Supplementary Materials

P.fragile_cox3	20 40 60 80 100	0 120 140	160 180	200 220 240 260	
B benerals	xx3 ···································	KUTNUN IHHNESNYNNYKAHUISYPSUTSUYGTSUKYESY-GUUET- ENP	INCLIEVES IR. ESH	SOMESEKESEAULEETYEWGELHESU- SPYP- UYN HGEELTSSRMUELTETETELASASCHTACUCHLEEKGI	MS
P.Knowiesi_cox3	xx3 ······	KITHUN HEFTESNYNNYKAHUYSYPSUTSUYGTSUKYESY-GUUET- ENP	TRANEWYS NR-ESH NSWESSUM	SOMESEERSEALLEEETEENGELHESE-SPEP-EENEGERETSSRMUELTETEEASASCMTACEOHUEEKG	MS 8 55 152
P.cynomolgi_cox3	xx3	KITYEIN HERIESNYNNYKAHEYSYPSETSEYG <mark>TSEKYESY</mark> -G elet - Enp	NTEXTEVYS	SOMESEEMSEATLEFETYEWGELHESE-SPYP-LYNEGENETSSRMENETETETELASASCHTACEOPETEKG	MS
P.inui_cox3	x3	KITYLIN HERIESNENNIKAHLISYPSETSENGTSEKNESN- GILET- ENP	TILLIEVYS	SOMESEES SEATCHEET YEWGELHESE. SPEP. LYNEGENETSSRMENETETEN LASASCHTACLOPLERKG	MS. 5 152
P.simium_cox3	xi)		TREATEVYS	SOMESTERSEATER TYPE WOLLHESE. SPRP. EXNEGETTSSRMETETETELASASCHTAC CHEEKG	MSLEISS 138
P.ovale_curtisi_cox3	xx3 METAH	KETNEEN HEFERNYNNYKAHEESYPSEASEYGTSEKYFSY- GELET- ENP	TREATEVYS	SOMESETESEALLEETTYENGELHESE-SPSP-EYNEGENETSSRMEELTETENEASASCHTACEOHENEKG	MS
P.malariae_cox3	nci	KITYEIN IMPRENNYNNYKAHEISYPSETSENGTSEKNESY-CHEMT- ENP	INCLINYS	SOMEST FISTALLEFT YEWCHINESE- SPCP- LYD	MS
P.falciparum_cox3	xx]		TTULTEVYS	SONESEVESEALLEETYNWGELHESE. SPWP-ESNEGENETSSRMEELTETHEASASCHTAC.OVHERKO	MS
P.billbrayi_cox3	xx3	SETREENNMERSERSSESNEKAHENSEPAETSENGTTEKNESS-CEERT-ENP	EXMXXMVXS	SCHEST HISEALLINET YNWCHEHNSE- SPWP- CCS COMMETSSHMENETET THEEASASCETACECENTERS	MS
Earanadina cov3			TURINET	IGU EXTELLED AND A VISION STATE SPECIAL SET OF A STATE AND A ST	
E annaren can?					
E-pratox_cox5			THEFT AND A CONTRACT OF A CONTRACT		
E.mitis_cox3	xx;]	WWWNEREKKEYSNCSYLKETKESELWATTERVETW CELESPEL	THREEFNET	EGYESTELLERVSRWGA-NTCLESPSK-ITOTTLESPERGEVSESSNGELVTETHLESSASV EGYCRETSERA	TENECH 150
E.necatrix_cox3	xx]	······································	THELLENET	IGH ISTELLERYSFRWCA- WSSILSPSN- ITOTTLESPTECENSISSICE. ITELESTASY ICHCALTSEKA	ENENECK 149
E.tenella_cox3	xx3	······································	THREEFNET	UGY ESTELLERYSREWGA- MSSILSPSM- WTOTTLESPTEGUESISSICENTITHLESTASMILGYGALTSEKA	INENEOK 149
E.maxima_cox3	xx3		THEYERNET	UGYESTELLERISHEWGA-NSSEESPSN-HADITESSPUEGUSSESNNCLEVIETHLESSASUNLONCULTSEKA	EVENECK 149
E.brunetti_cox3	xx3	MWERENERKEYSNYSYERENERENATTERYNT - CHENSPHIE	THEREENET	LON ESTELLERNSHEWGA- HTCLESPSN- ETOTTENSPTECENSISSICENTITELEASASENEGROWNTSEKA	TENEOR 150
8.cabali_cox3	ax3	ME- MERNOSENATSEOSASE- CERNERS-	······································	- OM THE SETTICE SCHEWARDHARW- ONDE- HOT PENER PENER TSTE NAASSETSEDNHKNATS	NSDAEK 117
B.bigemina_cox3	xx3	MI- NERNOTERSTSEQSASE- GEVELVS	······································	. ON	ERSESEK 117
8.bovis_cox3	xx3	MI- MEKNHSERATSEQSESE- GERREES		· OM· · EXESTICATEDETWONTHARW· ONTE· TOM· · · · · PUNCE· · PYUNE· · · TSSE- · NETSSESTIVEMENTES	HESDEEK 117
B.gibsoni_cox3	xx3	MITHIEHNSENATSEKAASE- GEVENES-	······································	- OM	NSETEK 118
B.rodhaini.cov3	xx3	NI. FILLFINERSINSSTEREESI. GEREGAS.	·· DETERNT. ····· EQ. ETS. ···· ETERNERS	LENS. NILSESNIFUSILUST LULLI. NKO	SENSEA 123
R mirmti cov2	xx3	MITTELEVEN INTERSET ANTERSE	······································	IVSS. VIUSELLUFASUIVSVUHSTV. CRD	SEWSKEN 124
T orleast in and	w3				
		MERENAULE OF ENGINE			118
T.parva_cox3	wi	MEMERGATIERYESY GERENNS.	TATELNS	· CM· · ELESELLVESTELWOVEHLEN. SN· · · · · · · PILLAELNVEAMUONSDVU · · NTOSILVSIILHENGES	ANEETOE 118
T.annulata_cox3	DX3	TENERYSTOEDTERYIDSTYKNEIIMYNNCHEENGTTERNESN- GEFFENS		. OM HIH SEIMWEST HIWGEHHERD. SN	GYFEYOY 160
Durinskia_baltica_cox3	xx3 ···································	H LEYA SOSECEL IN SSWELVE I LYVER I FSGEN LYCWKGI HESXNN FLI FLL	UHIESEHWERDU ···· UR· ELD···· KNNEQUUTTEE	LYPELIELYSECULEYSERWISEHSES SPI COS WPOEGHYLPEPCELIEANTELLSNAALSEGNAEISEELS	SOFSIFF 181
Karlodinium_veneficum_cox3	2X3	VASCPECILINSPWISTEALLEELYTIGINEYCWNGIHESWSI: DETEECLES	GUENTUHEWERDU····UR·EEN····KKYELUUMVUE	ENFREE ASEGULFESHEWASEHSUS. SPS NOW WPAEGENNPORCELTEANTULESNAANSEGGAENSEERS	SSEWVILE 173
Hematodinium_spcox3	xx3 N¥NEN¥KKE	USSNPRUSUINCPWEURNNUESEWKSTUS-SUSTSUGRU	YEEVALNYWERDY ···· HR· EYE···· KKYEYEVIMLE	NYRETRYRSEGULEUSREWYSEHSSCSUNWTOGNYYPDPCEUTNTTEUESNAAUSEGCAENCEEAA	000 N- 150
Oxyrrhis_marina_cob/cox3	DX3 UFTUHUHMEHSRSHASENASENSKHESEHSEHSEHSEFSSUTEOSCUPHEUFTYYCHUHTUHMUUSHYPKOHHUTKUKCHNNUSSNSSSANFUPANRTUAASUF	YSSHPECULINSP STUFUAUUOUNSNUAUOYYUERAUSU FEUPEEU	TH. HUENNWIRDI WR. ESS KNCSUTEMULA	HANSERLINSELCHERSERWAAHHSTH. SPMI. U	LISANA 247
Kryptoperidinium_foliaceum_cox3	xx3	ELASWGAENEYTGGELYMHKMEGGWGEEEN- GE	ETTERMERWURDT VR. EATRECHHTERRORGERE	· GMELEEUSEUMEEEAEUWAEEHSSE· APUUN GS· · · SWPPHAEETMNT HTEPETNTEELESSGAT HTWAHHAUEUN.	AKRHTII 146
Chromera_velia_cox3_ORF	DRF	SRPATY SEDENEEHYYY NY MY CHYLEYY YN YNN STHEFMER. FYFRIMN H	NGMSWESTLERITYE IR. QPW SSK SE SWEEPA	motif 2 motif 2 	MDISNY 150
Vitrella_cox3_ORF	DRF RTWGHLERREAWRUHNSEYSNCPUAWTSAUNRSHKDQEUSUUSQGAUSAUFSKURRWCARMUSCYUYELETRSUASDHKSHSUQARWNIESQCIEUSC	KT BET EHKEYKENKOLRIISRSKOLRINKYYLGSYLFTKSIINNI YIHYNLLF		motif 2	KNKENKE 263
cha		. 6 00 7. No. 0. X . V .	- T	E FILLE STREET	
sequence logo	90 188	ĸŦŦŎĔŦŔĸĸĽŸĔŦŦĔŔŢŔŔŦĔĨĨĹŶŦŶŦĔŶŦĔĹŶĔĬĬĹŔŶĔŚŎĹŬĔĹĬŦŎĔĠĔ	ŦŸĹĨĨŀŎŶĬĸĸĸĸĸĹĬĹĹŦŦŤŸŧŹĬŦĬŠĨĨŦ	ŧĽGŴĊĸŤĬŖſĹĸĹĹĬŶĬŔĬŴĊŔĹŀŀġŚĽċſŶġŧĹŢġġŗŗijĸĸŊĸĔĊĊŦŸĔŸŔĸĹŶĔĬŶĬŀĹĹŶġŇſĔŴĹĚĔĔĸĔŖŸĔŔĔ	a de la
	200 300 320 340 340	380 400 420			
P.fragile_cox3	XX3 ARRENESECHASEGITE. NEH- BONCENBAUSOTERNENEGENBSHUNNG- LEULENNERNENENEENSESENGUSN. HEPHIOGITE	LIWHENELWWERE ERRESE			
P.knowlesi_cox3	XX3 HIFFFALLOECHASLOTTE- MH- LONCENDALSOTLEYCUTOLHESHYING- LLULLINGEREIENYDTHCEWSKSLNGESN- LUPHTRONTE		263		
P.cynomolgi_cox3	XX3 HIFFINGECHASEGTTE. HEH- ECHCHNEANSCTERNCHYCHYCEHESHWING- EELEENWEINEMKETNCEWSYSENGESN. HHEPHTEGNTE	EVWHEVEIWWERE.EFFENSE	263		
P.inul_cox3					
P.simium_cox3	XX3 HIELEYLLOECETSLOTTE. HLH- LCHYNNDAI SCTLEWCYTCHHESHYLWG- LELLLYYFIRWENYDTNCEWSYSLWGI SY- HWLPHTDCHT	FRODEXCENSIONER. BEEREOC			
	XX) EHRHIKELORCETSKATTE. KEH- EOKKINKALSOTERKOETOERSHVEVA- ERELEENSKAVERKOTNORKSEKOISK. EVEPHTEOITE XX) Evenemelorcenskatte. Keh- Eokoinkalsoterkoetoershvevao- Eleletynerhuvevertissenseskoisk. Evephteoite		249		
P.ovale_curtisi_cox3	na statistica (kurata, kurata, kurata, Distriktica (kurata, kurata, kurata Distriktica (kurata, kurata, kurat	EXAMPLE EXEMPLE EXEMPLE	····· 249 ····· 263		
P.ovale_curtisi_cox3 P.malariae_cox3	I TELEVISTE OF TELEVISTE UND TRANSPORTUNE UND TRANSPORTUNE UND TRANSPORTUNE UND TRANSPORTUNE UND TRANSPORTUNE U I I I I I I I I I I I I I I I I I I I		249 263 263		
P.ovale_curtisi_cox3 P.malariae_cox3 P.falciparum_cox3		ERFORMANNEN AUGUS S EPONEREENSEEL EPONEREENSEEL EPONEREENSEEL EPONEREENSEEL EPONEREENSEEL EPONEREENSEEL EPONEREENSEEL	249 		
P.ovale_curtisi_cox3 P.malariae_cox3 P.falciparum_cox3 P.bilibrayi_cox3			249 		
P.ovale_curtisi_cox3 P.malariae_cox3 P.falciparum_cox3 P.billbrayi_cox3 E.acervulina_cox3			249 249 263 250 259		
P.ovale_curtisi_cox3 P.malariae_cox3 P.falciparum_cox3 P.bilibrayi_cox3 E.acervulna_cox3 E.acervulna_cox3			263 263 263 263 205 205 205 205 205 205 205 205		
P.ovale_curtisi_cox3 P.malariae_cox3 P.falciparum_cox3 P.bilibray(_cox3 E.acervulina_cox3 E.praecox_cox3 E.miss_cox3			243 243 243 243 243 243 243 245 259 259 259 259		
Powale_curtisi_cox3 P.malarise_cox3 P.falciparum_cox3 P.bilibrayi_cox3 Eacervulna_cox3 Epraecox.cox3 Emitis_cox3 Emitis_cox3			243 243 253 265 255 259 259 259 259 259 259 259 259 25		
Ponile_curtisi_cox3 P.malariae_cox3 P.falciparum_cox3 P.falciparum_cox3 E.acerrvnina.cox3 E.praecox_cox3 E.mrits_cox3 E.mecatric_cox3 E.mecatric_cox3			263 263 263 263 265 255 259 266 266 269 269 269 269 279 269		
Ponile.curtisi.cox3 Philoiptum.cox3 Philoiptum.cox3 Philoiptum.cox3 Exercisine.cox3 Emitty.cox3 Enecarris.cox3 Extendit.cox3			249 243 243 243 255 259 259 259 259 259 259 259		
Ponile.curtisi.cox3 P.nilarise.cox3 P.hiloparum.cox3 P.hiloparum.cox3 E.aceronium.cox3 E.arretox.cox3 E.metis.cox3 E.metis.cox3 E.tenelia.cox3			243 243 243 243 255 255 259 266 259 259 259 259 259 259 259		
Ponik, curisi, cond Prailuita, cond Philoprum, cond Philoprum, cond Exercutina, cond Enectiva, cond Enectiva, cond Enectiva, cond Eneratima, cond Eneratima, cond Eneratima, cond			243 243 243 243 245 245 245 245 249 249 249 249 249 249 249 249		
Ponik,curisi,coxi Pinaliria.coxi Pilotopurun,coxi Pilotopurun,coxi Eaconolini.coxi Eaconolini.coxi Enectric.coxi Enectric.coxi Enectric.coxi Enectric.coxi Enectric.coxi Enumeti.coxi Enumeti.coxi Exabili.coxi			243 243 243 253 255 259 259 259 259 259 259 259		
Pank, unst, and Pinkink, gad Pikiapan, and Pikink, gad Pikink, gad Earen, gad Earenk, gad Earenk, gad Earenk, gad Earenk, gad Earenk, gad Earenk, gad Earenk, gad			243 243 243 255 255 259 220 229 229 229 229 229 229 22		
Poels, ansi, and Pratrice, and Pratrice, and Pratrice, and Pratrice, and Extended, and Extension, and Extension, and Extension, and Extension, and Extension, and Extension, and Extension, and Extension, and Extension, and			249 243 243 243 245 259 259 259 259 259 259 259 25		
Postek, conti, così Postate.com Philipuro, così Philipuro, così Lennenta, così Lennis, così			200 201 201 201 201 205 205 205 209 200 209 200 209 209 209 209		
Coo, Simo, Jahon Coo, Simo, Jahon Coo, Simo, Jahon Coo, Simona Coo			249 243 243 255 255 255 259 259 260 259 259 259 259 259 259 259 259		
Con Janua Aneo Con Janua A Con Janua C Con Annora Con Annora Con Janua Con Janua Con Janua Con Janua Con Janua Con J			200 201 201 201 202 202 202 203 203 203 203 203		
Cao, Java, Jawa Palaba, Cao, Ya Palaba, Cao, Ya Palaba, Cao, Ya Cao, San, Cao, Cao, Cao, Cao, Cao, Cao, Cao, Cao			200 201 201 201 201 205 205 205 209 200 209 209 209 209 209 209		
Cao, Jamo, Jako, Y Cao, Jamo, Jako, Y Cao, Jano, Cao, J Cao, Jano, J Cao, J Cao			200 201 201 201 201 201 202 202		
Constraint, and Constraint,			200 201 201 201 201 205 205 209 209 209 209 209 209 209 209		
Cao, Sano, Jakov, T Cao, Sano, Jakov, T Cao, Sano, Cao, Cao, Cao, Cao, Cao, Cao, Cao, Ca			249 243 243 255 259 259 259 259 259 259 259		
Con Jimo Anno Con Jimo Anno Con Annon Con Annon Con Annon Con Anno Con Anno			200 201 201 201 202 203 205 209 209 209 209 209 209 209 209		
Panka canta, casa Panka canta, casa Palabar, casa Palabar, casa Exercita, casa			249 241 243 243 243 245 245 245 249 249 249 249 249 249 249 249		
Ceo, Simo, Jaiwa, T Ceo, Simo, Jaiwa, T Ceo, Simo, S Ceo, Simo, S Ceo, Simo, S Ceo, S			249 243 243 243 243 243 243 243 243 243 243 243 243 243 243 243 243 244 245 245 249 240 249 240 242 212 213 221 213 221 213 225 225 226 227 228 242		
ետեն, որուց հայտուց հայտո հայտուց հայտուց հա			200 201 201 201 205 205 205 209 209 209 209 209 209 209 209		
Շաչնար չեստ չեստ չ Շաչությունին Շաչությունին Շաչությունին Շաչություն Casababababababababababababababababababa			249 243 243 243 245 245 245 249 249 249 249 249 249 249 249		
الجائبة المعالم المراجع الحيمة المراجع الم المراجع المراجع الم المراجع المراجع ال المراجع المراجع المراح لمراجع المراجع المراجع المراجع المراجع المراجع الم			249 243 243 243 244 245 259 259 259 259 259 260 259 259 260 259 260 259 260 260 279 260 279 260 279 260 261 279 260 279 260 261 279 260 261 279 260 260 261 261 261 261 261 261 261 261		
Vente, cente, ces Patiente, ce			203 203 203 203 204 205 205 207 208 209 209 209 209 209 209 201 202 203 204 205 205 206 201 202 203 204 205 205 201 203 204 205 205 205 205 205 205 204 205 205 205 206 207 208 209 201 202 203 204 205 206		
			249 243 243 243 244 245 259 259 259 269 269 269 269 269 269 269 26		

