                   AAEH--GVDGVVLSN HGGRQLDGAPPSLQV LAECMPVLRQR-G-- -------LDK--KLE VFVDGGIRRGTDIMK ALCLGAKGVGLGRPF
Perkinsus_marinus_ATCC_50983                  RNQL--GIRGIVVSN HGARQVDTVKSGVQM LYECTRALKKA-G-- ---WRGRIDP--EFS VFVDGGVRRGTDIIK CIALGASAVGIGRPF
Karlodinium_micrum_1                          AYEM--GMDGLVVSN HGGRQLDYARSGIEM LVEIMDALSSIGA-- -------DLE--KFT VLVDGGFRRGSDVFK ALALGAKGVGLGRPT
Karlodinium_micrum_2                          AAQH--GCAGVLVSN HGGRQLDHARPTFDI LVEVMQDLEEA-D-- -------LKD--KIE VYLDGGVRRGTDVYK ALALGAKAVGIGRPC
Chromera_velia_gene4425.t1                    GVAR--GADALILSN HGGRQVDTARPGIDV LPEVVATLRAAYGGE GRQLRGPNGA--PVE VFVDGGIRRATDVFK CLAMGATGVGIGRPV
Chromera_velia_gene5763.t1                    ARQH--RVDAVWVSN HGGRQLDCTIAPIDA LPEVVSALEG----- -----------SGID AWLDGGVRRGTDVVK ALALGAKLVGVGRPV
Vitrella_brassicaformis_gene7868.t1           AREH--GVSAIWVSN HGGRQLDGVDAPITV LPSIKQALAEKPV-- ---MESHFRMPTRMA IWVDGGVRRGSDILK ALALGADMVAVGRPV
Vitrella_brassicaformis_gene19079.t1          AKEQ--GVAAVWVSN HGARQLDGVDATIDV LEEVVDALR------ ----------GTSIE VYLDGGIRRGTDILK ALALGARAVAFGRPA

                                              541
Cryptococcus_neoformans_var._neoformans_JEC21 LYSL-TYGEEGVVHA IEIMRDEIETTMRLL GVTKLDQLGPHLLNT KALDPL--VFDQPMW GPGEKQ--------- ---------------
Neosartorya_fischeri_NRRL_181                 LFAT-NYGQEGVEHL IDIMKDELETAMRNN GITSLDEAGPHMVHT GDIDHL--VPDSRSH PYARAVAKGRRSREV AK-L-----------
Aspergillus_flavus_NRRL3357                   LYGLAAGGQSGVERT LQILADETATAMRLL GVQHVDQLSLQHVSW DILTIT--LSLTSIQ VNTR---------LV DSQISDSESSMLNAE
Ajellomyces_capsulatus_NAm1                   LYAMSAYGMPGVERA MQLLKDEMVMNMRLI GCSNIGQLCPDLVDA RGLAV-KTVPNPVDS LGL------------ --EVYD-PLGLPTDR
Uncinocarpus_reesii_1704                      LYAMSTYGVPGVERA MQLLKDEMTMNMRLL GCTSVDQLTPDLLDI RGLGH-HSVPNPVDR LAE------------ --SVYD-PLITPSDR
Neurospora_crassa_OR74A                       LYAMSAYGFDGVDRA MQLLKDEMEMNMRLI GATKIEDLNPGMLDL RSLYN--HGAPPTDT LSN------------ --ISYD-ALISPAQR
Pyrenophora_tritici-repentis_Pt-1C-BFP        MFAN-IYGADGVKKA ADLLKNELIMDAANM GVSDLKNVPLDLIDW KAASP----WVGNDQ I-------------- ---------------
Talaromyces_stipitatus_ATCC_10500             LYAMSAYGQEGVERA FQLLKDELEMNMRLI GAATIDDLKPSMVDT RGLVGGHY-SAPSDT LGL------------ --NVYD-PLVSP---
Paracoccidioides_brasiliensis_Pb01            LYAMSAYGVPGVERA MQLLKDELVMNMRLI GCSSIEQLCPDLVDT KGLAV-RSVPNPVDS LGM------------ --GVYE-SLVLPSER
Verticillium_albo-atrum_VaMs.102              MYSN-VYGLEGVTKA IDIMRTEIVRDGAQL GATNLQNISTSFLNT RQLESN--VYLFDQ- --------------- ---------------
Arthroderma_otae_CBS_113480                   LYAMSAYGTEGVEKA MQLLKDEMEMNMRLL GCTSIDQLGPDLLDT RGLSV-HSVSGPVDH LSR------------ --SVYE-PLTLPRER
Saccharomyces_cerevisiae_S288c                LYANSCYGRNGVEKA IEILRDEIEMSMRLL GVTSIAELKPDLLDL STLKA-RTVGVPNDV LYN------------ --EVYEGPT-LTEFE
Lachancea_thermotolerans                      LYAMAGYGEEGVSHL LDILRNEMKNNMRLL GVDKVEDLNENLVDI GSLKFRN--PRANDV LYD------------ --EAYQ-PLRFPPVR
Trichophyton_rubrum_CBS_118892                LFSLAAGGPEGVERM LEILCAETMTAMRLL GVDKVEDLGMQHINA RAVEQ---------- --------------- --QLYDGGAGLDILG
Candida_dubliniensis_CD36                     LYAMTGYGDAGVNKA IQLLKDEMVMNMRLL GVNKLEELNESFVDT RFMQ-TR--YVPDDV LFN------------ --RVYE-PVPLPKFK
Schizophyllum_commune_H4-8                    LYAQGAYGVAGVKRI THILETEIVTAMRLM GASRIKDLTPELVER VDWQVG--LKARL-- --------------- ---------------
Komagataella_pastoris_GS115                   LYANSAYGPDGVEKA IDILKNELIMNMRLL GVTKISDLSPEFVDT RPLFG----LTANDR LFN------------ --NNYL-DIEFPKFK
Perkinsus_marinus_ATCC_50983                  MTAMAAFGEAGMVRL AALLKEEILVNMRLL GCRSLEELSEEFLDA RAMEQPP----PGKV R-------------- --ARY----------
Karlodinium_micrum_1                          LVGMAAYGEEGVEKV VQIFKDEMEMHMRLM GTPTVADMVPKMVIT RNVAD-HFSPAPTDF LSH------------ --QVYD-PLPLATRV
Karlodinium_micrum_2                          MYAL-TFGQDGVEKC LQLIRDEFMLTMKLM GVTSIDQIRKKDIVL KAFGGTPF------- --N------------ --DGWP----IPSRL
Chromera_velia_gene4425.t1                    LWSLAAYGQPGVEKV ISLLKDEMTMLMRLM GAPRLADMRPEMCVY KGPNGGVPV--QGRV VGL------------ --RDMTKSQFLPSSS
Chromera_velia_gene5763.t1                    LWGLAAGGEKGVSRV LSMLKEELGTSLQLA GAPSVRDLSPDFLVM PGMTPF---CRPSSQ DTRKKMRRKTQR--- -----AAPAVLTRNR
Vitrella_brassicaformis_gene7868.t1           LWGLACGGQQGVVDV LSMIREELTLAMKLA GCNDVKKVSRQLVRL AHEPV---------- --------------- ---------------
Vitrella_brassicaformis_gene19079.t1          LWGLTYNGKEGVSQV LQMMKDELSLAMKLS GATNVSSVPRDLIRM PYDAP---------- -YKSKL--------- ---------------

                                              631
Cryptococcus_neoformans_var._neoformans_JEC21 -------------
Neosartorya_fischeri_NRRL_181                 -------------
Aspergillus_flavus_NRRL3357                   SVEGAFPARAKF-
Ajellomyces_capsulatus_NAm1                   LPAP--AAKSRL-
Uncinocarpus_reesii_1704                      PPAS--VGGGSK-
Neurospora_crassa_OR74A                       VKFA--EK-SKI-
Pyrenophora_tritici-repentis_Pt-1C-BFP        -------------
Talaromyces_stipitatus_ATCC_10500             -RFA--EK-SKL-
Paracoccidioides_brasiliensis_Pb01            LPPP--AEKSKL-
Verticillium_albo-atrum_VaMs.102              -------------
Arthroderma_otae_CBS_113480                   PISN--SKL----
Saccharomyces_cerevisiae_S288c                DA-----------
Lachancea_thermotolerans                      SAI----------
Trichophyton_rubrum_CBS_118892                S---HI--KAKI-
Candida_dubliniensis_CD36                     DT--------KL-
Schizophyllum_commune_H4-8                    -------------
Komagataella_pastoris_GS115                   DE-----------
Perkinsus_marinus_ATCC_50983                  -------------
Karlodinium_micrum_1                          SKM----------
Karlodinium_micrum_2                          -------------
Chromera_velia_gene4425.t1                    MR-P--KSGGALS
Chromera_velia_gene5763.t1                    TAVRGCTDAQRLG
Vitrella_brassicaformis_gene7868.t1           -------------
Vitrella_brassicaformis_gene19079.t1          -------------
```

Fig. S7. Multiple alignment of amino acid sequences of L-LDH (cytochrome) oxidoreductase. Due to occurrence of the enzyme limited to fungi, perkinsids, dinoflagellates, *Chromera* and *Vitrella*, we did not constructed the tree. Alignment was shortened.