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 cox1-cox3* ORF sequence downstream of the conserved *cox1* fragment is shown, along with the corresponding portion of the *Vitrella cox3* ORF (see also supplementary fig. S5 showing a full *cox3* 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-cox1* fusion ORF, and *cox3* 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 -³²P]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

Commission and formand and profession TECOL	1		CHUNTDORWDUMP		ACEDARA TRODUURD	
Neceartorya fischeri NREL 181	LI STERVSEH		CWIVUDNKUWDMTDF	LEE-HDCCSTTTLEV	AGRDATAIFOPIHPP	GIVEDGEDPRAMVGE
Aspergillus flavus NRRL3357	MI EPSEVAKH	NSADS	CWVVLYGKVYDVTNF	LEN-HPGGSAAILAL	AGKDATEEYDTIHPS	GLLEEYLDPKACLGV
Ajellomyces_capsulatus_NAm1	LLTGPEIQKH	NSKNS	CWIIVHGKAFDVTEF	LPE-HPGGQDIILKY	AGKDATDEFEPIHPP	DTLDKYLDSSKHLGP
Uncinocarpus_reesii_1704	PLTGAEIAKH	<mark>NSAK</mark> S	CWVIVHGKAYDVTDF	LPE-HPGGQKIILKY	AGKDATEEFDPIHPP	D TLDKYLDPS<mark>K</mark>HL GP
Neurospora_crassa_OR74A	PLTGVEVAKH		CWVIVHGKAYDITEF	LPE-HPGGMKIILKY	AGKDATEEFDPIHPP	DTLEKYLAKDKHLGP
Talaromyces stipitatus ATCC 10500	KT TCODVAOH	NNENS	CWVIVHCEAVDVTFF	LDE-HDCCONTTLEY	ACKDATEFEEDTHDD	DTUDEVIDASEHLCE
Paracoccidioides brasiliensis Pb01	LL TGPEIOKH	NSKDS	CWVIVHGKAYDVTDF	LPE-HPGGOKIILKY	AGKDATEEFEPIHPS	DTLDKYLDPSKHLGP
Verticillium albo-atrum VaMs.102					MHSS	LLLITQLAAV
Arthroderma_otae_CBS_113480	KLTGEEIAKH	<mark>SS</mark> PSS	CWVVVHGHAYDVTDF	LPEEHPGGQQIILKH	AGKDATAEFDPVHPP	DTLDKYLDRSKHLGP
Saccharomyces_cerevisiae_S288c	KI <mark>SPAEVA</mark> KH	<mark>NK</mark> PDD	CWVVINGYVYDLTRF	LPN-HPGGQDVIKFN	AGKDVTAIFEPLHAP	NVIDKYIAPEKKLGP
Lachancea_thermotolerans	GI EVSEVMQH	<mark>NK</mark> PDD	CWIVLNGEVYDVTSF	LQM-HPGGAARIMEV	AGNDATRKFYSIHSD	STLEKMKEQLVYIGK
Trichophyton_rubrum_CBS_118892	MI DAEEVAKH		CWVVLYGKVYDVTNF	LPN-HPGGANIILQL	AGRDATEEYDPIHPT	GTLEDSLPPEALVGR
Candida_dubliniensis_CD36	VLSLQEVSQH	NKKDD	CWVIIHDKAYDLSDF	IDE-HPGGSAIIMKY	AGKDATKAFDPIHPG	DTLTKYLQPKYHKGE
Schizophyllum_commune_H4-8	MINWTLEEVKKHNT	RECPPREAPAPEPRS	CHUA INCRUVAL TER	ROD-HPGGAAIILKY	AGRDATAAYEPIHPA	DALEKNLPASAHLGP
Perkingus marinus ATCC 50983	The states of th	MIGIN	CWVALNORVIDDIEF	DFQ-HFGGRAVIDA	AGREADING VENETHEF	DATORELEADIE
Karlodinium micrum 1						
Karlodinium micrum 2						
Chromera_velia_gene4425.t1	MISFEEIAKH	<mark>NNR</mark> NS	CWVVLHGVVYDMTKF	LDE-HPGGAAIILKQ	AGKDGTKAFDAIHPR	DIV-KMLPPEAVLGK
Chromera_velia_gene5763.t1						
Vitrella_brassicatormis_gene7868.tl						
vitreila_brassicarormis_genei9079.tl						
	91					
Cryptococcus neoformans var. neoformans JEC21	VDPATLPKV		VDKKKEDGEO	RRVDLAEII GLPD	FDAAAKANLTSKAWA	YMSSGATDOY
Neosartorya fischeri NRRL 181	LDTSAIDEE		LKOPPSENPOVVLDN	EKPPLHTLI NSHD	FELVASKTASKKTWA	FYSSASTDLI
Aspergillus_flavus_NRRL3357	LSASAPEAAEPTR		SSQSASKEAS	EETPLSSLL NLAE	IEQAAKRKLSPKGWA	YYSSATDDSI
Ajellomyces_capsulatus_NAm1	VAMDTV		AVDPEEEERQFR-IQ	HMPPLEQCY NLLD	FEAVARRIMKKTAWA	YYSSGADDEM
Uncinocarpus_reesii_1704	VDMNTV	MQEDK	GVDPEEEARQER-VK	RMPLLEQCY NLMD	FEAVARRVMKKTAWG	YYSSGADDET VGQSH
Purepophora tritici-repentie Dt-1C-DED			CID-DO-ECAEVANK	TI DDI NGLV- CLDD	FOWUCENEMNASSET	VVDNCAACEW
Talaromyces stipitatus ATCC 10500	VDMSTV	EHDEK	VKDPLETEREER - TK	RMPALAOCY NLLD	FESVAREVMENTAWA	YYSSGADDEI
Paracoccidioides brasiliensis Pb01	VAMDTV		AVDPEEEEROLR - IO	RMPLLEOCY NLLD	FEAVARRILKKSAWA	YYSSGADDEI
Verticillium_albo-atrum_VaMs.102	VL	GEPWINEVDT	GLE - TYLLGTNYTEG	TLPLLKDIR ALPD	FDWAARQHLNDQMYS	FYRTGTAGEQ
Arthroderma_otae_CBS_113480	VDMASM	<mark>AQ</mark> EKK	AVDPEEEQRLER - IR	RMPSLDQCY NLMD	FEAVASRVMKKTAWG	YYSSGTEDEM
saccharomyces_cerevisiae_S288c	LOGSMPPELVC	YAP	GETKEDIARKEQLKS	LLPPLDNII NLYD	FEYLASQTLTKQAWA	YYSSGANDEV
Lachancea_thermotolerans	UDEKCI NELADED		TIS-EEDIRIRESRK	CODDI CTU	FEAVAREVLPRSTYA	YFATGSSDEF
Candida dubliniensis CD36	VEKKKKKKKKCDCTTTT	NNDADEYP DDDDD	PPTEERLEPROL WE	NEPDISOLY	FEFVARHTMDDICHA	YYSSSADCRA
Schizophyllum commune H4-8	VNTDAASTLEAAV	RNR	RKTK-DEIRIEEAHK	YKPALSHVL SLRD	MEDVAKKVLPHKAYA	YYSSSTEDTI
Komagataella pastoris GS115	LDGELP	BEEE	VLTDAEIERQER-IA	NKPPLSSMF NVYD	FEYVAONILDEAAWA	YYSSAADDEI
Perkinsus marinus ATCC 50983				ML NAMD	FKAVAEKKVSAEGWA	YLYSAAGDEF
Karlodinium_micrum_1		<mark>EEPLEPV</mark>	EDEEE <mark>QV</mark> EEE <mark>KA</mark> - AD	DTPDISQMV NVWD	FEVIAKENVTKEAWA	YLMSGADDEI
Karlodinium_micrum_2				AHMLNLND	FERAAAKSMDSQGYG	YYASAANDEV
Chromera_velia_gene4425.tl	ADMSTAPAPSAEA	<mark>BT</mark> -	-APPSGPASA-AA	EKPPIDHML NIFD	FESLAQQSMQKPGWD	YYCSGADDEK
Vitralla braccicaformic cone7868 ±1			MAN	CANCOPUDI - MIDD	VENTAREPITTOAUC	VECCONDON
Vitrella brassicaformis gene19079.tl			-MSMSRSK	ASAEESVPVNLDD	FERIADKRLTKAVHG	YYSSGANDEV
1	81					
Cryptococcus_neoformans_varneoformans_JEC21	TLD	LNRKAFNSILFRPRV	LVDVE - IA DTRTQ	MLGQDTSLPIFISPA	GMAKLAHP-EGECLL	AKAAGQSNII
Neosartorya_fischeri_NRRL_181	TRD	ANKSCFDRIWFRPRV	LRNVR-SVDSRTR	ILGVDCSMPLFVSPA	AMAKLIHP-DGECAI	ARACER KGIM
Aspergillus_flavus_NRRL3357	TIN	YNNLIYRSILLRPRV	FIDCR-ECDLSTR	FLGLKLGLPIYISPA	AMARLAHP-QGEAGI	AAACRKFGAM
Uncinocarpus reesii 1704	ARETPTADNCKOTMR	ENHSAFHKTWFRPRT	LVDVE-NVDISTT	MLGTPVSVPFIVIAT	ALGKIGHA - DGEVCL	TKAAASHDVV
Neurospora crassa OR74A	TLR	ENHAAFHRIWFRPKV	LVDVE - KV DFSTT	MLGTKVDIPFYVTAT	ALGKLGHV-EGEVLL	TRAAKKHNVV
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	TKAAHNHEVI
Paracoccidioides_brasiliensis_Pb01	SLR	ENHSAFHKIWFRPRV	LVDVQ-NVDITST	MLGTPVSAPFYVTAA	ALGKLGHP-EGEVCL	TRAANTHNII
Verticillium_albo-atrum_VaMs.102	SYR	HNLDIWKSVQFRSRH	LSDVT-KLNETLATT	ILGYNFSAPFFIAPA	ARGVYGDPERAELNF	VEAAGKENIL
Saccharomyces cerevisiae S288c	THP	FNHNAVHRIFFKOKI		MLGSHUDVDEVVSAT	ALCKLONPLECEKDV	ARCCCOCV
Lachancea thermotolerans	SIR	ENHYAYSRVFFKPMI	LOENEYDV DTSTE	FLGSKVSLPVYISAF	AGSKWAHP-LAELNL	OSAAYE ADIM
Trichophyton rubrum CBS 118892	SKS	LNNSIYRSILLRPRV	FVDCK-NCSLATS	MLGYKLNTPIIASPT	AMARLAHP-SGEAGI	AAACAKFGAM
Candida_dubliniensis_CD36	TFR	LNTGSYQRILFKPRV	MIDVT-EVDTSTT	MLGTNVSAPFYITAT	ALGKLGHP-DGEKVL	TRGAYKHDII
Schizophyllum_commune_H4-8	<mark>SKN</mark>	ENAHAFS <mark>R</mark> FFFHARV	MRPVS-RCDPSTT	IL <mark>GF<mark>K</mark>SSIPVFVSGA</mark>	AMAKLAHP-DGELNI	TRGCAAQGII
Komagataella_pastoris_GS115	TIR	ENHFAYHKVFFRPRI	LVDVT-NIELETE	MLGIKTSAPFYISAT	ALAKLGHP-EGEVGI	AKGAGR GDII
Perkinsus_marinus_ATCC_50983	SYA	ENEDAFSRIALVPRV	LVDVS-KVDCSSS	VLGRHFDVPFYMTAV	AMAKLYNV-DGEKCV	ARGIGKTKEAGIDMA
Karlodinium micrum 2	TKR	DNCAAFSRAWLKPRV	MRNVL-STNTRCT	LIGTEFAFPIFISPA	AMAGLAHE - DAEPAL	ARAAGK I.GAL
Chromera velia gene4425.tl	TLR	ENKETFONIWLVPRV	LRDVS-SIDFATT	MLGCSSSIPVYLTAT	AVAKLAHP-DGELGI	IRAAHAHGLV
Chromera velia gene5763.tl						
Vitrella_brassicaformis_gene7868.t1	TLR	ENREAFNRIRLRPRV	LNDVS - KV NLNTS	ILGHTSTMPVLVAPT	GMQKLAHP-EGETAT	ARAARE VGVP
Vitrella_brassicaformis_gene19079.t1	<mark>SLR</mark>	ENRVAFERIRLRPRF	LIDVD - NI QLQTS	ILGYNVSMPLIVSPT	AMQKMAHP-DGETGV	AIACKEMGIP
Cruptococcus peoformans war peoformans TEC21	OMTOTNACADI DOTT	CON	MOL VUDDND DETEC	LICETN-SICIENTE	WWWDA DA DCEPEADE	POPARU - PUACCT
Neosartorva fischeri NRRL 181	OGVSNNSSYTLDOLK	EAAPSANEE	FOLYVNRDRSKSAA-	LLHOCSANPNVKAIF	VTVDAAWPGKREADE	RVKADE NLSVPM
Aspergillus flavus NRRL3357	QLISHNASMTTOOIV	ANAH PDOIFG	WOLYCLKDVKRSEK-	RIAEINSIKEIKFIC	LTLDAPFPGKREIEE	ROKMEEL
Ajellomyces capsulatus NAm1	QMI PTLASCSFDEIV	DARGPDQVQW	LQLYVNKDRTITKR-	IVQHAQ-QRGCKALF	ITVDAPQLGRREKDM	RSKFSDRGS-AVQAA
Uncinocarpus_reesii_1704	QMI PTLASCSFDEIV	DAAI DKQTQW	LQLYVNKDRDITRK-	IVNHAE-KRGCKGLF	ITVDAPQLGRREKDM	RSKFSDPGS - DVQQT
Neurospora_crassa_OR74A	QMI PTLASCAFDE IM	DAAE GDQVQW	LQLYVNKDRAITER-	IIKHAE-KRGCKALF	ITVDAPQLGRREKDM	RVKFTDDGS-NVQK-
Pyrenophora_tritici-repentis_Pt-IC-BFP	INPSLISSIPMEDIX	DARASKDQVMF	QQIYLNGNLSSTKA-	LFDKAK-SLGAKGLV	TTVDSAGSATRHRAA	RIGVGSANT
Paracoccidioides brasiliensis Pb01	OMIPTLASCSFDFIV	DARGPNOTOW	LOLYVNKDRGITKR	IVOHAE - KRGCKALF	ITVDAPOLGRREKDM	RTKFSDRGS - DVOAS
Verticillium albo-atrum VaMs.102	YIPSMYASKTIEEIA	AGKSNSTLNGPOVIF	OOIYTNANLSVTWD-	NIRRAE-RTGAKAIV	FTIDAPGNSVRHRAA	RYDTTNANS
Arthroderma_otae_CBS_113480	QMI PTLASCSFDQIV	DAKT PROTOW	LQLYVNKDRDITRR-	IVEHAE - ARGCKGLF	ITVDAPQLGRREKDM	RSKFAEQGS-NVQAS
Saccharomyces_cerevisiae_S288c	QMISTLASCSPEEII	EAAPSDKQIQW	YQLYVNSDRKITDD-	LV <mark>KNVE - KL</mark> GVKALF	VTVDAPSLGQREKDM	KLKFSNTKAGP-KAM
Lachancea_thermotolerans	QMVPKQNSYSIEEFY	ENVPEDQKHW	SQYHFDSREEFNETL	IKKLEA-QPSVKALF	LNVDLRDIGNREKDS	RQRALDVESSKSLSA
Candida dubliniangia CBS_118892	QIISNNASMTPEEIV	KGAT PDQVFG	WQLYVQIERKKSEA-	MLARINKI KAIKFIC	LTLDAPVPGKRELDE	RTKAIASTPAVADIV
Candida_dubiiniensis_CD36	OMUGGNAGVGVAFTA	UDDRPNOTOW	FOLVENEDRIALO	TUQHAE - ARGMRGLF	LTUDATUDGKREKDM	RIKSIVULS-FVQGE
Komagataella pastoris GS115	OMISTLASCSLDETV	AAAK EGOSOW	FOLYVNSDREVAYN-	MIKHCE-ELGIKGIF	VTVDAPSLGNREKDR	RMKETEDTDVDLSGD
Perkingung marinug ATCC E0083	YMI PTLASCGNGEFY	GGLK		DKAD-AR	GRRMD-V	FLKGI
Ferkinsus_marinus_Arcc_50905	OT COMPLETE OCCUPATION	AARS PGOTOW	WOLVVNKDRELTKT-	VVOKAR-SLGENALE	ITVDAPQLGRRERDM	RNKAKMSAN VQTK
Karlodinium_micrum_1	QLCPTLASCIMDEMH				THE PARTY OF THE PARTY	PNKUTDSSNLSLVOK
Karlodinium_micrum_1 Karlodinium_micrum_2	ULCPTLASCIMDEMH HVVANMASRELEEIT	DARVPGQTQW	YQIYVNPERSKTEA-	II <mark>KR</mark> AV-QAGVKALL	VIVDIPOLGRRERDM	Alter TDDDDHIDH VQR
Karlodinium_micrum_1 Karlodinium_micrum_2 Chromera_velia_gene4425.t1	OLCPTLASCIMDEMH HVVANMAS <mark>R</mark> ELEEIT YML <mark>PTLSSYAFDD</mark> ML	DARVPGQTQW DAIQPGQPFW	YQIYVNPERSKTEA- MQLYVNRDRNMTAE-	II <mark>KRAV-QAGVK</mark> ALL HV <mark>RKAA-KRGCRALC</mark>	VTVDTPQLGRRERDM	RNKLTKTEA - SIQKD
Ferfilisus marinus file 5050 Karlodinium micrum 1 Karlodinium micrum 2 Chromera, velia gene4425.t1 Chromera, velia gene5763.t1 Vitrella brassicaformic gene7062 t1	OLCPTLASCIMDEMH HVVANMASRELEEIT YMLPTLSSYAFDDML	DARVPGQTQW DAIQPGQPFW	YQIYVNPERSKTEA- MQLYVNRDRNMTAE-	IIKRAV-QAGVKALL HVRKAA-KRGCRALC	VTVDTPOLGRRERDM VTVDAPQLGRRERDM VDLPPLGHREADI	RNKLTKTEA-SIQKD RNRFHLPAPLRLENF RNHFTL PDU
Karlodinium micrum 1 Karlodinium micrum 2 Chromera velia gene4425.t1 Chromera velia gene5763.t1 Vitrella brassicaformis gene7868.t1 Vitrella brassicaformis gene19070 t1	QLCPTLASCIMDEMH HVVANMASRELEEIT YMLPTLSSYAFDDML MTLSTMSTTPLEEVA MTLSTLATTCVTDVC	DARVPGQTOW DAIQPGQPFW -ESSGKGLRF ANCGGFGLPF	YQIYVNPERSKTEA- MQLYVNRDRNMTAE- FQLYVFQQRELTVQ- FQLYTFEDRCTTVA-	IIKRAV-QAGVKALL HVRKAA-KRGCRALC LIKRAE-AAGYSAIV	VTVDTPQLGRRERDM VTVDAPQLGRRERDM VDLPPLGHREADI LTVDVPAMGRREADL LTVDTPTLGPRNAPT	RNKLTKTEA-SIQKD RNRFHLPAPLRLENF RNHFTLPPHLTLANF RNOFSLPTGLGLANF
FarAnisda marinda Arco-1995 Karlodinium micrum 1 Karlodinium Terum 2 Chromera_velia_gene4425.t1 Chromera_velia_gene5763.t1 Vitrella_brassicaformis_gene19079.t1	QLEPTLASCIMDEMH HVVANMASRELEEIT YMLPTLSSYAFDDML MTLSTMSTTPLEEVA MTLSTLATTSVTDVS	DARVPGQTQW DAIQPGQPFW -ESSGKGLRF ANCGGEGLRF	YQIYVNPERSKTEA- MQLYVNRDRNMTAE- FQLYVFQQRELTVQ- FQLYTFRDRGITYD-	IIRRAV - QAGVKALL HVRKAA - KRGCRALC LIKRAE - AAGYSAIV LVKRAE - AAGFKALV	VIVDIPQLGRRERDM VIVDAPQLGRRERDM VDLPPLGHREADI LIVDVPAMGRREADL LIVDTPILGRRNADI	RNKLTKTEA-SIQKD RNRFHLPAPLRLENF RNHFTLPPHLTLANF RNQFSLPTGLSLANF
FarAnisda marinus Arcoussos Karlodinium micrum 1 Karlodinium micrum 2 Chromera velia gene4425.t1 Chromera velia gene5763.t1 Vitrella brassicaformis gene7868.t1 Vitrella brassicaformis gene19079.t1 3	HVVANMASELEEIT HVVANMASELEEIT MILPTLSSYAFDDML MTLSTMSTTPLEEVA MTLSTLATTSVTDVS 61	DARV PGOTOW DAIO PGOPFW - ESS GKGLRF ANCG GEGLRF	YQIYVNPERSKTEA- MQLYVNRDRNMTAE- FQLYVFQQRELTVQ- FQLYTFRDRGITYD-	IIKRAV-QAGVKALL HVRKAA-KRGCRALC LIKRAE-AAGYSAIV LVKRAE-AAGFKALV	VIVDIPQLGRRERDM VDLPPLGRRERDM LIVDVPAMGRREADI LIVDVPAMGRREADI LIVDTPTLGRRNADI	RNKLTKTEA-SIQKD RNRFHLPAPLRLENF RNHFTLPPHLTLANF RNQFSLPTGLSLANF
FarAnisda marinda Arcousdou Karlodinium micrum_1 Karlodinium crum_2 Chromera_velia_gene4425.t1 Chromera_velia_gene5763.t1 Vitrella_brassicaformis_gene1868.t1 Vitrella_brassicaformis_gene19079.t1 3 Cryptococcus_neoformans_varneoformans_JEC21	HVVANNASELEEIT YMLPTLSSYAPDDML MTLSTMSTTPLEEVA MTLSTLATTSVTDVS 61 SGG	DARVPGOTOW DATOPGOPFW -ESSGEGLRF ANCGGEGLRF	YQIYVNPERSKTEA- MQLYVNRDRNMTAE- FQLYVFQQRELTVQ- FQLYTFRDRGITYD- GGGIGE	IIKRAV-QAGVKALL HVRKAA-KRGCRALC LIKRAE-AAGYSAIV LVKRAE-AAGFKALV SVGGFIDPKLSWK-D	VIVDIPDGSKERDM VDIPPGSKERDM LIVDVPAMGREADI LIVDTPILGRRNADI IEWLRQHTKL	RNKLTKTEA-SIQKD RNRFHLPAPLRLENF RNHFTLPPHLTLANF RNQFSLPTGLSLANF PIGLKGVQTAEDAMK
Arlodinium micrum 1 Karlodinium micrum 1 Karlodinium micrum 2 Chromera_velia_gene4425.tl Chromera_velia_gene5763.tl Vitrella_brassicaformis_gene7868.tl Vitrella_brassicaformis_gene19079.tl Cryptoccoccus_neoformans_varneoformans_JEC21 Neosartorya_fischeri_NRRL_181	OLCPTLASCIMDEMH HVVANMASELEEIT YMLPTLSSYAPDDML MTLSTMSTTPLEEVA MTLSTLATTSVTDVS 61 SGGA	DANU PGCTON DAIO PGOPFW -ESS CKGLEF ANCG GEGLEF KIGSDNK	YQIYVNPERSKTEA- MQLYVNRDRNMTAE- FOLYVFOCRELITVO- FOLYTFRDRGITYD- GGGIGR GGGIGR	IIKRAV-QAGVKALL HVRKAA-KRGCRALC LIKRAE-AAGYSAIV LVKRAE-AAGFKALV SVGGFIDPKLSWK-D VMAGFIDPGLTWD-D	VIVDIPUSCHERDM VDLPPLGHREADI LIVVDVPAMCREADL LIVVDTPILGRRNADI IEWLRQHTKL LVWVRKHTHL	RNKLIKTEA-SIOKD RNEFHLPAPLRLENF RNHFTLPPHLTLANF RNQFSLPTGLSLANF PIGLKGVQTAEDAMK PVGLKGVMSADDAIL
SarAdishum micrum_1 Karlodishum micrum_2 Chromera_vella_gene4425.tl Chromera_vella_gene425.tl Vitrella_brassicaformis_gene1868.tl Vitrella_brassicaformis_gene19079.tl Cryptococcus_neoformans_var.neoformans_JEC21 Neosartorya_fischeri_NRRL181 Aspergillus_flavus_NRRL357	DUCPTLASCIMDEMH HVVANMASRELELIT YMLPTLSSYAFDDML MTLSTMSTTPLEEVA MTLSTLATTSVTDVS 61 SGG	DARY PGCTON DAIO PGCPFW -ESS GEGLRF RNCG GEGLRF 	YQIYVNERSKTEA- MQLYVNEDENMIRE- FOLYVFQORELTVQ- FQLYTFRDEGITYD- GGGIGE GGGIGE AG	IIKRAV-QAGVKALL HVRAA-KRGCRALC LIKRAE-AAGYSAIV LVKRAE-AAGYKALV SVGGPIDPKLSWK-D VMAGPIDPGLTWD-D PQVWGTDASLTWENT	VTVDIPUGERERDM VDLPPLGHREADI LTVDVPAMCREADI LTVDVPTLGRRADI IEWLROHTKL LVWWRKHTHL LNWLRMHTSL	RNKUTKTEA-SICKD RNRFHLPAPLRLENF RNRFHLPAPLRLENF RNQFSLPTGLSLANF PIGLKGVQTAEDAMK PVCLKGVMSADDAIL PIVLKGIQTYEDAIL
Arlodinium micrum_1 Karlodinium micrum_2 Chromera_velia_gene4425.tl Chromera_velia_gene4425.tl Vitrella_brassicaformis_gene7868.tl Vitrella_brassicaformis_gene19079.tl Cryptococcus_neoformans_varneoformans_JEC21 Neosartorya_fischeri_NREL_181 Aspergillus_flavus_NRRL3357 Ajellomyces_capsulatus_NAm1 Uncincarrumg_reseii_1704	QLCPTLASCHMDEMH HVVANMASRELEDIT MTLSTNSTFFLERVA MTLSTLATTSVTDVS 61 SGG APA DGB	DARY PGOTON DAIO PGOPPW -ESS	YOIYVNERSKTEA MOLYVNERSKTEA FOLYVFORELIVO FOLYTFEDEGITYD GG	IIKRAV-QACVKALL HVRKAA-KRGCRALC LIKRAE-AAGYKALV LVKRAE-AAGYKALV SVGGPIDPKLSWK-D UMAGFIDFGLTWD-D PQVWGTDALTWEET AISSFIDPELSWK-D AISSFIDPELSWK-D	VTVDPUGGREEDM -VDLPPLGREEDDI LTVDVPAMGREEDDI LTVDVPIGREEDDI LTVDTPTLGREADL LTVDTPTLGREADL LVWVR	RNKUTKTEA-SICKD RNRFHLPALLENF RNHFTLPHILINF RNPSLPTGLSLANF PIGLKGVCTAEDAMK PVCLKGVCTAEDAML PIVLKGICTYEDAIL PIVLKGICTYEDAIL PILLKGVCRVDDVLR PILLKGVCRVDDVLR
Karlodinim minten_1_ Karlodinim minten_1 Chromera_velia_gene425.tl Chromera_velia_gene575.tl Chromera_velia_gene575.tl Vitrella_brassicaformis_gene19079.tl Zryptococccus_neoformans_var.neoformans_JEC21 Aceosartorya_fischeriNRELJa5 Aspergillus_flavus_NREL3357 Ajellomyces_capsulatus NAM1 Uncinocarpus_reesii_1704 Neurospora_crassa_0Z74A	DCCPTLASCIMDEMH HVVANMASRELEEIT YMLPTLSSVAFDDML MTLSTMSTTPLEEVA MTLSTLATTSVTDVS 61 86	DARY	VOIVAIPERSETEA MOLYVNEDENMIAE FOLYVFORELIVO FOLYFROEGITYD GG	IIERAV - QAGVEALL HVRKAA - KRGCEALC LIERAE - AAGVEALV LVERAE - AAGVEALV SVGGPIDPELSWE - D AISSFIDPELSWE - D AISSFIDPELSWE - D AISSFIDPELSWE - D	VTVDIPGGREERDM -VDLPPGGREERDM LTVDVPMGREADI LTVDTPTLGRRNADI LVWWR	RNKLIKTEA-SIGED RNETHLPAPLRLENF RNHTLPPHLTLANF RNQFSLPTGLSLANF PVCLKGVMSADDAIL PIVLKGIGTYEDAIL PIVLKGVQVDDUL PILLKGVQVDDUL PILLKGVQVDDUL
Arlodinium micrum_1 Karlodinium micrum_2 Chromera_velia_gene4425.t1 Chromera_velia_gene4425.t1 Vitrella_brassicaformis_gene7868.t1 Vitrella_brassicaformis_gene19079.t1 Cryptococcus_neoformans_var. neoformans_JEC21 Aspergillus_flavus_MRRL3357 Ajellomyces_capsulātus_NAm1 Uncinocarpus_ressi! 1704 Neurospora_crassa_OR74A Pyrenophora_tritici-repentis Pt-1C-BFP	DLOFILASCIPUERA HVVANNASEBLEET YMLPTLSSVAPDONL MTLSTINSTPLEEVA MTLSTLATSVTDVS G1 SG - G1 SG - G - G - G	DAU	YOIYWIPESSTER MQLYVNBDRMMTAE FOLYVFRORSLITVO- FOLYVFRORSITVD- GGGLG SG	IIERAY-ORGYEALL HYRAA-REGCEALC LINRAE-AAGYSAIV UNERAE-AAGYKALV SVGGPIDPKLSWE-D PQVWGTDASLTWERT AISSTIDPSLSWK-D AISSTIDPSLSWK-D AISSTIDPLSWK-D MISUTUPE-UNERT	VIVDIPOLSREERD VIVDAPOLSREERD LIVDPPLSHERADI LIVDTPILGREADI LIVUTTIGREADI LIVUTTIGREADI LIVUTTIGREADI LIVUTTIGREADI IEWLEOHTKL LNWLESITEN IPWFQSITEN IPWFQSITEN IPWFQSITEN	RNKLIKTER-SIORD RNEFHLPR-LELEN RNFTLPR-LILANF RNFTLPR-LILANF PVCLKOVMSADDAIL PIVLKKOVGVDOUL PIVLKKOVGVDOUL PILLKGVQRUDVLR PILLKGVQRUDVLR PILLKGVQRUDVLR PILLKGVQRUDVLR
Arlodinium micrum_1 Karlodinium micrum_2 Chromera_velia_gene4425.tl Chromera_velia_gene4425.tl Vitrella_brassicaformis_gene7868.tl Vitrella_brassicaformis_gene7868.tl Cryptococcus_neoformans_varneoformans_JEC21 Neosartorya_fischeri_NREL1357 Aspergillus_flavus_NREL3357 Ajellomyces_capsulatus_NAm1 Uncinocarpus_resii_1704 Neurospora_crassa_0R74A Pyrenophora_tritici_repentis_Pt-1C-BFP Talaromyces_stipitatus_ATCC_10500	CLOP LANGE THURSHING WYONANA SELECT YML PILSE VAPDOWL MILSILATEV VIDY 61 EG. G. G. J. DG. B. 	DALU POOPO DALO POOPO ESS. GEOLO ANG GEOLO ANG GEOLO ANG GEOLO DALU CARDO DISVI HEINS	VOIVAPEESSTER MOLVAPEESSTER FOLVAPEELIVA OLVAPEELIVA GC GLG GC GLG GC GLG GC GLG SC GLA SC GLA SC GLA SC GLA SC GLA SC GLA	II DEAN - OAGVEALL HVINAA - REGEALC II GAR - AAGVEALV LI GAR - AAGVEALV VIRAE - AAGVEALV SVGGEIDELING - D PUWGGEDASLIWED 2000 CDASLIWED AISSTIDELING - D AISSTIDELING - D AISSTIDELING - D RELING - V AISSTIDELING - D - CALLING - V - CALLING - V	VIDE DURBERN -VDLPEGEREN LIVOTPILGERNADI IEWLEOHTKL LIVOTPILGERNADI IEWLEOHTKL LIVOTPILGERNADI IEWLEOHTKL IEWLE	RNLITTER STORD RNFHLPARLER RNFTLPARLTANF RNFTLPARLTANF PIGLKOVGTAEDAMK PVCLKOVMSADDATL PIVLKGOVGVDOVLE PIVLKGOVGVDOVLE PILLKOVGVDOVLE PILLKOVGVDOVLE PILLKOVGVDOVLE PILLKOVGVDOVLE