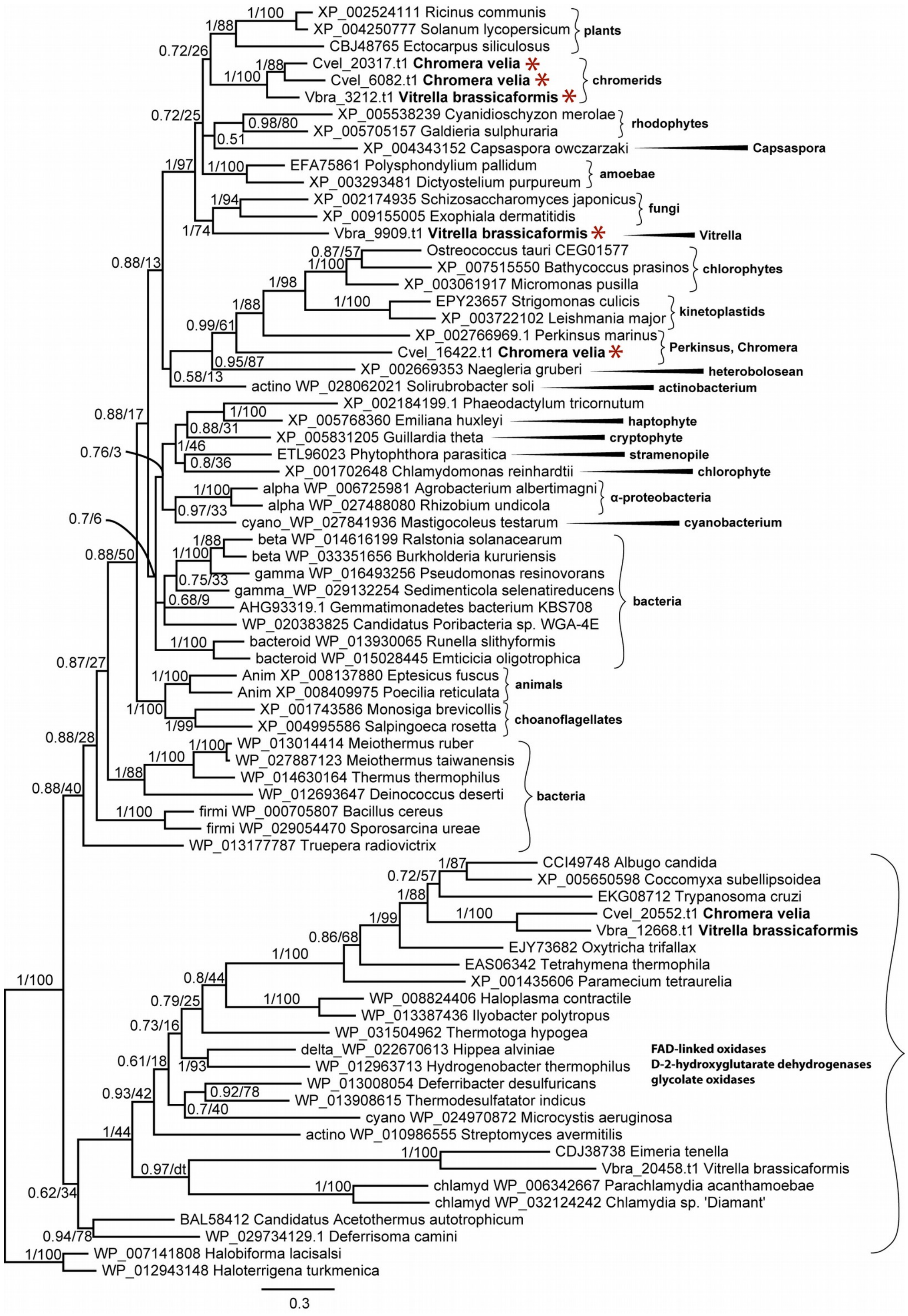


Fig. S8. Bayesian phylogenetic tree as inferred from D-lactate: cytochrome *c* oxidoreductase amino acid sequences. The supposed mitochondrial-located enzymes are marked by *. Numbers above branches indicate Bayesian PP/ML bootstrap support (1000 replicates).

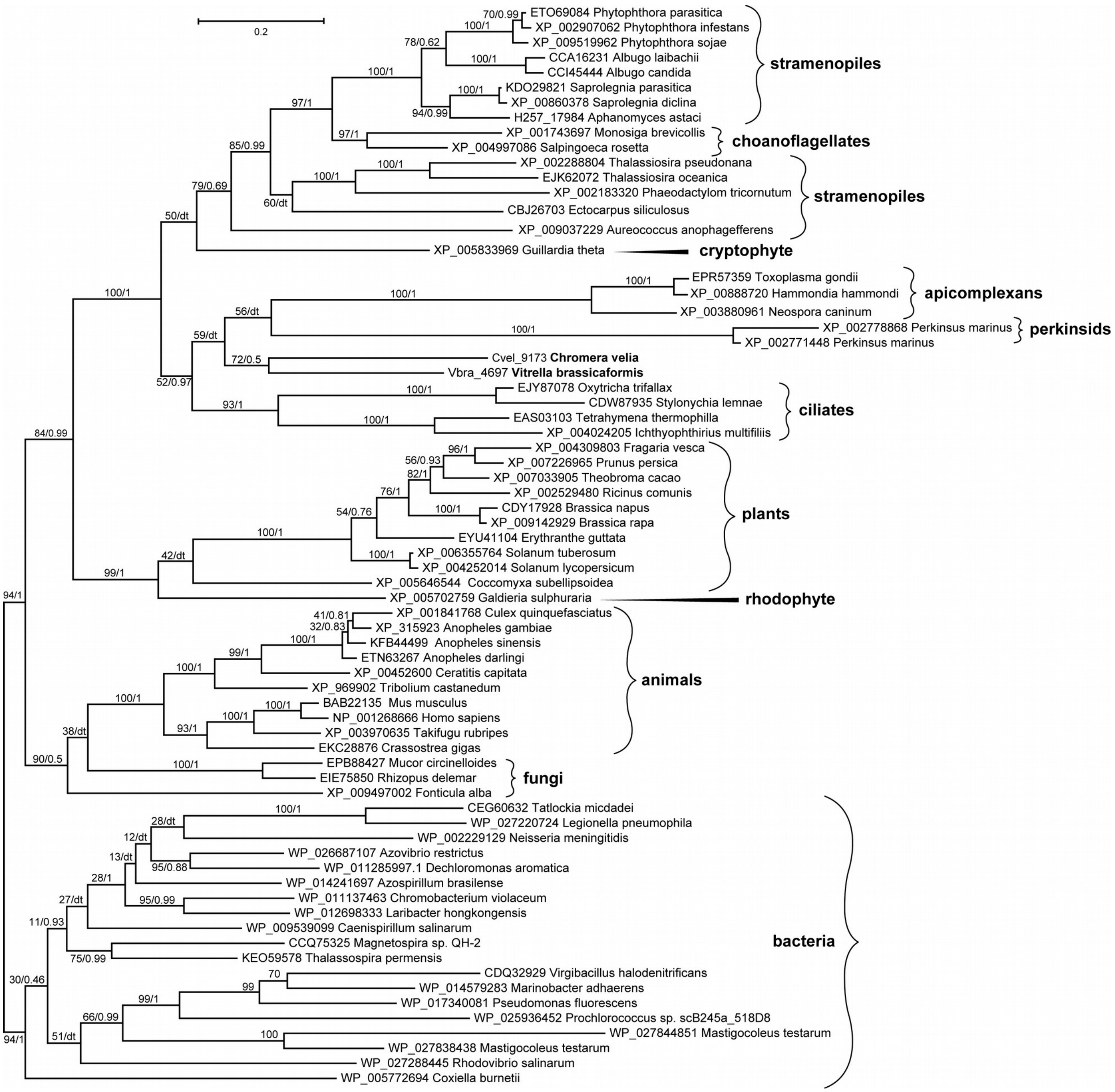


Fig. S9. Maximum likelihood as inferred from electron-transfer flavoprotein: ubiquinone oxidoreductase (ETFQO) amino acid sequences. Numbers above branches indicate ML bootstrap support (1000 replicates) /Bayesian PP.

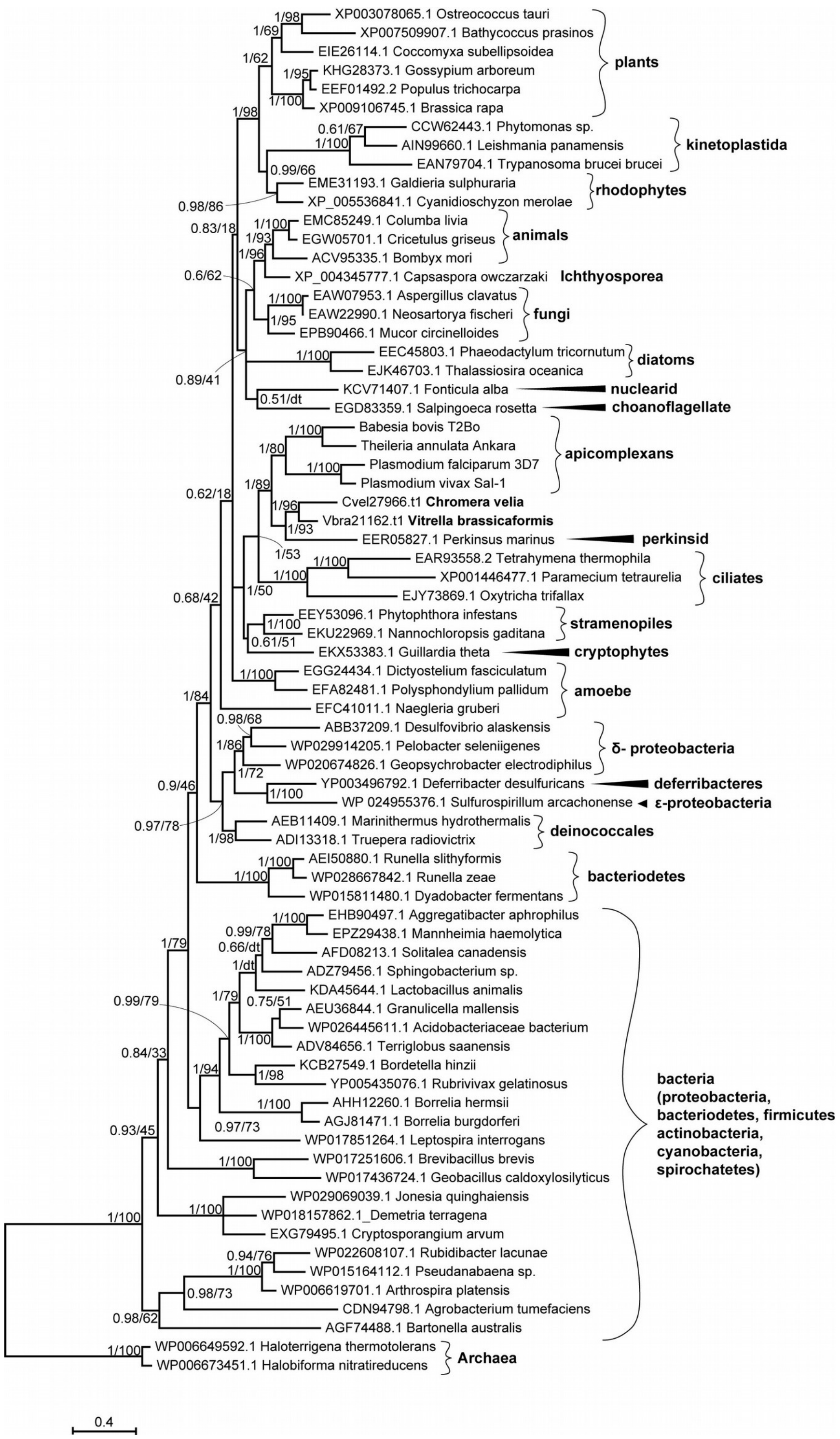


Fig. S10. The extended Bayesian phylogenetic tree as inferred glycerol-3-phosphate: ubiquinone oxidoreductase amino acid sequences. Numbers above branches indicate Bayesian PP/ML bootstrap support (1000 replicates).