Arlodinium micrum_1 Karlodinium micrum_1 Karlodinium crum_1 Chromera_velia_gene4425.tl Chromera_velia_gene4425.tl Vitrella_brassicaformis_gene1868.tl Vitrella_brassicaformis_gene19079.tl Reosartorya_fischeri_NREL_181 Aspergillus_flavus_NREL357 Ajellomyces_capsulatus_NAM1 Uncinocarpus_reesi1_1704 Neurospora_crassa_0R74A Pyrenophora_tritLof-repentis_Pt-1C-BFP Talaromyces_stipitatus_ATCC_10500 Faracoccidicides_brasiliensis_Pb01	DLOFILASCIPULSH HVVANASEBLEET YMLFILSVAFDONL MILSTLATSVTDVS 61 SG GG GG AP	DAU - POOPO DALQ - POOPO -ESE - GROLK ANGE - GROLK - RIGSDNE - OANDEK - OAN	YOIYVAPEESETEA MQLYVARDENMTAE FOLYVFOCRELIVO FOLYVFOCRELIVO FOLYVFROEGITYD GG GIGE GG GIGE GG GIGE GG GIGE GG GARE GG GARE GG GARE GG GARE GG GARE GG GARE	II LEAN - OAGVEALL HVIKAA - ERGCEALC LIKEAE - AAGVSALV LVIEAE - AAGVSALV LVIEAE - AAGVSALV LVIEAE - AAGVSALV VIAGTIDELING - D POWGDDALLWED AISSTIDELING - D AISSTIDELING - D AISSTIDELING - D AISSTIDELING - D AISSTIDELING - D	MANDA COLOR EDU - NDLPJCH EDU INNOVALCE EDU INNOVALE EDU INNOVALCE EDU INNOVALCE EDU INNOVALCE EDU INNOV	NALITYTER- STOR NURTLPPHLTAIP NURTLPPHLTAIP PUCLEOWRADDAIL STULESUNGADDAIL STULESUNGADDAI STULESUNGADDAI STULESUNGADDAI STULESUNGADDAI STULESUNGADDAI STULESUNGADDAI STULESUNGADDAI STULESUNGADDAI STULESUNGADDAI STULESUNGADDAI STULESUNGADDAI
Rarlodinium micrum_1 Karlodinium micrum_2 Chromera_velia_gene4425.tl Chromera_velia_gene4425.tl Vitrella_brassicaformis_gene7868.tl Vitrella_brassicaformis_gene7868.tl Cryptococcus_neoformans_varneoformans_JEC21 Neosartorya_fischeri NRRL 181 Aspergillus_flavus_NRRL3357 Ajellomyces_capsulatus_NAm1 Uncinocarpus_ressii_1704 Neurospora_rassa_OR74A Pyrenophora_ratifici-repentis_Pt-1C-BFP Talaromyces_stipitatus_ATCC_10500 Paracoccioides_brasilieneis_Pb01 Vertocioides_orarum_VAMS_102	DUCE TRANSCIPULSEN WILE TLASS APODAL MILE TLASS APODAL MILE TLATTS VIDUS 61 EG - G - G - G - G - G - G - G - G - G -	DALU POCTOR DALG	VOIVAPEESETEA MOLVAPEESETEA POLVAPEELIVA POLVAPEGEITYO COCCELTIVA	II DEAM - GACYEARI HYDRAA - RECEARC LI DEAE - AACYEARU HYDRAE - AACYEARU AISEPID - BLOWN - D AISEPID - BLOWN - D AISEPID - BLOWN - D AISEPID - BLOWN - D AISEPID - BLOWN - D	VIDIO CONCERENT VIDA POLOSE ELM VIDA POLOSE EL	NALITETES STORD NARTH PAPLALEN RUPEL PAPLALEN PIGLEGVQTAEDAME PVCLEGVQTAEDAME PVCLEGVQTAEDAME PVLEGVQTYEDAIL PIVLEGVQTYEDAIL PIVLEGVQTYEDAIL PILLEGVQCVEDVLE PILLEGVQCVEDVLE PILLEGVQCVEDVLE PILLEGVQCVEDVLE PIVLEGVQCVEDVLE PIVLEGVQCVEDVLE PIVLEGVQCVEDVLE PIVLEGVQCVEDVLE
<pre>serials a maring file Jobdy Karlodinium micrum_1 Karlodinium micrum_2 Chromes, velia gene425.11 Chromes, velia gene425.11 Vitrella_brassicaformis_gene19079.11 Vitrella_brassicaformis_gene19079.11 Accession of the series o</pre>	DLOFILASCIPUSH HAVANAASHEBEBT YMLFILSSVAPDDML MILGILAITSVTDVS 61 BGG AP	DALU - POOPEN DALQ - POOPEN -ESS - GEGUE ANCC - SEGUE - CANDOR - BESSNDR - BESSNDR - BESSNDR - BESSNDR - BESSNDR - BESSNDR - SGRVDR	VOIVVA PEESKIEA MOLVVA PEESKIEA FOLVIPRORELIVQ FOLVIPRORELIVQ GG GIG GG GG GG GG GG GG GG GG GG GG GG GG GG	II REAM - ORGYNALL HYTRAA - REGCALC LINERE - ARGYSAIV LYTERE - ARGYSAIV LYTERE - ARGYSAIV WIAGFIDEGLTWD-D POWGTDASLTWEET AISSTDPELSWE-D AISSTDPELSWE-D AISSTDPELSWE-D AISSTDPELSWE-D AISSTDPELSWE-D 	VIUL QUEEKEE BUN - YDL PLGEE BUN - YDL PLGEE BUN IVUD PLGEE BADL IVUD	NELTHYER - FICE NEPTLPRIATEN NEPTLPRIATEN PICKEVOTAEDANE PICKEVOTAEDANE PICKEVOYAEDANI PICKEVOYAEDANI PICKEVOYAEDANI PICKEVOYAEDANI PICKEVOYAEDANI PICKEVOYAEDANI PICKEVOYAEDANI PICKEVOYAEDANI PICKEVOYAEDANI PICKEVOYAEDANI
<pre>serviness marines rice_objoint Karlodinium micrum_1 Karlodinium micrum_2 Chromera_velia_gene4425.t1 Chromera_velia_gene4425.t1 Vitrella_brassicaformis_gene7868.t1 Vitrella_brassicaformis_gene7868.t1 Vitrella_brassicaformis_gene7868.t1 Neosartorya_fischeri NREL 181 Aspergillus_flavus NREL3357 Ajellomyces_capsulātus NAn1 Uncinocarpus_ressii 1704 Neurospora_tritici-repentis Pt-1C-BFP Talaromyces stipitatus ATCC 10500 Paracoccioides brailiensis Pb01 Verticillium albo-atrum YaMs.102 Arthroderma_otac CBS 113480 Saccharomyces_carevisiae_S288c Lachancea thermotolerans</pre>	DECENTRASCIPUESAN WILSTUSSYAPDONL WILSTUSYAPDONL MILSTUSYAPDONL 61 SG 	DALU POCTA BALC POCTA BALC CONTACT BALC CONTACT BALC CONTACT C	VOILVVA PERSATIA MOLVVA PERSATIA POLVVA PERSATIA POLVVA PERSATIA COLVANA PERSATIA COLVANA COLV	IT DEAM - OAGVEANT IT DEAM - ANGVEANT IT SPIDPELAW - D IT SPIDPELAW -	VIDIOUSSEEUR WIDAPUGEEER -VDLPIGHEADI IIWIDAPUGEERDI IIVIDYPHEERDI IIVIDYPHEERDI IIVIDYPHEERDI IIVIDY IIVIDY VIDIOUSSEE VIDIOUSS	BRIGHTTER BOOM BRIGHTER BOOM BRIGHTER BOOM BRIGHTER BOOM PUCKSOWS ADDALE PUCKSOWS ADDALE PUCKSOWS ADDALE PUCKSOWS ADDALE PUCKSOWS ADDALE PILL SOUR VADULE PILL SOURVED ADDALE PILL SOURVED ADDALE
Karlodinium mirr_1 Karlodinium mirr_2 Chromera velia genef425.tl Chromera velia genef573.tl Vitrella brassicaformis genef968.tl Vitrella brassicaformis genef9079.tl Accession of the second se	CLOP LARCHELSET WILPILSEVAPDONL MILSINSTIPLEEVA GI GG GG DA 	DALU	VOIVAPEESETEA MOLVAPEESETEA FOLVEROCELUIVA GG GG GG GG GG GG GG GG GG GG GG GG GG	II REAM-ORGYEALL HYTRAA-REGGEALC HYTRAE-ARGYEALY EVGGEIDPELSWE-D WRAGFIDPELFWE-D POWSGDASLIWEET AISSTIDPELSWE-D AISSTIDPELSWE-D AISSTIDPELSWE-D AISSTIDPELSWE-D AISSTIDPELSWE-D SALTWD-Y AISSTIDPELSWE-D SALTWD-Y AISSTIDPELSWE-D SALTWD-Y AISSTIDPELSWE-D SALTWD-Y AISSTIDPELSWE-D SALTWD-Y AISSTIDPELSWE-D SALTWD-Y AISSTIDPELSWE-D CHEAGUDPELIWEDT	VIDDEDGESERUM -VTDAPGGESERUM -VTDAPFGESERUM ITVDYPAGESERUM ILVVDTPTLGERNADI IEWLE	NRITHTER-BIOD NRFHLERFLEFT NRFHLERFLEFT NRFSLFTCLSLAHF PICLSCWERADDATL PICLSCWERADDATL PICLSCWERADDATL PICLSCWERADDATL PICLSCWERADDATL PILLSCVERADDATL PILLSCVERADDATL PILLSCVERADDATL PILLSCVERADDATL PILLSCVERADDATL PILLSCVERADDATL PILLSCVERADDATL PILLSCVERADDATL PILLSCVERADDATL PILLSCVERADDATL PILLSCVERADDATL PILLSCVERADDATL
<pre>Selfision and files file of of</pre>	DLOF TARSCIPULS HAVEN TARS RELEAT YMD PTLSS VAPDONL MTLSTLATSVTDVS 61 96 	DALU	VOIVAPEESSTER MOLVVAPEESSTER MOLVVAPEESSTER GG GG <td< td=""><td>LI DEAN - OAGVIALL LI DEAN - ARGCEALC LI DEAN - ARGCEALC LI DEAN - AAGVIALU WINAGED - AAGVIALU WINAGED - CLINNO - DOWNGUDALLINNET AISSFIDPELSWA - DOWNGUDALLINNET - TRUTNE - MISSFIDPELSWA - - SALINNO - NISSFIDPELSWA - SALINNO - NISSFIDPELSWA - NISSFIDPELSWA</td><td>VIDD 2002 E 200 -VDL 2002 E 200 -VDL 2002 E 200 IIV/DY PAGE E 200 IIV/DY PAGE E 200 IIV/DY PAGE E 200 IIV/DY PAGE A 200 IIV/DY - 041 IIV/DY - 041 IIV/DY - 041 IV/DY - 041 IV/DY</td><td>NHITHTER-BIOD NEFHIGADIEN NHIFTLPHLITAN PHILTAN PYCLKOVERDAN PYCLKOVERDAN PYCLKOVERDAN PYLKGVCYBDVL PILLGVCYBDVL PILLGVCYBDVL PILLGVCYBDVL PILLGVCYBDVL PILLGVCBDVL PYLKGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL</td></td<>	LI DEAN - OAGVIALL LI DEAN - ARGCEALC LI DEAN - ARGCEALC LI DEAN - AAGVIALU WINAGED - AAGVIALU WINAGED - CLINNO - DOWNGUDALLINNET AISSFIDPELSWA - DOWNGUDALLINNET - TRUTNE - MISSFIDPELSWA - - SALINNO - NISSFIDPELSWA - SALINNO - NISSFIDPELSWA	VIDD 2002 E 200 -VDL 2002 E 200 -VDL 2002 E 200 IIV/DY PAGE E 200 IIV/DY PAGE E 200 IIV/DY PAGE E 200 IIV/DY PAGE A 200 IIV/DY - 041 IIV/DY - 041 IIV/DY - 041 IV/DY	NHITHTER-BIOD NEFHIGADIEN NHIFTLPHLITAN PHILTAN PYCLKOVERDAN PYCLKOVERDAN PYCLKOVERDAN PYLKGVCYBDVL PILLGVCYBDVL PILLGVCYBDVL PILLGVCYBDVL PILLGVCYBDVL PILLGVCBDVL PYLKGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL PILLGVCBDVL
Karlodinium micrum 1 Karlodinium micrum 2 Chromera velia genef425.tl Chromera velia genef573.tl Vitrella brassicaformis genef968.tl Vitrella brassicaformis genef9079.tl Acception of the second secon	CLOP LARCH L	DALU - POOTO DALO - POOPEN ESS - GEGLEP ANG - GEGLEP - GE	VOIVAPEESETEA MOLVAPEESETEA POLVAPEELIVA GG GG GIG GG GG GIG GG GG GIG GG GG GG	II DEAN - OAGVEALL HVIKAA - REGEALC II DEAE - AAGVEALV LVIRAE - AAGVEALV LVIRAE - AAGVEALV UVRACTIDGITWD-D POWGTDAGITWD-D POWGTDAGITWD-D AISSTIDGILWW-D AISSTIDGILWW-D AISSTIDGILWW-D AISSTIDGILWW-D AISSTIDGILWW-D SALTWD-V AISSTIDGILWW-D SALTWD-V AISSTIDGILWW-D SALTWD-V AISSTIDGILWW-D GULWADDUMWET	VIUD POUGEERUM VDLPEGHERDI ITVDPAGEERDI ITVDPAGEERDI ITVDPTIGERNDI IEWLEOHIKL IVVVFHITKL IVVVFHITKL IVVFSITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0SITKM IPVF0	NRICHTER-BIOD NIFHER-BILEFF NIFFLPFLLLAHF NOFSLPTCLSLAHF PVCLSCWARADDATL PVCLSCWARADDATL PVCLSCWARADDATL PVCLSCWARADDATL PVCLSCWARADDATL PIVLSCVCVCNDVLR PIVLSCVCVCNDVLR PILLSCVCVCNDVLR PILLSCVCVCNDVLR PILLSCVCVCNDVLR PILLSCVCVCNDVLR PILLSCVCVCNDVLR PILLSCVCVCNDVLR PILLSCVCVCNDVLR PILLSCVCVCNDVLR PILLSCVCVCNDVLR PILLSCVCVCNDAL PILLSCVCVCNDAL PILLSCVCVCNDAL PILLSCVCVCNDAL
Arlodinium micrum_1 Karlodinium micrum_1 Karlodinium micrum_2 Chromera_velia_gene4425.tl Chromera_velia_gene4425.tl Chromera_velia_gene425.tl Vitrella_brassicaformis_gene1868.tl Vitrella_brassicaformis_gene19079.tl Sapergillus_flavus_NRKL3357 Ajellomyces_capsulatus_NAM1 Uncinocarpus_ressi1704 Neurospora_crassa_OR74A Pyrenophora_tritic1repentis_Pt-1C-BFP Talaromyces_stipitatus_ARCC_10500 Paracoccidicides_brasiliensis_Pb01 Verticillium_albo-atrum_VAMs_102 Arthroderma_otae_CBS_113480 Saccharomyces_crevisiae_S289c Lachancea_thermcolerans Trichophyton_rubrum_CBS_118992 Candida_dubliniensis_CD36 Schizophyllum_commune_H-8 Komagataella_pastoris_GS115_	DLOF HASCHDENN WILGTLSSVAPDDUL MILGTLSSVAPDDUL 61 86	DALU	YOI VAPERSHTA MOLYVAPERSHTA FOLVYFORGITYD FOLVYFFORGITYD GG GIGE GG GIGE GG GIGE GG GIGE GG GAA SQ GAA SP G	LIERA-ORGVIALI HVERA-ORGVIALI LIERA-ARGVIAL LIERA-ARGVIALU UVERA-ARGVIALU POVGCPALINU- TALSTIDPLLW-D LISTIDPLLW-D TALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-A ALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-D ALSTIDPLLW-D	MINDA CORE EDIN - NDLPIGE EDIN - NDLPIGE EDIN INVOVRACE EDI	IN LITTER STORE STORE