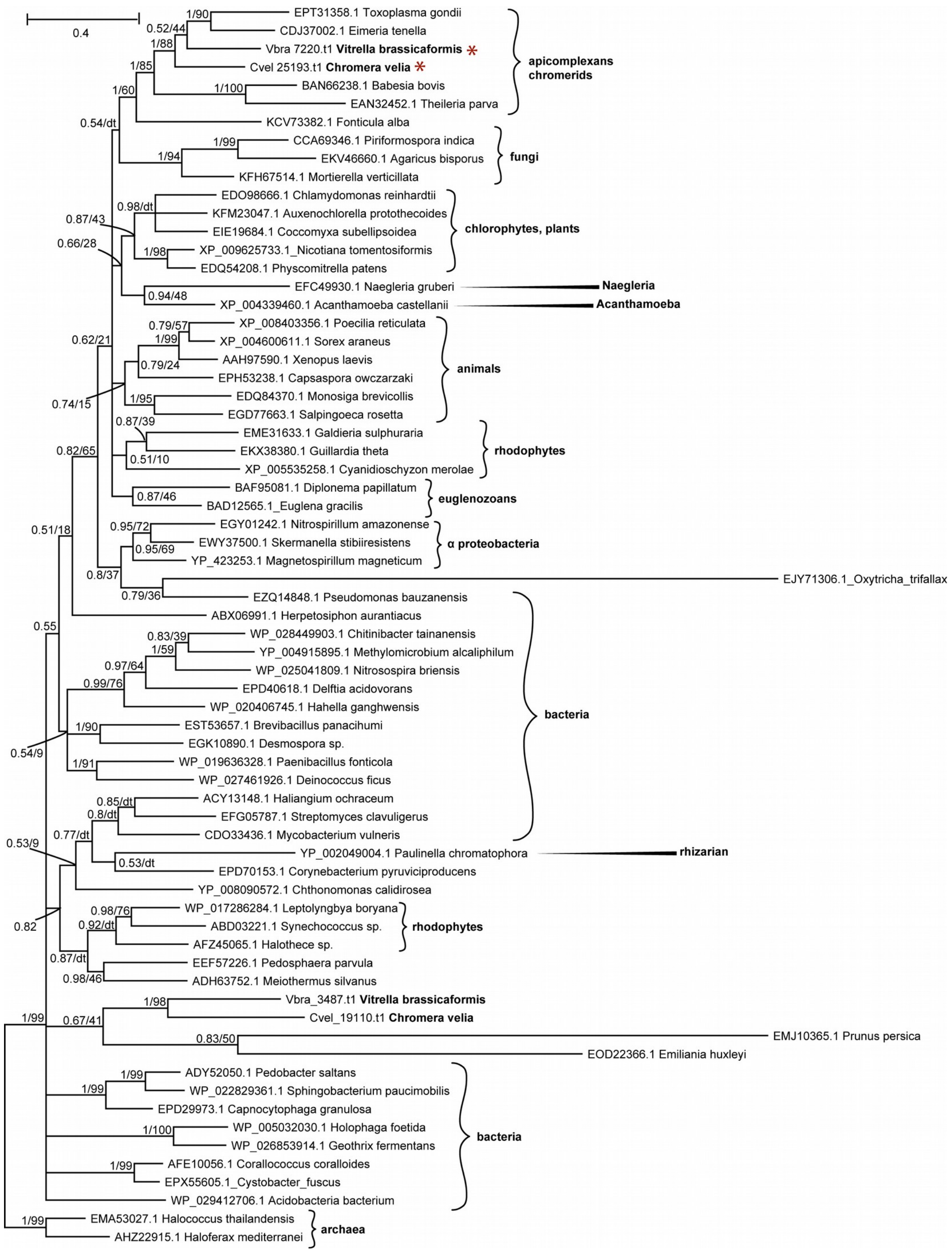


Fig. S11. Bayesian phylogenetic tree as inferred from dihydroorotate: ubiquinone oxidoreductase (DHODH) amino acid sequences. Targeting predictions were inconclusive in the case of this enzyme (see Table 2. For details). We suppose that proteins from *Chromera* and *Vitrella* (marked by *) branching together with apicomplexan homologs are used in the respiratory chain. Numbers above branches indicate Bayesian PP/ML bootstrap support (1000 replicates).

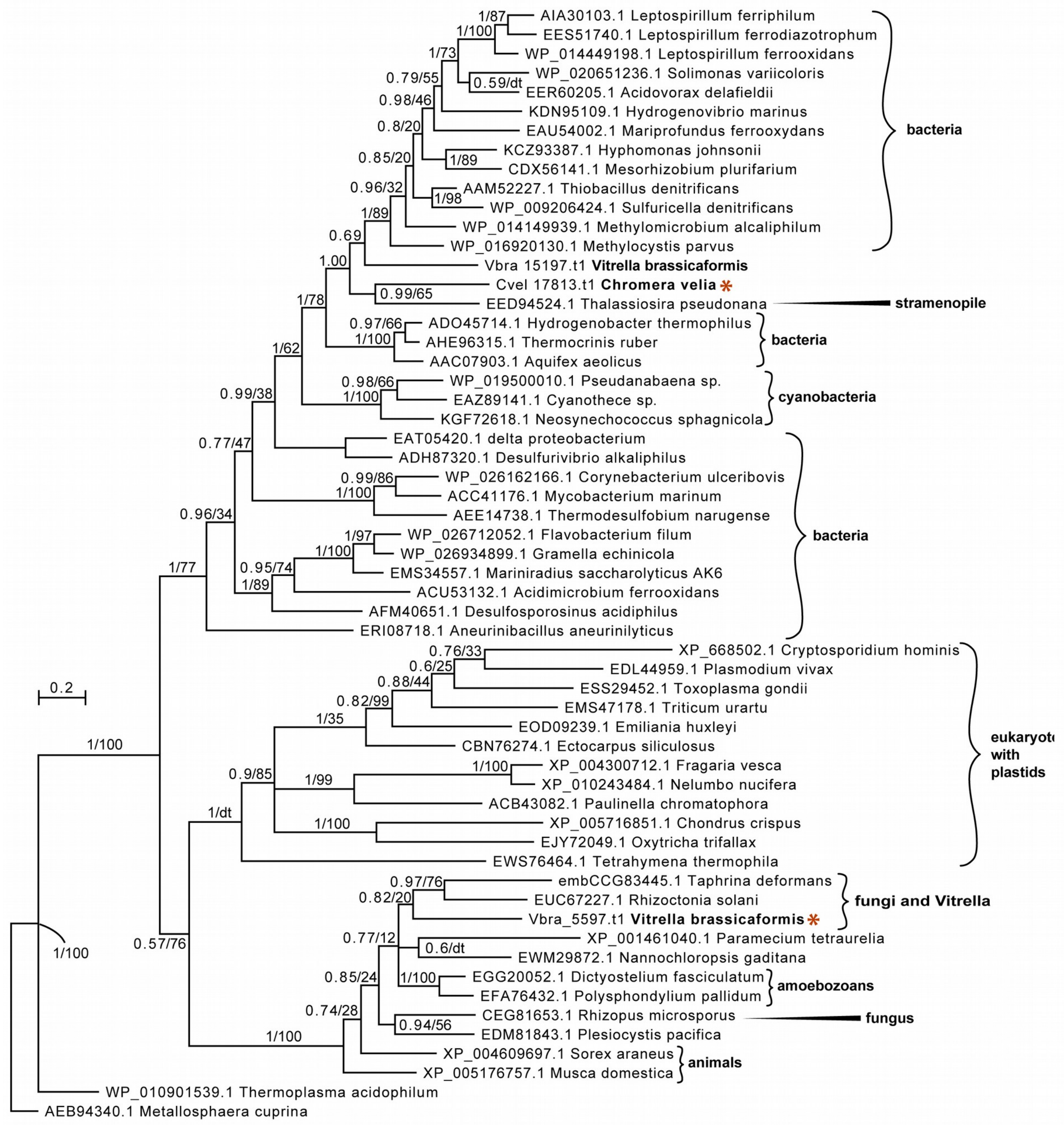


Fig. S12. Bayesian phylogenetic tree as inferred from sulfide: ubiquinone oxidoreductase (SQO) amino acid sequences. The supposed mitochondrial-located enzymes are marked by *. Numbers above branches indicate Bayesian PP/ML bootstrap support (1000 replicates).

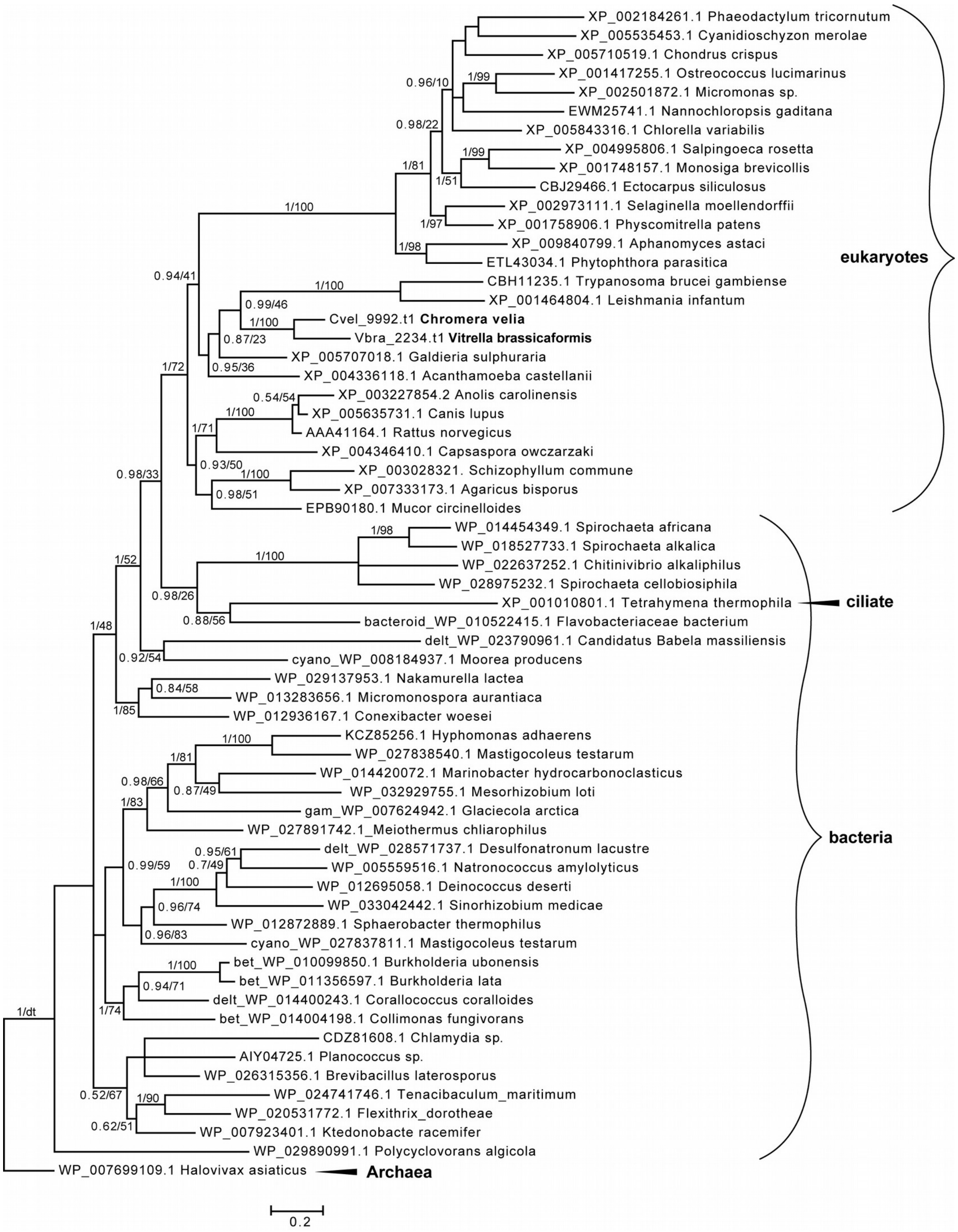


Fig. S13. Bayesian phylogenetic tree as inferred from galacto-1,4-lactone: cytochrome *c* oxidoreductase amino acid sequences. Numbers above branches indicate Bayesian PP/ ML bootstrap support (1000 replicates).

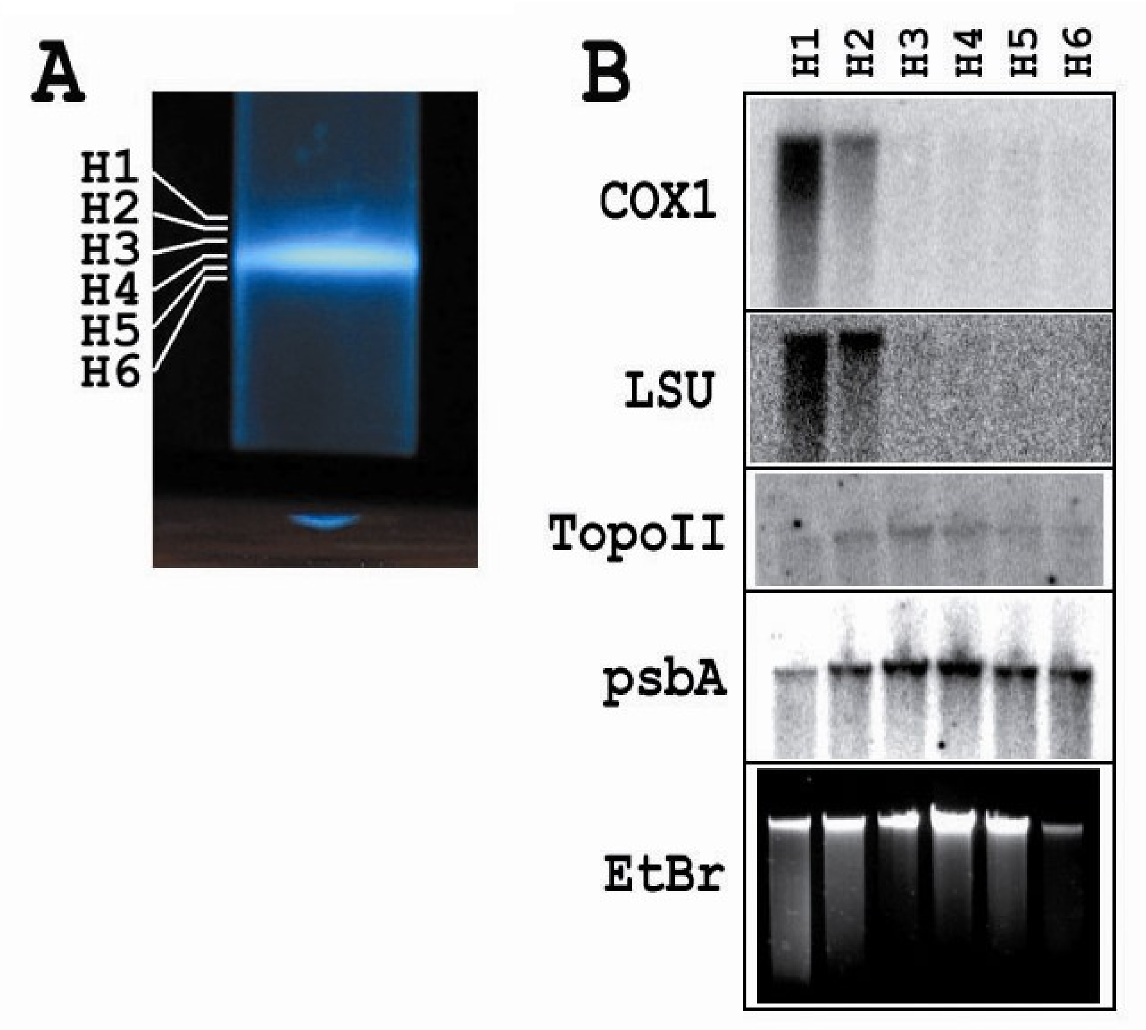


Fig. S14. Analysis of total DNA from *C. velia*. (**A**) Part of the CsCl-Hoechst 33258 density gradient of total cell DNA. The position of fractions H1 through H6 is indicated. (**B**) Southern hybridization of blotted agarose gel containing undigested DNA isolated from fractions H1 through H6 run in the presence of EtBr (bottom image). The probes used are indicated: *cox*1 (cytochrome *c* oxidase subunit 1), LSU (large subunit of mitoribosomal RNA *rnl*8), topoII (nuclear topoisomerase II) and *psbA* (photosystem complex subunit A).

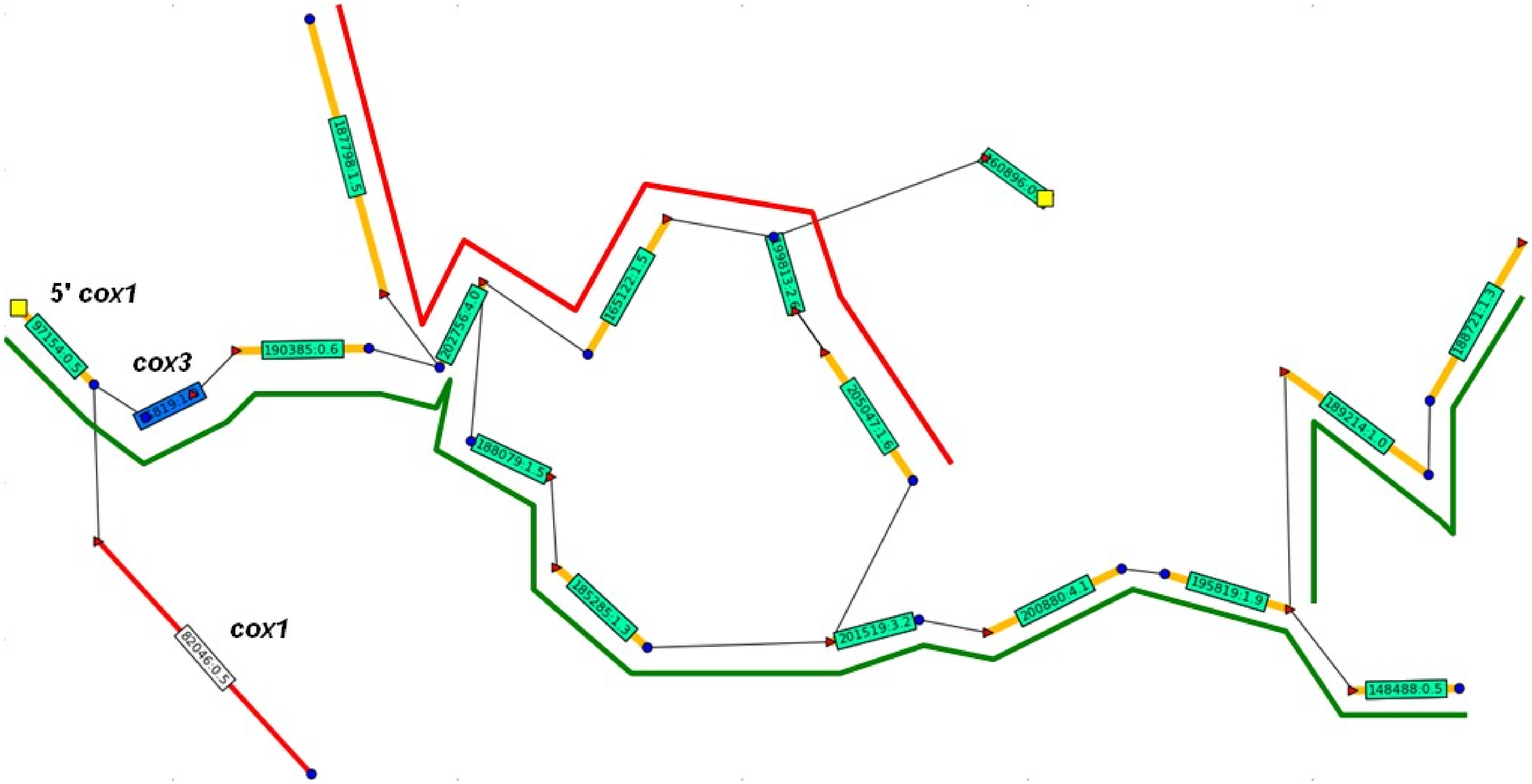


Fig. S15. An example of a contig graph, an output of the Newbler assembly software, visualized by an in-house script. Contigs are represented with bars in the following way: rectangles contain contig numbers and their relative coverage (absolute coverage divided by *cox1* contig coverage in this case) in the format number:coverage. Blue circles and red triangles and mark contig 5’ and 3’ ends, respectively. Direct connections between contigs, i.e. reads flowing from one contig into another, are shown with thin lines, and yellow rectangles mark contig termini with no direct connections to any other contigs. Contigs containing the *cox1* and *cox3* gene sequences are indicated. This figure shows pathways in the graph starting from one of the highest-coverage contigs, contig 97154, containing a 5'-fragment of *cox1* (amino acids 1-192). From the 5' end this contig reads flow either into a downstream part of the *cox1* gene (contig 82046, marked *cox1*), or into the *cox3* gene (contig 75819, marked *cox3*), and then into contigs 190385 and 202756. From that point there are two possible pathways: either through contig 165122, or through contig 188079. Based on the analysis of 454 read flow and on mapping of paired Illumina reads of the mitochondrial DNA-enriched fraction, it can be concluded that of the two alternative pathways at this point, the pathway indicated with a green line probably represents a genomic sequence, and the 'red' pathway represents no actual sequence. The 'red' pathway contains regions not covered with Illumina reads, and not all contig connections in this path are supported by 454 read flow. All *Chromera* mitochondrial sequences in this paper were constructed and validated in a similar way: i/ contigs with hits to the *cox1* or *cox3* genes, or to rRNA fragments acted as seeds, i.e. starting contigs, for contig graph construction (*cox3* contig 75819 in the example presented); ii/ all possible pathways in the graph in both 5' and 3' directions from seed contigs were considered; iii/ only pathways supported by all of the following were taken as final 'supercontigs': a) by flows of 454 or Illumina MiSeq reads between contigs; b) by complete coverage with paired-end Illumina reads of the mitochondrial DNA-enriched fraction (see an example in supplementary fig. S16); c) by complete coverage with Illumina mate pair inserts 1700-5500 bp in length (data not shown).