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 CsCI-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 MiSeg 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,	average coverage, rea @ enBank
	bp	of the mitochond aac ession #
		DNA-enriched
		fraction
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_1	2043	1934
0		
114504 cox1 3 c1 1	2113	1909
1		
114504_cox1_3_c1_1	2127	1204
2		
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-	920	1892	
coding c2 2			
114504_non-	1343	2031	
coding c2 3			
114504 non-	1413	1871	
coding c2 4			
114504 non-	1427	1911	
coding c2 5			
114504 non-	917	1892	
coding c2 7	517	1032	
<u>114504 pop</u>	1240	2022	
114504_11011-	1340	2032	
	1410	1071	
114504_NON-	1410	1871	
coding_c2_9	4.40.4	1010	
114504_non-	1424	1912	
_coding_c2_10			
114504_non-	618	2035	
_coding_c3_1			
114504_non-	632	2122	
_coding_c3_2			
114504_non-	721	2013	
coding_c4_1			
114504_non-	791	1728	
coding c4 2			
114504_non-	805	1802	
codina c4 3			
114504 non-	677	1971	
coding c5 2			
114504 non-	747	1673	
TT4204_11011-	141	1012	

coding c5 3			
<u>114504 non-</u>	761	1753	
coding cE 4	701	1100	
<u>4</u> 114504_pop_	868	2207	
114504_11011-	000	2291	
<u></u>	6670	107	
<u>13057_mil2_1</u>	5579	187	
<u>13113_rns7_1</u>	7644	128	
13113_rhs/_2	4519	80	
154998_non-coding_1	604	1460	
<u>154998_non-coding_3</u>	1042	2803	
<u>154998_non-coding_4</u>	925	2755	
<u>154998_non-coding_5</u>	481	1951	
154998_non-coding_7	919	2668	
154998_non-coding_8	802	2593	
154998_non-coding_9	375	2152	
154998_non-	328	4291	
_coding_10			
_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	3879	16	
1			
<u></u> 46689_non-	1189	2475	
coding c13 1	1100	2410	
<u></u>	1179	1566	
40919_{0000}	1066	2020	
40919_{0000}	1900	2057	
40313_0003_0001_3	1065	30/1	
40313_00X3_00X1_4	7302 T202	2050	
40919_00X2_00X1_5	2230	2000	
40919_00X3_00X1_0	2001	2020	
40919_00X3_00X1_/	2100	3029 2001	
_40919_COX3_COX1_8	2104	1805	