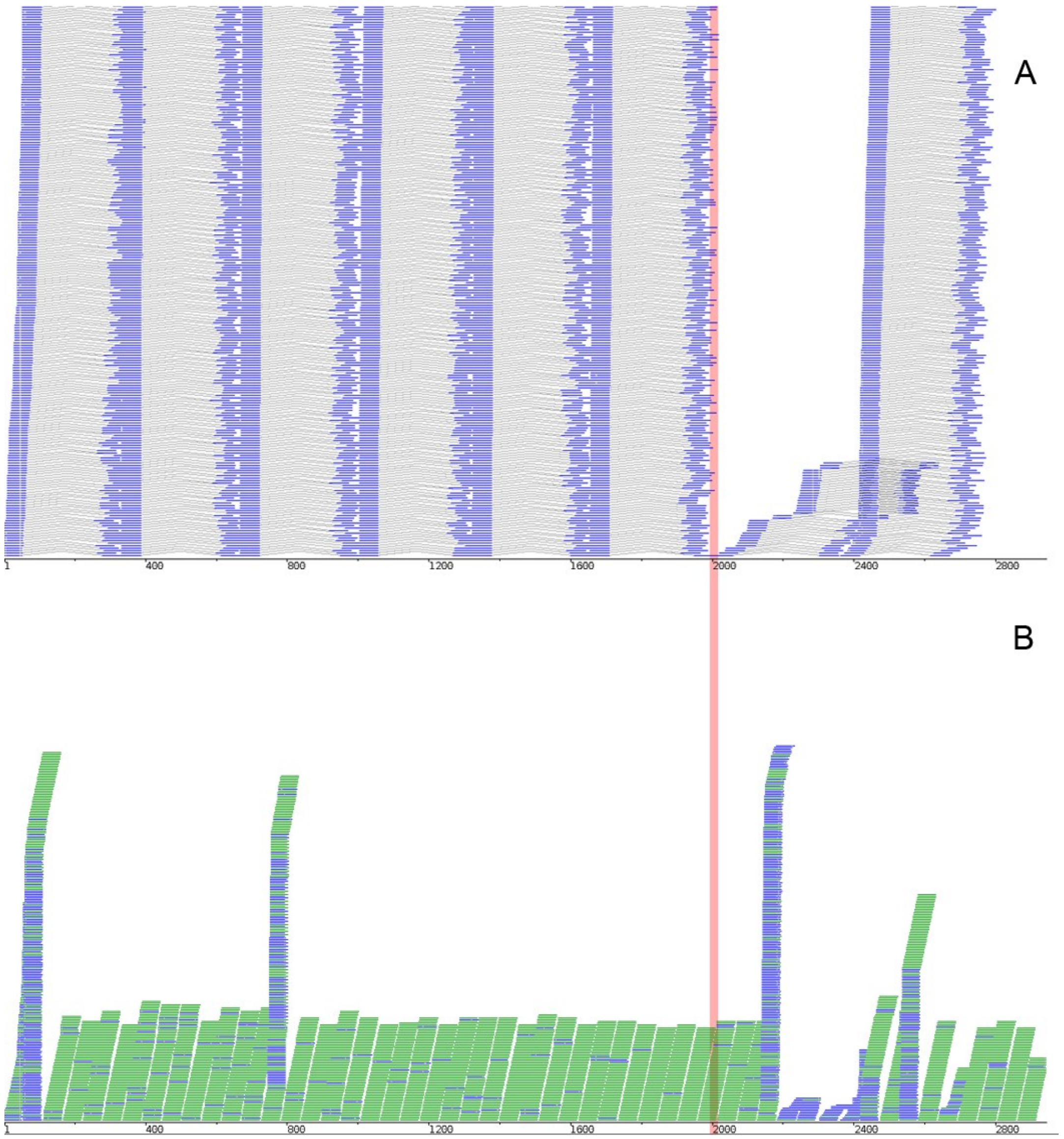


Fig. S16. An example of *Chromera* mitochondrial 'supercontig' with paired Illumina reads of the mitochondrial DNA-enriched fraction (supplementary table S3) mapped on it. Bowtie2 v.2.1.0 was used for mapping with the 'very sensitive' and 'end-to-end' options, and a narrow insert size range of 300-355 bp. Read pairs and individual reads are shown in panels (**A**) and (**B**), respectively. As seen in panel B, the full 'supercontig' is covered with reads of the expected insert size, making it a highly supported contig.

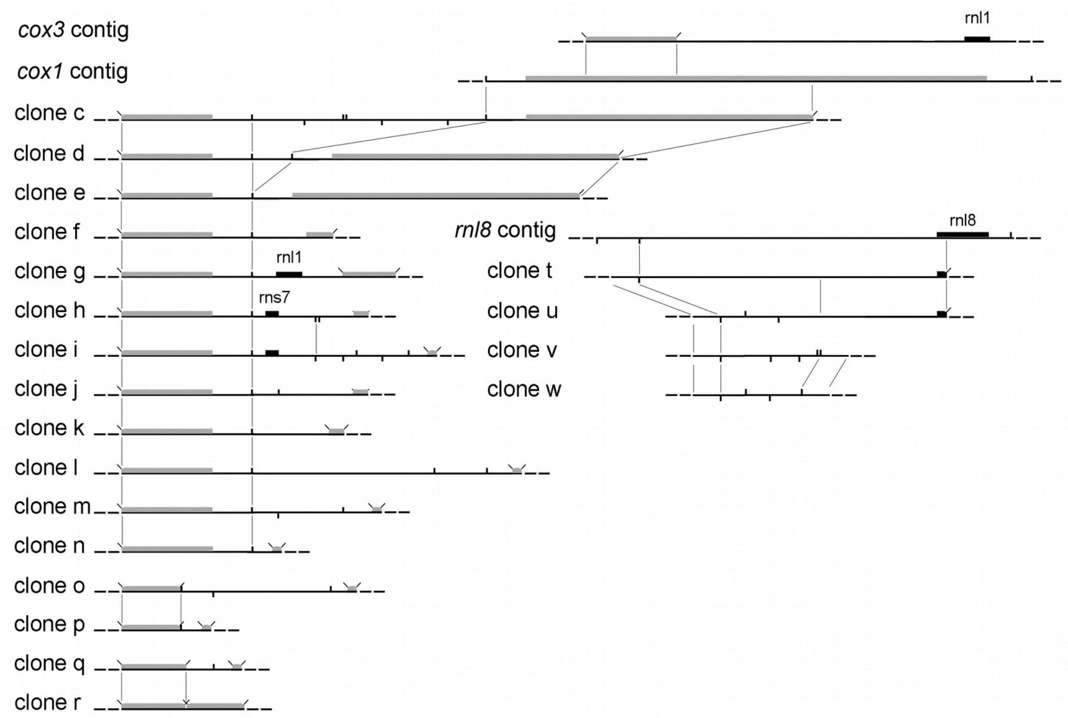


Fig. S17. Diversity of gene arrangements as determined by PCR (clones c-r, t-w). Most clones represent highly variable spacers between *cox1* genes/gene fragments. Few contigs assembled from Illumina and 454 data are shown for comparison (*cox1*, *cox3*, *rnl8* contigs). PCR products were amplified either by ‘outward’ *cox1*-specific oligonucleotides or by ‘outward’ and ‘inward’ rRNA-specific oligonucleotides. Grey rectangles indicate *cox1* gene sequences, black boxes indicate rRNA gene fragments, and short slanted lines at their ends mark truncated gene ends. Small vertical bars indicate short dispersed repeats, probably acting as recombination hot-spots. Sequence regions nearly identical among the clones are aligned between thin vertical or slanted lines.

Table S1. A list of *Chromera* mitochondrial contigs. The contigs shown here were generated with the contig graph analysis approach (supplementary fig. S15) based on initial contigs assembled with Newbler v.2.9 from 454 and Illumina reads (supplementary table S3), compared with PCR-generated clones (supplementary fig. S17), and validated with mapping mate pair reads and paired-end reads of the mitochondrial DNA-enriched fraction (supplementary fig. S16). Length, coverage with reads of the mitochondrial fraction, and GenBank accession numbers are shown.