46010 0012 0011 0	2225	2120	
46919_C0X3_C0X1_9	2323	3139	
67435_non-coding_1	3670	810	
67435_non-coding_2	835	944	
6/435_non-coding_3	6338	185	
<u>67435_non-coding_4</u>	666	1248	
<u>67435_rns7_1</u>	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	
<u>68201_non-coding_13</u>	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	
v			

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 noncoding 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	transcri pt name	size , bp	strands transcr ibed	feature name	feature coordinates within transcript	full OR F leng th	start codo n
11012_rnl8_1	transcri	124					
	pt_1						
	transcri	86	both	cox3 fragment,	complement(174		
	pt_2			449-476 bp	5)		
				ORF1	complement(<1>		
					86)		
	transcri	144	both	ORF2	<159	306	CAT
	pt_3						
	transcri	139	both				
	pt_4						

114504_cox1_3_	transcri	148	both	ORF2	52>148	315	CTA
c1_1	_pt_1						
	transcri	214	both				
	_pt_2						
	transcri	121	both				
	_pt_3						
	transcri	168	both				
	pt_4						
114504_cox1_3_	transcri	132	both	putative rRNA	1202		
2	pt_2						
	transcri	487					
	pt_3						
	transcri	700		ORF1	5>700	723	ATG
	pt_4						
114504_cox1_3_	transcri	881	both	ORF2	<1153	315	CTA
c1_4	pt_2			ORF3	complement(496	387	TAT
					>881)		
				Plasmodium rnl8 hit	complement(104		
					196)		
				<i>rnl</i> 8 transcript	complement(122		
					8)		
114504_cox1_3_	transcri	744	both	ORF2	<150	315	CTA
c1_5	pt_2			ORF3	complement(393	675	ATA
					>744)		
				Plasmodium rnl8 hit	complement(193		
)		
				<i>rnl</i> 8 transcript	complement(<11		
					25)		
114504_cox1_3_	transcri	135	both	ORF2	<1153	315	СТА
c1_7	pt_2	7		ORF3	complement(496	462	ACC
					957)		
				Plasmodium rnl8 hit	complement(104		
					196)		
				rnl8 transcript	complement(122		
					8)		
114504_cox1_3_	transcri	435	both	ORF3	138>435	675	ATA
_c1_10	pt_3						