contig name	**length, bp**	**average coverage, reads of the mitochondrial DNA-enriched fraction**	**GenBank accession #**
11012_rnl8_1	**3186**	**466**	
114504_cox1_3_c1_1	**1782**	**2264**	
114504_cox1_3_c1_2	**11724**	**368**	
114504_cox1_3_c1_4	**1623**	**1881**	
114504_cox1_3_c1_5	**2046**	**1975**	
114504_cox1_3_c1_6	**2116**	**1870**	
114504_cox1_3_c1_7	**2130**	**1897**	
114504_cox1_3_c1_9	**1620**	**1881**	
114504_cox1_3_c1_10	**2043**	**1934**	
114504_cox1_3_c1_11	**2113**	**1909**	
114504_cox1_3_c1_12	**2127**	**1204**	
114504_cox1_c1_5	**3030**	**2200**	
114504_cox1_c1_6	**12972**	**649**	
114504_cox1_c1_8	**2871**	**2494**	
114504_cox1_c1_9	**3294**	**2473**	
114504_cox1_c1_10	**3364**	**2397**	
114504_cox1_c1_11	**3378**	**2412**	
114504_cox1_c1_13	**2868**	**2494**	
114504_cox1_c1_14	**3291**	**2474**	
114504_cox1_c1_15	**3361**	**2397**	
114504_cox1_c1_16	**3375**	**2412**	
114504_cox1_c2_5	**3303**	**2282**	
114504_cox1_c2_6	**13245**	**701**	

114504_cox1_c2_8	3144	2554	
114504_cox1_c2_9	3567	2528	
114504_cox1_c2_10	3637	2456	
114504_cox1_c2_11	3651	2469	
114504_cox1_c2_13	3141	2554	
114504_cox1_c2_14	3564	2528	
114504_cox1_c2_15	3634	2456	
114504_cox1_c2_16	3648	2470	
114504_cox1_c3_5	3251	2249	
114504_cox1_c3_6	13193	687	
114504_cox1_c3_8	3092	2523	
114504_cox1_c3_9	576	2285	
114504_cox1_c3_10	646	1907	
114504_cox1_c3_11	660	1993	
114504_cox1_c3_13	3089	2524	
114504_cox1_c3_14	576	2285	
114504_cox1_c3_15	646	1907	
114504_cox1_c3_16	660	1993	
114504_cox1_c4_5	3229	2279	
114504_cox1_c4_6	13171	692	
114504_cox1_c4_8	3070	2557	
114504_cox1_c4_9	3493	2530	
114504_cox1_c4_10	3563	2457	
114504_cox1_c4_11	3577	2471	
114504_cox1_c4_13	3067	2558	
114504_cox1_c4_14	3490	2531	
114504_cox1_c4_15	3560	2457	
114504_cox1_c4_16	3574	2471	
114504_cox1_c5_5	3126	2259	
114504_cox1_c5_6	13068	674	
114504_cox1_c5_8	2967	2546	
114504_cox1_c5_9	3390	2519	
114504_cox1_c5_10	3460	2444	
114504_cox1_c5_11	3474	2458	
114504_cox1_c5_13	2964	2546	
114504_cox1_c5_14	3387	2520	
114504_cox1_c5_15	3457	2444	
114504_cox1_c5_16	3471	2459	
114504_cox1_c6_5	3390	2357	

114504_cox1_c6_6	13332	731	
114504_cox1_c6_8	3231	2625	
114504_cox1_c6_9	3654	2543	
114504_cox1_c6_10	3724	2520	
114504_cox1_c6_11	3738	2533	
114504_cox1_c6_13	3228	2626	
114504_cox1_c6_14	3651	2544	
114504_cox1_c6_15	3721	2521	
114504_cox1_c6_16	3735	2534	
114504_non-coding_c2_2	920	1892	
114504_non-coding_c2_3	1343	2031	
114504_non-coding_c2_4	1413	1871	
114504_non-coding_c2_5	1427	1911	
114504_non-coding_c2_7	917	1892	
114504_non-coding_c2_8	1340	2032	
114504_non-coding_c2_9	1410	1871	
114504_non-coding_c2_10	1424	1912	
114504_non-coding_c3_1	618	2035	
114504_non-coding_c3_2	632	2122	
114504_non-coding_c4_1	721	2013	
114504_non-coding_c4_2	791	1728	
114504_non-coding_c4_3	805	1802	
114504_non-coding_c5_2	677	1971	
114504_non-	747	1673	

coding_c5_3			
114504_non-coding_c5_4	761	1753	
114504_non-coding_c9_9	868	2297	
13057_rnl12_1	5579	187	
13113_rns7_1	7644	128	
13113_rns7_2	4519	80	
154998_non-coding_1	604	1460	
154998_non-coding_3	1042	2803	
154998_non-coding_4	925	2755	
154998_non-coding_5	481	1951	
154998_non-coding_7	919	2668	
154998_non-coding_8	802	2593	
154998_non-coding_9	375	2152	
154998_non-coding_10	328	4291	
19169_cox1_frag_1	7774	116	
203480_non-coding_3	393	3223	
274408_non-coding_3	564	3015	
274408_non-coding_4	320	2511	
274408_non-coding_6	584	3131	
33769_cox1_frag_1	4645	694	
33769_cox1_frag_2	4604	700	
33769_cox1_frag_3	5611	605	
33769_cox1_frag_4	5707	596	
33769_cox1_frag_5	4663	693	
33769_cox1_frag_6	4622	698	
33769_cox1_frag_7	5629	603	
33769_cox1_frag_8	5725	595	
33769_cox1_frag_9	4644	695	
33769_cox1_frag_10	4603	737	
33769_cox1_frag_11	5610	607	
33769_cox1_frag_12	5706	566	
33769_cox1_frag_13	4662	692	
33769_cox1_frag_14	4621	735	
33769_cox1_frag_15	5628	605	
33769_cox1_frag_16	5724	564	
33769_non-coding_1	1490	154	

33769_non-coding_2	1449	158	
33769_non-coding_3	2456	163	
33769_non-coding_4	2552	161	
33769_non-coding_5	1508	159	
33769_non-coding_6	1467	158	
33769_non-coding_7	2474	163	
33769_non-coding_8	2570	161	
33769_non-coding_9	1489	159	
33769_non-coding_10	1448	277	
33769_non-coding_11	2454	167	
33769_non-coding_12	2551	94	
33769_non-coding_13	1507	154	
33769_non-coding_14	1466	275	
33769_non-coding_15	2473	167	
33769_non-coding_16	2569	92	
34169_cox1_frag_1	2146	197	
36603_cox1_3_1	3706	596	
36603_non-coding_1	4069	161	
36603_non-coding_2	4919	193	
36603_non-coding_3	3863	562	
36603_non-coding_4	3698	166	
36603_non-coding_5	3604	180	
36603_non-coding_6	1579	27	
3725_rRNA_1	12088	196	
37286_rns7_1	2690	170	
41777_rnl12_rns7_1	2021	438	
43250_rns7_1	2736	582	
46689_cox3_frag_c12_1	3879	16	
46689_non-coding_c13_1	1189	2475	
46919_cox3_cox1_1	4478	1566	
46919_cox3_cox1_2	1966	3039	
46919_cox3_cox1_3	2239	3057	
46919_cox3_cox1_4	1965	3041	
46919_cox3_cox1_5	2238	3059	
46919_cox3_cox1_6	2061	3090	
46919_cox3_cox1_7	2185	3029	
46919_cox3_cox1_8	2164	3081	

46919_cox3_cox1_9	2325	3139	
67435_non-coding_1	3670	810	
67435_non-coding_2	835	944	
67435_non-coding_3	6338	185	
67435_non-coding_4	666	1248	
67435_rns7_1	2860	751	
68201_non-coding_2	779	2195	
68201_non-coding_3	658	2108	
68201_non-coding_7	454	1709	
68201_non-coding_10	393	3223	
68201_non-coding_12	702	2191	
68201_non-coding_13	608	1783	
73351_non-coding_8	3169	433	
73351_non-coding_9	3489	470	
73351_non-coding_10	1741	600	
73351_non-coding_11	1413	656	
73351_non-coding_12	2441	573	
73351_non-coding_13	3554	468	
73351_rRNA_3	3427	628	
73351_rRNA_4	2112	646	
75915_rnl8_1	444	1719	
95117_non-coding_1	945	5088	
95117_non-coding_2	2392	2619	
95117_non-coding_6	1391	1349	
95117_non-coding_8	1967	1735	
95117_non-coding_10	1429	1385	
95117_non-coding_11	1509	1341	
95117_non-coding_12	1812	1471	
95117_non-coding_18	1524	1030	
95117_non-coding_19	1541	1270	
95117_non-coding_28	2741	526	
95117_non-coding_29	2413	549	
95117_non-coding_30	3441	522	
95117_non-coding_31	4554	430	
95117_non-coding_32	4489	452	
95117_non-coding_33	3216	436	
95117_non-coding_35	3950	509	
95117_non-coding_36	3379	520	
95117_non-coding_37	4554	430	

95117_non-coding_38	3743	641	
95117_non-coding_39	3415	668	
95117_non-coding_40	4443	620	
95117_non-coding_41	5556	524	
95117_non-coding_42	5491	544	
95117_non-coding_45	4952	599	
95117_non-coding_46	4381	620	
95117_non-coding_47	5556	524	
95117_non-coding_48	2729	566	

Table S2. A non-redundant set of transcripts detected in *Chromera* mitochondrial contigs, excluding full-length *cox1* transcripts or long *cox1* fragments. Strand-specific transcriptomic reads (supplementary table S3) were mapped on mitochondrial contigs (supplementary table S1) with Bowtie2 v.2.1.0, and transcripts were identified using a minimum coverage threshold of 10. For each contig, transcripts are numbered sequentially (see the ‘transcript name’ column). A substantial number of genome regions, including, e.g., the *cox1* gene, is transcribed from both strands which is indicated in a separate column. Annotated features overlapping each transcript (genes, gene fragments, ORFs, BLAST hits to *Plasmodium* rRNA fragments) are shown in the following way: feature name; feature coordinates within the transcript (‘<’ and ‘>’ symbols denote features extending beyond the transcript 5’ and 3’end, respectively); length of a full ORF (not necessarily contained within the transcript) and its start codon. It should be noted that unidentified transcripts had no reliable BLAST hits to the NCBI nr database at the E-value threshold of 10^{-5} (data not shown). The bottom section of the table lists high-coverage transcripts of rRNA fragments presented in the same fashion. *Rnl1*, *rnl8*, *rnl12*, *rns1*, and *rns7* transcripts were detected, and very short hits to *Plasmodium rns5* were found within a much longer transcript (see contigs 68201_non-coding_2 and 68201_non-coding_3). Due to their high-coverage, a number of short transcripts with no sequence similarity to *Plasmodium* rRNA fragments was annotated as putative rRNAs, until further investigation.