114504_cox1_c3	transcri	696		ORF2	1>696	723	ATG
6 – –	pt 4						
	transcri	694	both	ORF1	complement(309	387	TAT
coding_c2_2	pt_1				>694)		
				Plasmodium rnl8 hit	complement(<19		
)		
114504_non-	transcri	117	both	ORF1	complement(309	462	ACC
coding_c2_5	pt_1	0			770)		
				Plasmodium rnl8 hit	complement(<19		
)		
114504_non-	transcri	594	both				
_coding_c3_2	pt_1						
114504_non-	transcri	711	both	ORF1	138545	408	ATA
_coding_c4_1	pt_1						
114504_non-	transcri	761	both				
_coding_c4_3	pt_1						
114504_non-	transcri	714	both				
_coding_c5_4	pt_1						
114504_non-	transcri	856	both	ORF1	complement(213	309	AGG
_coding_c9_9					521)		
13057_rnl12_1	transcri	758		ORF1	complement(462	366	CTT
	pt_1				>758)		
				ORF2	466>758	387	GGG
				ORF3	612>758	537	TGT
				ORF4	complement(703	480	CIC
		400	la a tla	0052	>758)	F07	тот
	transcri	420	DOTH		<1205	537	
	pt_2			URF4		480	CIC
					99)	200	TCC
	transcri	222		URFJ	13294	309	166
	nt 2	222					
	_µo transcri	03	both		complement(<1 >	108	ΛΛΤ
	nt 1	30	JUIT		03)	400	
	μ 4			ORE7	<u> </u>	2/10	GT^
				ORF8	 comnlement(0 >0	<u></u> <u></u> 1	
					2)	-+ 1	
	transcri	92		ORE8	 78 >92	<u>1</u> /1	ΤΤΔ
	uunsun	ĴΖ			10.1202	-++T	

	_pt_5						
	transcri	116					
	pt_6						
13113_rns7_1	transcri	254		ORF2	<1>254	876	TTT
	_pt_1						
	transcri	175					
	pt_2			ORF3	complement(244	117	AAA
					1422)	9	
				ORF3a	230679	450	CGT
				ORF4	complement(630	639	СТА
					1268)		
	transcri	146					
	_pt_4						
	transcri	227					
	pt_5						
	transcri	102					
	pt_6						
13113_rns7_2	transcri	123					
	_pt_1						
	transcri	803		ORF1	complement(90>	744	CCA
	pt_2				803)		
				ORF2	46801	756	GAC
	transcri	217					
	pt_3						
154998_non-	transcri	280	both	ORF1	complement(<11	318	ATG
_coding_1	pt_2				04)		
154998_non-	transcri	751	both				
_coding_3	pt_1						
154998_non-	transcri	157	both	ORF1	complement(<11	318	ATG
_coding_5	pt_2				04)		
154998_non-	transcri	362	both				
coding_9	pt_1						
154998_non-	transcri	328	both				
_coding_10	pt_1						
19169_cox1frag_	transcri	273		ORF1	complement(<11	106	ATG
1	pt_1				06)	2	
	transcri	184		ORF1	complement(<1>	106	ATG

	pt_2				184)	2	
				ORF2	<1>184	843	ATG
	transcri	206		ORF3	complement(<1>	156	ATG
	pt_3				206)	6	
				ORF4	<197	122	ATG
						7	
	transcri	143		ORF3	complement(<1>	156	ATG
	pt_4				143)	6	
				ORF4	<1>143	122	ATG
						7	
	transcri	119		ORF10	<162	153	ATG
	pt_5					9	
				ORF6	<1>119	134	ATG
						4	
				ORF8	complement(<1>	909	ATG
					119)		
				ORF9	complement(<1>	693	ATG
					119)		
	transcri	298		ORF10	<1>298	153	ATG
	_pt_6					9	
	transcri	107	both				
	pt_9						
203480_non-	transcri	393	both				
coding_3	pt_1						
274408_non-	transcri	319	both				
coding_4	pt_1						
274408_non-	transcri	584	both				
coding_6	pt_1						
33769_cox1frag_	transcri	118	both				
1	_pt_1						
	transcri	167	both				
	_pt_2						
	transcri	152	both				
	_pt_4						
	transcri	229	both				
	_pt_5						
	transcri	135	both	ORF2	complement(<17	315	ATG

	pt_6				6)		
				Plasmodium rns7	complement(134		
				hit	>135)		
	transcri	133	both	cox1 fragment, 1-	complement(149		
	pt_7			79 bp	2)		
				ORF1	complement(9>1		
					33)		
	transcri	127	both	ORF3	complement(123	348	ATG
	pt_8				>127)		
33769_non-	transcri	150		ORF1	2941292	999	ATG
coding_5	pt_1	5		ORF2	complement(282	993	ATG
					1274)		
				ORF3	complement(290	411	ATG
					700)		
				ORF4	complement(825	513	ATG
					94)		
33769_non-	transcri	146		ORF1	2531251	999	ATG
coding_6	pt_1	4		ORF2	complement(241	993	ATG
					1233)		
				ORF3	complement(249	411	ATG
					659)		
				ORF4	complement(145	540	ATG
					53)		
33769_non-	transcri	247		ORF1	12602258	999	ATG
coding_7	pt_1	1		ORF2	complement(1248	993	ATG
					2240)		
				ORF3	complement(1256	411	ATG
					1666)		
				ORF4	complement(1021	540	ATG
					1560)		
				ORF5	7481251	504	ATG
				ORF6	300944	645	ATG
				ORF7	complement(537	354	ATG
					890)		
				ORF8	complement(137	312	ATG
					448)		
				ORF9	20367	348	ATG
33769_non-	transcri	256		ORF1	13562354	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	8441347	504	ATG
				ORF6	3961040	645	ATG
				ORF7	complement(633 986)	354	ATG
				ORF8	complement(233 544)	312	ATG
				ORF9	116463	348	ATG
34169_cox1frag_	transcri	990		ORF1	<1276	375	ATG
1	pt_1			ORF2	complement(<12 87)	378	ATG
				ORF3	complement(172 657)	486	ATG
				ORF4	269688	420	ATG
				ORF5	286708	423	ATG
				ORF6	complement(716 >990)	315	ATG
36603_cox1_3_1	transcri	572	both	cox1 fragment,	<1560		
	pt 1			1193-1766 bp			
				ORF1	49>572		
	transcri	132	both				
	pt_2						
	transcri	355					
	pt_3						
	transcri	137		ORF1	complement(473	408	ATG
	pt_4	6			880)		
				ORF2	complement(436	321	ATG
					756)		
				ORF3	39416	378	ATG
				ORF4	20379	360	ATG
36603_non-	transcri	166		ORF1	109486	378	ATG
coding_1	pt_1	4		ORF2	complement(227 595)	369	ATG
					/		

			ORF3	complement(622	414	ATG
			ORF4	>1664)	306	ATG
	transcri pt 2	172 both	ORF5	<1>172	330	ATG
	transcri pt_3	417	ORF5	complement(304 >417)	330	ATG
			ORF6	<113	318	ATG
	transcri pt 4	294	ORF6	complement(<15 1)	318	ATG
			ORF7	complement(<12 29)	423	ATG
			ORF8	177>294	417	ATG
	transcri	171	ORF10	<1>171	441	ATG
	pt_5		ORF9	complement(<1> 171)	381	ATG
	transcri pt_6	169	ORF11	complement(<1> 169)	495	ATG
36603 non-	transcri	950 both	ORF1	661>950	378	ATG
coding_2	pt_1		ORF2	complement(779 >950)	369	ATG
	transcri	115	ORF2	complement(<19	369	ATG
	pt_2	8		6)		
	. –		ORF3	complement(864 06)	321	ATG
			ORF4	201617	417	ATG
			ORF5	complement(858 >1158)	306	ATG
36603_non- coding 4	transcri pt 2	341				
	transcri	114	ORF5	complement(<14	537	ATG
	pt 3	3		70)		
	· _		ORF6	163888	726	ATG
			ORF7	complement(265 870)	606	ATG
			ORF8	complement(635	372	ATG

					1006)		
36603_non-	transcri	156		ORF1	152538	387	ATG
coding 5	pt 1	3		ORF2	complement(535	408	ATG
3_	. –				942)		
				ORF3	complement(982	309	ATG
					1290)		
36603_non-	transcri	109		ORF1	complement(<14	486	ATG
coding 6	pt 1	7			24)		
				ORF2	117842	726	ATG
				ORF3	complement(219	606	ATG
					824)		
				ORF4	complement(589	372	ATG
					960)		
3725_rRNA_1	transcri	143					
	pt 1						
	transcri	367		ORF1	<1119	615	ATG
	pt 2						
	transcri	159		ORF1	complement(<1>	615	ATG
	pt 3				159)		
	transcri	832		ORF3	1528	612	ATG
	pt 4			ORF4	<1161	606	ATG
	transcri	854		ORF4	complement(<14	606	ATG
	pt 5				26)		
				ORF4a	<1633		
	transcri	172		ORF5	complement(<1>	678	ATG
	pt_6				172)		
				ORF6	<1>172	624	ATG
				ORF7	<1>172	606	ATG
	transcri	98					
	pt_7						
37286_rns7_1	transcri	108					
	pt_1						
	transcri	133					
	pt_2						
41777_rnl12_rns7	transcri	114	both				
_1	pt_1						
43250_rns7_1	transcri	550		ORF1	<1>550	642	GTG
	pt_1						

	transcri	99	both				
	pt 2						
	transcri	133	both				
	pt_3						
	transcri	118	both				
	pt_4						
46689_cox3frag_	transcri	293		ORF1	<11073	125	ATG
c12_1	pt_1	2				4	
				ORF2	complement(139	702	ATG
					840)		
				ORF3	complement(194	852	ATG
					1045)		
				ORF4	218922	705	ATG
				ORF5	complement(861	885	ATG
					1745)		
				ORF6	9011686	786	ATG
				ORF7	9321579	648	ATG
				ORF8	complement(1310	960	ATG
					2269)		
				ORF9	21252745	621	ATG
46689_non-	transcri	871		ORF1	complement(527	519	ATG
coding_c13_1	pt_1				>871)		
				ORF2	complement(621	363	ATG
					>871)		
				ORF3	complement(280	459	ATG
					738)		
	transcri	130	both				
	pt_2						
46919_cox3_cox	transcri	134	both	ORF2	32>134	462	TGT
_1_1	pt_2						
46919_cox3_cox	transcri	232		cox1 fragment, 1-	complement(1452		
1_9	pt_1	5		579 bp	2030)		
				cox3	complement(582		
					1451)		
				ORF1 cox1-cox3	complement(582	170	TCA
					2282)	1	
				ORF2	19782286	309	TCA
				Plasmodium rnl1 hit	170247		

				rnl1 transcript	137263		
67435_non-	transcri	145	both				
coding_1	pt_1						
	transcri	176	both				
	pt_2						
	transcri	127	both	ORF1	<1>127	327	CAA
	pt_3						
	transcri	143	both	ORF1	<182	327	CAA
	pt_4			ORF2	complement(69>	318	TCG
					143)		
	transcri	109	both	ORF4	complement(<1>	360	GGA
	pt_5				109)		
	transcri	196	both	ORF4	<161	360	GGA
	pt_6						
	transcri	361	both				
	pt_7						
	transcri	134	both				
	pt_8						
67435_non-	transcri	120	both				
coding_2	pt_1						
	transcri	133	both	ORF1	complement(<1>	555	GAA
	pt_2				133)		
	transcri	116	both	ORF1	complement(<1>	555	GAA
	pt_3				116)		
67435_non-	transcri	158	both				
coding_3	pt_2						
	transcri	311	both				
	pt_3						
	transcri	148		ORF1	<1546	747	GAA
	pt_4	5		ORF2	complement(<19	102	AAC
					97)	9	
				ORF3	1021001	900	TTG
				ORF4	149853	705	CAG
				ORF5	complement(279	128	AAT
					>1485)	1	
	transcri	532		ORF6	complement(<12	723	AAG
	pt_5				32)		
				ORF7	complement(<14	750	AAG