contig	transcript name	size, bp	strands transcribed	feature name	feature coordinates within transcript	full ORF length	start codon
11012_rnl8_1	transcript_1	124					
	transcript_2	86	both	*cox*3 fragment, 449-476 bp	complement(17..45)		
				ORF1	complement(<1..>86)		
	transcript_3	144	both	ORF2	<1..59	306	CAT
	transcript_4	139	both				

114504_cox1_3_c1_1	transcript_1	148	both	ORF2	52..>148	315	CTA
	transcript_2	214	both				
	transcript_3	121	both				
	transcript_4	168	both				
114504_cox1_3_c1_2	transcript_2	132	both	putative rRNA	1..202		
	transcript_3	487					
	transcript_4	700		ORF1	5..>700	723	ATG
114504_cox1_3_c1_4	transcript_2	881	both	ORF2	<1..153	315	CTA
				ORF3	complement(496..>881)	387	TAT
				*Plasmodium rnl*8 hit	complement(104..196)		
				*rnl*8 transcript	complement(1..228)		
114504_cox1_3_c1_5	transcript_2	744	both	ORF2	<1..50	315	CTA
				ORF3	complement(393..>744)	675	ATA
				*Plasmodium rnl*8 hit	complement(1..93)		
				*rnl*8 transcript	complement(<1..125)		
114504_cox1_3_c1_7	transcript_2	1357	both	ORF2	<1..153	315	CTA
				ORF3	complement(496..957)	462	ACC
				*Plasmodium rnl*8 hit	complement(104..196)		
				*rnl*8 transcript	complement(1..228)		
114504_cox1_3_c1_10	transcript_3	435	both	ORF3	138..>435	675	ATA

114504_cox1_c3_6	transcript_4	696		ORF2	1..>696	723	ATG
114504_non-coding_c2_2	transcript_1	694	both	ORF1	complement(309..>694)	387	TAT
				*Plasmodium rnl*8 hit	complement(<1..9)		
114504_non-coding_c2_5	transcript_1	1170	both	ORF1	complement(309..770)	462	ACC
				*Plasmodium rnl*8 hit	complement(<1..9)		
114504_non-coding_c3_2	transcript_1	594	both				
114504_non-coding_c4_1	transcript_1	711	both	ORF1	138..545	408	ATA
114504_non-coding_c4_3	transcript_1	761	both				
114504_non-coding_c5_4	transcript_1	714	both				
114504_non-coding_c9_9	transcript_1	856	both	ORF1	complement(213..521)	309	AGG
13057_rnl12_1	transcript_1	758		ORF1	complement(462..>758)	366	CTT
				ORF2	466..>758	387	GGG
				ORF3	612..>758	537	TGT
				ORF4	complement(703..>758)	480	CTC
	transcript_2	420	both	ORF3	<1..265	537	TGT
				ORF4	complement(<1..299)	480	CTC
				ORF5	13..294	309	TGG
	transcript_3	222					
	transcript_4	93	both	ORF6	complement(<1..>93)	408	AAT
				ORF7	<1..>93	348	GTA
				ORF8	complement(9..>93)	441	TTA
	transcri	92		ORF8	78..>92	441	TTA

	pt_5						
	transcript_6	116					
13113_rns7_1	transcript_1	254		ORF2	<1..>254	876	TTT
	transcript_2	175					
				ORF3	complement(244..1422)	1179	AAA
				ORF3a	230..679	450	CGT
				ORF4	complement(630..1268)	639	CTA
	transcript_4	146					
	transcript_5	227					
	transcript_6	102					
13113_rns7_2	transcript_1	123					
	transcript_2	803		ORF1	complement(90..>803)	744	CCA
				ORF2	46..801	756	GAC
	transcript_3	217					
154998_non-coding_1	transcript_2	280	both	ORF1	complement(<1..104)	318	ATG
154998_non-coding_3	transcript_1	751	both				
154998_non-coding_5	transcript_2	157	both	ORF1	complement(<1..104)	318	ATG
154998_non-coding_9	transcript_1	362	both				
154998_non-coding_10	transcript_1	328	both				
19169_cox1frag_1	transcript_1	273		ORF1	complement(<1..106)	1062	ATG
	transcri	184		ORF1	complement(<1..>	106	ATG

	pt_2				184)	2	
				ORF2	<1..>184	843	ATG
	transcript_3	206		ORF3	complement(<1..>206)	1566	ATG
				ORF4	<1..97	1227	ATG
	transcript_4	143		ORF3	complement(<1..>143)	1566	ATG
				ORF4	<1..>143	1227	ATG
	transcript_5	119		ORF10	<1..62	1539	ATG
				ORF6	<1..>119	1344	ATG
				ORF8	complement(<1..>119)	909	ATG
				ORF9	complement(<1..>119)	693	ATG
	transcript_6	298		ORF10	<1..>298	1539	ATG
	transcript_9	107	both				
203480_non-coding_3	transcript_1	393	both				
274408_non-coding_4	transcript_1	319	both				
274408_non-coding_6	transcript_1	584	both				
33769_cox1frag_1	transcript_1	118	both				
	transcript_2	167	both				
	transcript_4	152	both				
	transcript_5	229	both				
	transcri	135	both	ORF2	complement(<1..7	315	ATG

	pt_6				6)		
				*Plasmodium rns*7 hit	complement(134..>135)		
	transcript_7	133	both	*cox*1 fragment, 1-79 bp	complement(14..92)		
				ORF1	complement(9..>133)		
	transcript_8	127	both	ORF3	complement(123..>127)	348	ATG
33769_non-coding_5	transcript_1	1505		ORF1	294..1292	999	ATG
				ORF2	complement(282..1274)	993	ATG
				ORF3	complement(290..700)	411	ATG
				ORF4	complement(82..594)	513	ATG
33769_non-coding_6	transcript_1	1464		ORF1	253..1251	999	ATG
				ORF2	complement(241..1233)	993	ATG
				ORF3	complement(249..659)	411	ATG
				ORF4	complement(14..553)	540	ATG
33769_non-coding_7	transcript_1	2471		ORF1	1260..2258	999	ATG
				ORF2	complement(1248..2240)	993	ATG
				ORF3	complement(1256..1666)	411	ATG
				ORF4	complement(1021..1560)	540	ATG
				ORF5	748..1251	504	ATG
				ORF6	300..944	645	ATG
				ORF7	complement(537..890)	354	ATG
				ORF8	complement(137..448)	312	ATG
				ORF9	20..367	348	ATG
33769_non-	transcri	256		ORF1	1356..2354	999	ATG

coding_8	pt_1	7		ORF2	complement(1344..2336)	993	ATG
				ORF3	complement(1352..1762)	411	ATG
				ORF4	complement(1117..1656)	540	ATG
				ORF5	844..1347	504	ATG
				ORF6	396..1040	645	ATG
				ORF7	complement(633..986)	354	ATG
				ORF8	complement(233..544)	312	ATG
				ORF9	116..463	348	ATG
34169_cox1frag_1	transcript_1	990		ORF1	<1..276	375	ATG
				ORF2	complement(<1..287)	378	ATG
				ORF3	complement(172..657)	486	ATG
				ORF4	269..688	420	ATG
				ORF5	286..708	423	ATG
				ORF6	complement(716..>990)	315	ATG
36603_cox1_3_1	transcript_1	572	both	*cox*1 fragment, 1193-1766 bp	<1..560		
				ORF1	49..>572		
	transcript_2	132	both				
	transcript_3	355					
	transcript_4	1376		ORF1	complement(473..880)	408	ATG
				ORF2	complement(436..756)	321	ATG
				ORF3	39..416	378	ATG
				ORF4	20..379	360	ATG
36603_non-coding_1	transcript_1	1664		ORF1	109..486	378	ATG
				ORF2	complement(227..595)	369	ATG

				ORF3	complement(622..1035)	414	ATG
				ORF4	complement(1364..>1664)	306	ATG
	transcript_2	172	both	ORF5	<1..>172	330	ATG
	transcript_3	417		ORF5	complement(304..>417)	330	ATG
				ORF6	<1..13	318	ATG
	transcript_4	294		ORF6	complement(<1..51)	318	ATG
				ORF7	complement(<1..229)	423	ATG
				ORF8	177..>294	417	ATG
	transcript_5	171		ORF10	<1..>171	441	ATG
				ORF9	complement(<1..>171)	381	ATG
	transcript_6	169		ORF11	complement(<1..>169)	495	ATG
36603_non-coding_2	transcript_1	950	both	ORF1	661..>950	378	ATG
				ORF2	complement(779..>950)	369	ATG
	transcript_2	1158		ORF2	complement(<1..96)	369	ATG
				ORF3	complement(86..406)	321	ATG
				ORF4	201..617	417	ATG
				ORF5	complement(858..>1158)	306	ATG
36603_non-coding_4	transcript_2	341					
	transcript_3	1143		ORF5	complement(<1..470)	537	ATG
				ORF6	163..888	726	ATG
				ORF7	complement(265..870)	606	ATG
				ORF8	complement(635..	372	ATG

					1006)		
36603_non-coding_5	transcript_1	1563		ORF1	152..538	387	ATG
				ORF2	complement(535..942)	408	ATG
				ORF3	complement(982..1290)	309	ATG
36603_non-coding_6	transcript_1	1097		ORF1	complement(<1..424)	486	ATG
				ORF2	117..842	726	ATG
				ORF3	complement(219..824)	606	ATG
				ORF4	complement(589..960)	372	ATG
3725_rRNA_1	transcript_1	143					
	transcript_2	367		ORF1	<1..119	615	ATG
	transcript_3	159		ORF1	complement(<1..>159)	615	ATG
	transcript_4	832		ORF3	1..528	612	ATG
				ORF4	<1..161	606	ATG
	transcript_5	854		ORF4	complement(<1..426)	606	ATG
				ORF4a	<1..633		
	transcript_6	172		ORF5	complement(<1..>172)	678	ATG
				ORF6	<1..>172	624	ATG
				ORF7	<1..>172	606	ATG
	transcript_7	98					
37286_rns7_1	transcript_1	108					
	transcript_2	133					
41777_rnl12_rns7_1	transcript_1	114	both				
43250_rns7_1	transcript_1	550		ORF1	<1..>550	642	GTG

	transcript_2	99	both				
	transcript_3	133	both				
	transcript_4	118	both				
46689_cox3frag_c12_1	transcript_1	2932		ORF1	<1..1073	1254	ATG
				ORF2	complement(139..840)	702	ATG
				ORF3	complement(194..1045)	852	ATG
				ORF4	218..922	705	ATG
				ORF5	complement(861..1745)	885	ATG
				ORF6	901..1686	786	ATG
				ORF7	932..1579	648	ATG
				ORF8	complement(1310..2269)	960	ATG
				ORF9	2125..2745	621	ATG
46689_non-coding_c13_1	transcript_1	871		ORF1	complement(527..>871)	519	ATG
				ORF2	complement(621..>871)	363	ATG
				ORF3	complement(280..738)	459	ATG
	transcript_2	130	both				
46919_cox3_cox1_1	transcript_2	134	both	ORF2	32..>134	462	TGT
46919_cox3_cox1_9	transcript_1	2325		*cox*1 fragment, 1-579 bp	complement(1452..2030)		
				*cox*3	complement(582..1451)		
				ORF1 *cox*1-*cox*3	complement(582..2282)	1701	TCA
				ORF2	1978..2286	309	TCA
				*Plasmodium rnl*1 hit	170..247		

				*rnl*1 transcript	137..263		
67435_non-coding_1	transcript_1	145	both				
	transcript_2	176	both				
	transcript_3	127	both	ORF1	<1..>127	327	CAA
	transcript_4	143	both	ORF1	<1..82	327	CAA
				ORF2	complement(69..>143)	318	TCG
	transcript_5	109	both	ORF4	complement(<1..>109)	360	GGA
	transcript_6	196	both	ORF4	<1..61	360	GGA
	transcript_7	361	both				
	transcript_8	134	both				
67435_non-coding_2	transcript_1	120	both				
	transcript_2	133	both	ORF1	complement(<1..>133)	555	GAA
	transcript_3	116	both	ORF1	complement(<1..>116)	555	GAA
67435_non-coding_3	transcript_2	158	both				
	transcript_3	311	both				
	transcript_4	1485		ORF1	<1..546	747	GAA
				ORF2	complement(<1..997)	1029	AAC
				ORF3	102..1001	900	TTG
				ORF4	149..853	705	CAG
				ORF5	complement(279..>1485)	1281	AAT
	transcript_5	532		ORF6	complement(<1..232)	723	AAG
				ORF7	complement(<1..4	750	AAG

					55)		
				ORF8	<1..>532	1557	TTA
				ORF9	complement(519..>532)	819	CCA