					55)		
				ORF8	<1>532	155	TTA
						7	
				ORF9	complement(519	819	CCA
					>532)		
	transcri	102		ORF10	<1>102	693	TCC
	pt_6			ORF11	complement(<1>	618	CCC
					102)		
				ORF8	<1>102	155	TTA
						7	
				ORF9	complement(<1>	819	CCA
					102)		
67435_non-	transcri	236	both	ORF1	<1159	519	GAA
_coding_4	pt_1						
68201_non-	transcri	776	both	ORF1	154588	435	CTT
coding_2	pt_1			Plasmodium rns5	complement(658		
				hit	675)		
68201_non-	transcri	607	both	ORF1	<1419	465	NNN
coding_3	pt_1			Plasmodium rns5	complement(489		
				hit	506)		
68201_non-	transcri	376	both				
_coding_7	pt_1						
68201_non-	transcri	700	both				
_coding_12	pt_1						
68201_non-	transcri	606	both				
_coding_13	pt_1						
73351_non-	transcri	160	both				
_coding_8	pt_4						
73351_non-	transcri	156	both	ORF2	complement(<1>	447	ATC
_coding_10	pt_1				156)		
73351_non-	transcri	150	both				
_coding_12	pt_1						
73351_non-	transcri	246	both	ORF1	complement(70>	384	GAA
coding_13	_pt_1				246)		
	transcri	105	both	ORF1	complement(<1>	384	GAA
	pt_2				105)		
_73351_rRNA_3	transcri	172					

	pt_1						
	transcri	122	both				
	pt_2						
73351_rRNA_4	transcri	153	both	ORF1	complement(<1>	414	AGA
	pt_3				153)		
	transcri	122	both				
	pt 4						
95117_non-	transcri	737	both	ORF2	complement(148	333	TAT
coding 2	pt 1				480)		
95117_non-	transcri	601	both				
coding_6	pt_1						
95117_non-	transcri	117	both	ORF1	6431059	417	TCC
coding_8	pt_1	4		ORF2	complement(207	306	ATA
					512)		
				Plasmodium rns5	complement(1056		
				hit	1073)		
95117_non-	transcri	639	both				
_coding_10	pt_1						
95117_non-	transcri	718	both	ORF1	186512	327	ACT
_coding_11	pt_1						
95117_non-	transcri	102	both	ORF1	490816	327	ACT
_coding_12	pt_1	2					
95117_non-	transcri	103	both	ORF1	complement(<11	390	AAG
coding_18	pt_1				03)		
				ORF2	complement(<17	375	TGT
					2)		
	transcri	577	both	ORF1	<1234	390	AAG
	pt_2			ORF2	<1250	375	TGT
				ORF3	complement(<13	327	TAT
					20)		
				ORF4	complement(119	321	CCC
					439)		
				putative rRNA	5123		
95117_non-	transcri	751	both	ORF1	2424	423	TAC
_coding_19	pt_1						
				ORF2	complement(293	321	CCC
					613)		
95117_non-	transcri	132	both				

coding_30	pt_3						
95117_non-	transcri	115	both	ORF1	<1>115	324	GTT
coding_35	pt_1			ORF2	complement(<11	432	ATG
					3)		
95117_non-	transcri	104	both	ORF3	complement(71>	333	GTT
coding_38	pt_3				104)		
	transcri	304	both	ORF3	complement(<11	333	GTT
	pt_4				71)		
	transcri	138	both	ORF4	36>138	321	GGC
	pt_5						
95117_non-	transcri	133	both				
coding_42	pt_5						
rRNA transcripts,	sorted by	trans	cript nam	е			
46919_cox3_cox	rnl1	127		cox1-cox3	complement(<1>		
1_1	transcri			transcript	127)		
	pt			transcript 1	<1>127		
				Plasmodium rnl1 hit	34111		
73351_rRNA_3	rnl1	196		Plasmodium rnl1 hit	80158		
	transcri						
	pt						
11012_rnl8_1	rnl8	214		Plasmodium rnl8 hit	4068		
	transcri						
	pt						
114504_cox1_3_	rnl8	228		ORF2	complement(76>	315	СТА
c1_1	transcri				228)		
	pt			Plasmodium rnl8 hit	33125		
114504_cox1_3_	rnl8	228		transcript 2	complement(<12		
c1 10	transcri				28)		
-	pt			ORF2	complement(76>	315	CTA
	•				228)		
				Plasmodium rnl8 hit	33125		
114504_cox1_c3	rnl8	231		transcript 2	complement(<12		
_16	transcri				31)		
—	pt			ORF2	complement(79>	315	СТА
	-				231)		
				Plasmodium rnl8 hit	36128		
114504_non-	rnl8	179		Plasmodium rnl8 hit	<184		

coding_c2_10	transcri pt						
33769_cox1frag_ 1	<i>rnl</i> 8 transcri pt	171		Plasmodium rnl8 hit	33125		
3725_rRNA_1	<i>rnl</i> 8 transcri pt	206		Plasmodium rnl8 hit	27100		
75915_rnl8_1	<i>rnl</i> 8 transcri pt	444	both	Plasmodium rnl8 hit	352444		
13057_rnl12_1	rnl12 transcri pt	137	both	<i>Plasmodium rnl</i> 12 hit	complement(751 06)		
41777_rnl12_rns7 _1	rnl12 transcri pt	164	both	<i>Plasmodium rnl</i> 12 hit	4475		
73351_non- coding_10	<i>rnl</i> 12 transcri	175	both	ORF1	complement(<16 5)	339	TGG
	pt			ORF2 Plasmodium rnl12 hit	142>175 3970	447	ATC
73351_rRNA_3	<i>rnl</i> 12 transcri pt	189	both	Plasmodium rnl12 hit	106137		
95117_non- coding_35	<i>rnl</i> 12 transcri	190	both	ORF2	complement(141 >190)	432	ATG
	pt			Plasmodium rnl12 hit	5384		
46919_cox3_cox 1_1	<i>rn</i> s1 transcri	261	both	ORF4	complement(<12 3)	318	TAC
	pt			rns1 hit	56108		
37286_rns7_1	rns7	96		ORF1	<187	672	GTG
	transcri			Plasmodium rns7	2160		
	pt			hit			
41777_rnl12_rns7	rns7	101		ORF1	complement(<1>	363	TAC
_1	transcri				101)		

	pt			<i>Plasmodium rn</i> s7 hit	2362		
43250 rns7 1	rns7	109	both	Plasmodium rns7	2766		
	transcri			hit			
	pt						
67435_rns7_1	rns7	110	both	Plasmodium rns7	2867		
	transcri			hit			
	pt						
73351_non-	rns7	103	both	Plasmodium rns7	2261		
coding_8	transcri			hit			
	pt						
73351_non-	rns7	106	both	ORF1	complement(<1>	339	TGG
coding_10	transcri				106)		
	pt			Plasmodium rns7	2463		
	-			hit			
73351_non-	rns7	105	both	Plasmodium rns7	2362		
coding_12	transcri			hit			
	pt						
95117_non-	rns7	103	both	Plasmodium rns7	2261		
coding_48	transcri			hit			
	pt						
114504_cox1_c3	putative	105		cox1 transcript	complement(<11		
_10	rRNA				05)		
				ORF1	complement(<12	185	TTA
					7)	7	
19169_cox1frag_	putative	141					
1	rRNA						
46919_cox3_cox	putative	122		cox1-cox3	complement(<11		
1_3	rRNA			transcript	22)		
40010		104		transcript 1	<1122		
46919_COX3_COX	putative	104		COX1-COX3	complement(<11		
1_8	rrna						
					<pre></pre>	150	TTC
				UKFI	1)	о 723	пс
05117 non	nutativo	110		transcript 2	 ∠1 \\110	U	
90TT/_II0II-	pulative	119			<u>~1/119</u> <1 >110	200	
					ヽエ< エエジ	030	770

ooding 19				<1 \110	275	тот
County_18	oding_18 IRNA		URF2	<1>119	3/5	IGI
			ORF4	complement(115		CCC
				>119)		
			ORF3	complement(<1> 32		TAT
				119)		
114504_cox1_3_	putative	202	transcript 3	1>202		
_c1_2	rRNA1					
114504_cox1_3_	putative	261				
_c1_2	rRNA2					

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	Seque ncing platfor m	Expe cted aver age inser t size	Aver age read leng th	Rea d nu mb er, mill ion	Used for mitocho ndrial genome assembl y	Estim ated insert size range	Aver age read leng th after filter ing	Read num ber in millio n after filteri ng
Genomi									
С									
libraries									
C. velia	Illumina	Illumin	350	101	173				
	PCR-	а							
	free	HiSeq							
C. velia	Illumina	Illumin	350	101	162				
	PCR-	а							
	free	HiSeq							
C. velia	Illumina	Illumin	350	101	219				
	PCR-	а							
	free	HiSeq							
C. velia	Illumina	Illumin	8,000	101	152				
	mate	а							
	pair	HiSeq							
C. velia	Illumina	Illumin	3,000	101	133				

	mate	a							
<u>C</u> volia	pair	HISeq	2 000	101	17				
C. Vella	mata		3,000	101	47				
	male-	a							
Curalia	pair	HISeq	2 000	101	100		1700	0	<u>г</u> о
C. Vella	numina	niumin	3,000	101	102	yes	1700-	95	53
	Nextera	a					5500		
	mate	HISeq							
<u> </u>	pair		0.000	050			220	100	0.5
C. Vella	liiumina	liiumin	8,000	250	8	yes	230-	190	3.5
	mate	a					800		
	pair	MiSeq	= 000	050				400	
C. Vella	iliumina	liiumin	5,000	250	9	yes	230-	183	3.7
	mate	a					800		
	pair	MiSeq		050				400	
C. velia	Illumina	Illumin	3,000	250	12	yes	230-	182	4.6
	mate	a					800		
	pair	MiSeq	N1/A	00.4	0.00		N1/A		
C. Vella ^	454	454	N/A	224	0.29	yes	N/A		
	single	GS							
C	reads		N1/A	005			N1/A		
C. Vella ^	454	454	N/A	225	0.3	yes	N/A		
	single	GS							
	reads		N1/A	00.4	0.01		N1/A		
C. Vella	454	454	N/A	294	0.61	yes	N/A		
	single	GS							
0	reads	FLX	N1/A	000	0.00		N1/A		
C. Vella	454	454	N/A	326	0.63	yes	N/A		
	single	GS							
	reads	FLX	N1/A		0.00		N1/A		
C. Vella	454	454	N/A	292	0.63	yes	N/A		
	single	GS							
	reads		N1/A	004	0.70		N1/A		
C. vella	454 · ·	454	N/A	304	0.73	yes	N/A		
	single	GS							
0	reads	FLX			4.0				
C. velia *	Illumina	Illumin	330	54	19	yes	240-	54	13

	TruSeq, mitocho ndrial DNA- enriched fraction	a GAII				820
V. brassicaf ormis	Illumina TruSeq + Kappa Hifi	Illumin a HiSeq	350	101	288	
V. brassicaf ormis	Illumina TruSeq + Kappa Hifi	Illumin a HiSeq	350	101	203	
V. brassicaf ormis Transcri ptomic libraries	Illumina PCR- free	Illumin a HiSeq	500	93	155	
C. velia	Illumina TruSeq strand- specific	Illumin a HiSeq	300	101	93	
C. velia	Illumina TruSeq strand- specific	Illumin a HiSeq	300	101	118	
V. brassicaf ormis	Illumina TruSeq strand- specific	Illumin a 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	3,188,77	N/A
	4	
# Illumina mate pair reads	26,674,2	1700-5500
	72	
# Illumina paired-end MiSeq reads	11,789,8	230-800
	18	
# Illumina paired-end reads of the	13,136,9	240-820
mitochondrial DNA-enriched fraction	66	
# total reads	54,789,8	
	30	
# reads assembled	45,731,2	
	47	
total assembly size, Mbp	256	
# scaffolds	34,422	
scaffold length threshold, bp	1,000	
largest scaffold, bp	5,019,58	
	7	
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.