	transcript_6	102		ORF10	<1..>102	693	TCC
				ORF11	complement(<1..>102)	618	CCC
				ORF8	<1..>102	1557	TTA
				ORF9	complement(<1..>102)	819	CCA
67435_non-coding_4	transcript_1	236	both	ORF1	<1..159	519	GAA
68201_non-coding_2	transcript_1	776	both	ORF1	154..588	435	CTT
				*Plasmodium rns*5 hit	complement(658..675)		
68201_non-coding_3	transcript_1	607	both	ORF1	<1..419	465	NNN
				*Plasmodium rns*5 hit	complement(489..506)		
68201_non-coding_7	transcript_1	376	both				
68201_non-coding_12	transcript_1	700	both				
68201_non-coding_13	transcript_1	606	both				
73351_non-coding_8	transcript_4	160	both				
73351_non-coding_10	transcript_1	156	both	ORF2	complement(<1..>156)	447	ATC
73351_non-coding_12	transcript_1	150	both				
73351_non-coding_13	transcript_1	246	both	ORF1	complement(70..>246)	384	GAA
	transcript_2	105	both	ORF1	complement(<1..>105)	384	GAA
73351_rRNA_3	transcri	172					

	pt_1						
	transcript_2	122	both				
73351_rRNA_4	transcript_3	153	both	ORF1	complement(<1..>153)	414	AGA
	transcript_4	122	both				
95117_non-coding_2	transcript_1	737	both	ORF2	complement(148..480)	333	TAT
95117_non-coding_6	transcript_1	601	both				
95117_non-coding_8	transcript_1	1174	both	ORF1	643..1059	417	TCC
				ORF2	complement(207..512)	306	ATA
				*Plasmodium rns*5 hit	complement(1056..1073)		
95117_non-coding_10	transcript_1	639	both				
95117_non-coding_11	transcript_1	718	both	ORF1	186..512	327	ACT
95117_non-coding_12	transcript_1	1022	both	ORF1	490..816	327	ACT
95117_non-coding_18	transcript_1	103	both	ORF1	complement(<1..103)	390	AAG
				ORF2	complement(<1..72)	375	TGT
	transcript_2	577	both	ORF1	<1..234	390	AAG
				ORF2	<1..250	375	TGT
				ORF3	complement(<1..320)	327	TAT
				ORF4	complement(119..439)	321	CCC
				putative rRNA	5..123		
95117_non-coding_19	transcript_1	751	both	ORF1	2..424	423	TAC
				ORF2	complement(293..613)	321	CCC
95117_non-	transcri	132	both				

coding_30	pt_3						
95117_non-coding_35	transcript_1	115	both	ORF1	<1..>115	324	GTT
				ORF2	complement(<1..13)	432	ATG
95117_non-coding_38	transcript_3	104	both	ORF3	complement(71..>104)	333	GTT
	transcript_4	304	both	ORF3	complement(<1..171)	333	GTT
	transcript_5	138	both	ORF4	36..>138	321	GGC
95117_non-coding_42	transcript_5	133	both				
rRNA transcripts, sorted by transcript name							
46919_cox3_cox1_1	*rnl*1 transcript	127		*cox*1-*cox*3 transcript	complement(<1..>127)		
				transcript 1	<1..>127		
				*Plasmodium rnl*1 hit	34..111		
73351_rRNA_3	*rnl*1 transcript	196		*Plasmodium rnl*1 hit	80..158		
11012_rnl8_1	*rnl*8 transcript	214		*Plasmodium rnl*8 hit	40..68		
114504_cox1_3_c1_1	*rnl*8 transcript	228		ORF2	complement(76..>228)	315	CTA
				*Plasmodium rnl*8 hit	33..125		
114504_cox1_3_c1_10	*rnl*8 transcript	228		transcript 2	complement(<1..228)		
				ORF2	complement(76..>228)	315	CTA
				*Plasmodium rnl*8 hit	33..125		
114504_cox1_c3_16	*rnl*8 transcript	231		transcript 2	complement(<1..231)		
				ORF2	complement(79..>231)	315	CTA
				*Plasmodium rnl*8 hit	36..128		
114504_non-	*rnl*8	179		*Plasmodium rnl*8 hit	<1..84		

coding_c2_10	transcript						
33769_cox1frag_1	*rnl*8 transcript	171		*Plasmodium rnl*8 hit	33..125		
3725_rRNA_1	*rnl*8 transcript	206		*Plasmodium rnl*8 hit	27..100		
75915_rnl8_1	*rnl*8 transcript	444	both	*Plasmodium rnl*8 hit	352..444		
13057_rnl12_1	*rnl*12 transcript	137	both	*Plasmodium rnl*12 hit	complement(75..106)		
41777_rnl12_rns7_1	*rnl*12 transcript	164	both	*Plasmodium rnl*12 hit	44..75		
73351_non-coding_10	*rnl*12 transcript	175	both	ORF1	complement(<1..65)	339	TGG
				ORF2	142..>175	447	ATC
				*Plasmodium rnl*12 hit	39..70		
73351_rRNA_3	*rnl*12 transcript	189	both	*Plasmodium rnl*12 hit	106..137		
95117_non-coding_35	*rnl*12 transcript	190	both	ORF2	complement(141..>190)	432	ATG
				*Plasmodium rnl*12 hit	53..84		
46919_cox3_cox1_1	*rns*1 transcript	261	both	ORF4	complement(<1..23)	318	TAC
				rns1 hit	56..108		
37286_rns7_1	*rns*7 transcript	96		ORF1	<1..87	672	GTG
				*Plasmodium rns*7 hit	21..60		
41777_rnl12_rns7_1	*rns*7 transcri	101		ORF1	complement(<1..>101)	363	TAC

	pt			*Plasmodium rns*7 hit	23..62		
43250_rns7_1	*rns*7 transcript	109	both	*Plasmodium rns*7 hit	27..66		
67435_rns7_1	*rns*7 transcript	110	both	*Plasmodium rns*7 hit	28..67		
73351_non-coding_8	*rns*7 transcript	103	both	*Plasmodium rns*7 hit	22..61		
73351_non-coding_10	*rns*7 transcript	106	both	ORF1	complement(<1..>106)	339	TGG
				*Plasmodium rns*7 hit	24..63		
73351_non-coding_12	*rns*7 transcript	105	both	*Plasmodium rns*7 hit	23..62		
95117_non-coding_48	*rns*7 transcript	103	both	*Plasmodium rns*7 hit	22..61		
114504_cox1_c3_10	putative rRNA	105		*cox*1 transcript	complement(<1..105)		
				ORF1	complement(<1..27)	1857	TTA
19169_cox1frag_1	putative rRNA	141					
46919_cox3_cox1_3	putative rRNA	122		*cox*1-*cox*3 transcript	complement(<1..122)		
				transcript 1	<1..122		
46919_cox3_cox1_8	putative rRNA	104		*cox*1-*cox*3 transcript	complement(<1..104)		
				transcript 1	<1..104		
				ORF1	complement(<1..51)	1530	TTC
95117_non-	putative	119		transcript 2	<1..>119		
				ORF1	<1..>119	390	AAG

coding_18	rRNA		ORF2	<1..>119	375	TGT
			ORF4	complement(115..>119)	321	CCC
			ORF3	complement(<1..>119)	327	TAT
114504_cox1_3_c1_2	putative rRNA1	202	transcript 3	1..>202		
114504_cox1_3_c1_2	putative rRNA2	261				

Table S3. Genomic and transcriptomic read datasets generated in frame of this work (marked with an asterisk) and in frame of the *Chromera/Vitrella* genome project (Pain A, Otto, TD, Keeling PJ, Lukeš L, unpublished data), with those used for mitochondrial genome assembly indicated. Adapter removal and the following trimming and filtering settings were applied in CLC Genomics Workbench v. 6.5 to selected Illumina read datasets prior to mitochondrial genome assembly: terminal base runs with Phred quality below 20 were trimmed; longest read fragments with no more than one undetermined base were kept; read length threshold for filtered reads was set to 75 bp. No length threshold was applied to 54 bp reads of the mitochondrial DNA-enriched fraction. 454 reads were trimmed by the Newbler genome assembler itself prior to assembly.

Species	**Library type**	**Sequencing platform**	**Expected average insert size**	**Average read length**	**Read number, million**	**Used for mitochondrial genome assembly**	**Estimated insert size range**	**Average read length after filtering**	**Read number in million after filtering**
Genomic libraries									
C. velia	Illumina PCR-free	Illumina HiSeq	350	101	173				
C. velia	Illumina PCR-free	Illumina HiSeq	350	101	162				
C. velia	Illumina PCR-free	Illumina HiSeq	350	101	219				
C. velia	Illumina mate pair	Illumina HiSeq	8,000	101	152				
C. velia	Illumina	Illumin	3,000	101	133				

	mate pair	a HiSeq							
C. velia	Illumina mate-pair	Illumina HiSeq	3,000	101	47				
C. velia	Ilumina Nextera mate pair	Illumina HiSeq	3,000	101	165	yes	1700-5500	95	53
C. velia	Illumina mate pair	Illumina MiSeq	8,000	250	8	yes	230-800	190	3.5
C. velia	Illumina mate pair	Illumina MiSeq	5,000	250	9	yes	230-800	183	3.7
C. velia	Illumina mate pair	Illumina MiSeq	3,000	250	12	yes	230-800	182	4.6
C. velia *	454 single reads	454 GS FLX	N/A	224	0.29	yes	N/A		
C. velia *	454 single reads	454 GS FLX	N/A	225	0.3	yes	N/A		
C. velia	454 single reads	454 GS FLX	N/A	294	0.61	yes	N/A		
C. velia	454 single reads	454 GS FLX	N/A	326	0.63	yes	N/A		
C. velia	454 single reads	454 GS FLX	N/A	292	0.63	yes	N/A		
C. velia	454 single reads	454 GS FLX	N/A	304	0.73	yes	N/A		
C. velia *	Illumina	Illumin	330	54	19	yes	240-	54	13

	TruSeq, mitochondrial DNA-enriched fraction	a GAII				820
V. brassicaformis	Illumina TruSeq + Kappa Hifi	Illumina HiSeq	350	101	288	
V. brassicaformis	Illumina TruSeq + Kappa Hifi	Illumina HiSeq	350	101	203	
V. brassicaformis	Illumina PCR-free	Illumina HiSeq	500	93	155	
Transcriptomic libraries						
C. velia	Illumina TruSeq strand-specific	Illumina HiSeq	300	101	93	
C. velia	Illumina TruSeq strand-specific	Illumina HiSeq	300	101	118	
V. brassicaformis	Illumina TruSeq strand-specific	Illumina HiSeq	300	101	195	

Table S4. Statistics for the best *Chromera* genome assembly made with Newbler v.2.9, from which mitochondrial contigs were extracted. The assembly was made from read sets indicated in supplementary table S3. Extensive tuning of assembly parameters (data not shown) was employed to achieve as complete assembly of mitochondrial coding sequences as possible.

assembly statistics		insert size estimated by Newbler v.2.9, bp
# single 454 reads	3,188,774	N/A
# Illumina mate pair reads	26,674,272	1700-5500
# Illumina paired-end MiSeq reads	11,789,818	230-800
# Illumina paired-end reads of the mitochondrial DNA-enriched fraction	13,136,966	240-820
# total reads	54,789,830	
# reads assembled	45,731,247	
total assembly size, Mbp	256	
# scaffolds	34,422	
scaffold length threshold, bp	1,000	
largest scaffold, bp	5,019,587	
scaffold N50, bp	9,382	
large contig length theshold, bp	500	
contig N50, bp	5,465	

Supplementary references:

Jirků M, Kolesnikov AA, Benada O, Lukeš J. 1995. Marine fish and ray trypanosomes have large kinetoplast minicircle DNA. Mol Biochem Parasitol. 73:279-283.

Lukeš J, Jirků M, Avliyakulov N, Benada O. 1998. Pankinetoplast DNA structure in a primitive bodonid flagellate, *Cryptobia helicis*. EMBO J. 17:838-846.