

Evolution of the Cytosolic Iron/Sulfur cluster Assembly machinery in *Blastocystis* sp. and other microbial eukaryotes

Anastasios D. Tsaousis, Eleni Gentekaki, Laura Eme, Daniel Gaston & Andrew J. Roger

Supplementary materials list:

1. Supplementary figure legends
2. Supplementary References
3. Figure S1
4. Figure S2
5. Figure S3
6. Figure S4
7. Figure S5
8. Figure S6
9. Figure S7
10. Figure S8
11. Figure S9
12. Figure S10
13. Figure S11
14. Figure S12
15. Figure S13
16. Figure S14
17. Figure S15
18. Table S1

Supplementary figure legends:

Figure S1:

Schematic model of the CIA pathway in opisthokonts. Dre2 along with the diflavin reductase Tah18, form an electron transfer chain with the use of NADPH for Tah18-dependent reduction of one of the two Fe/S clusters of Dre2. The nuclear binding proteins Nbp35 and Cfd1 form a scaffold complex responsible for assembling of transient bound Fe/S cluster in an early step of the biosynthetic process. In a later stage the Fe/S clusters are transferred to apo-proteins via Cia1 and Nar1 proteins. Lastly, MMS19 and Cia2 form a complex with Cia1 and Nar1 that facilitates the transfer of mature clusters to apo-proteins.

Figure S2:

Identification of functionally important structures in eukaryotic Cia1 sequences. Sequence alignment of Cia1 protein sequences from microbial eukaryotes against the published structure of the yeast Cia1p (3). Cia1 are WD40 proteins that contain 28 β -sheet structures highlighted in **green**. Potential helices are highlighted in **pink**. We used the PSIPRED server (1) for the structure prediction of the different proteins.

Figure S3:

Conservation of functionally important residues in Nar1 sequences from microbial eukaryotes. Sequence alignment of Nar1 protein sequences against orthologous proteins from microbial eukaryotes have demonstrated the presence of a ferredoxin-like domain

(highlighted in blue) along with conserved residues forming an H-cluster domain (highlighted in yellow). Cysteines (potentially involvement in Fe/S cluster binding) are marked in red.

Figure S4:

Conservation of functionally important residues in Nbp35 and Cfd1 sequences from microbial eukaryotes. Sequence alignment of Nbp35 and Cfd1 protein sequences against bacterial and eukaryotic orthologous proteins. The conserved N-terminal CX13CX2CX5C motif (2) found only in the Nbp35 eukaryotic homologues is highlighted in blue. The Mrp family signature is indicated in green, the conserved Walker A motif is highlighted in grey, whereas the C-terminal CPXC is highlighted in yellow. Cysteines (potentially involvement in Fe/S cluster binding) are marked in red. Cfd1 protein sequences are shaded in grey for comparison with the Nbp35 homologues.

Figure S5:

Western blot analyses of the expression of *Blastocystis* IscS, Nar1, Nbp35, SufCB and Tah18. Western blots demonstrating the cellular localization of the *Blastocystis* components of Fe/S cluster biosynthesis. The heterologous anti-yeast Nbp35, Nar1 and Tah18 antisera shows specific detection of *Blastocystis* Nbp35, Nar1 and Tah18 proteins with apparent relative molecular mass of 36 kDa, 83 kDa and 75 kDa respectively in the cytosolic fraction, but absent in the MRO fraction, that is consistent with the absence of this protein in the mitochondria. In the blot of Nbp35 protein, a double band appears, which could be a result of degradation of the protein. The full-sized western blots are

shown on the right panel. Noted that in the case of the Nbp35 protein, a second band is present in an apparent molecular weight of 72 kDa (indicated with “**”) in the cytosolic fraction, which could be a result of dimerisation of the aforementioned protein. The *Blastocystis* SufCB antiserum shows specific detection of *Blastocystis* SufCB with an apparent relative molecular mass of 77 kDa in the cytosolic fraction as well (positive control). The heterologous anti-*Trichomonas vaginalis* IscS antisera shows specific detection of *Blastocystis* IscS (4) with an apparent relative mass of 47 kDa in the MRO fraction (MRO positive control).

Figure S6:

Immunolocalization of Nbp35, Nar1 with MitoTracker in *Blastocystis* sp.

a. Cellular localization of the Nbp35 and MitoTracker in *Blastocystis* cells. i. Rabbit anti-yeast Nbp35 antibody (1:200) detects *Blastocystis* Nbp35 protein. ii. Localization of MitoTracker staining. iii. Overlapping of the previous images showing the localization pattern of Nbp35 and no co-localization with MitoTracker iv. Differential interference contrast (DIC) image of the cells used for immunofluorescence. **b.** Cellular localization of the Nar1 and MitoTracker in *Blastocystis* cells. i. Rabbit anti-yeast Nar1 antibody (1:100) detects *Blastocystis* Nbp35 protein. Localization of MitoTracker staining. iii. Overlapping of the previous images showing the localization pattern of Nbp35 and no co-localization with MitoTracker iv. Differential interference contrast (DIC) image of the cells used for immunofluorescence. Scale bar, 5 μ m.

Figure S7:

Immunolocalization of Nbp35 in *Blastocystis*. Image demonstrating the cytosolic localization of Nbp35 in *Blastocystis* cells by transmission electron microscopy. Inset is enlargement of specific regions of the cell, focusing in the localization of Nbp35 within the different compartments of the cell but also the distribution of the protein within the cytosol.

Figure S8:

Immunolocalization of Nar1 in *Blastocystis*. Image demonstrating the cytosolic localization of Nar1 in *Blastocystis* cells by transmission electron microscopy. Inset is enlargement of specific regions of the cell, focusing in the localization of Nar1 within the different compartments of the cell but also the distribution of the protein within the cytosol.

Figure S9:

Immunolocalization of Tah18 in *Blastocystis*. Image demonstrating the cytosolic localization of Tah18 in *Blastocystis* cells by transmission electron microscopy. Insets are enlargement of specific regions of the cell, focusing in the localization of Tah18 within the different compartments of the cell but also the distribution of the protein within the cytosol.

Figure S10:

Identification of functionally important structures in eukaryotic Dre2 sequences. Sequence alignment of Dre2 protein sequences from microbial eukaryotes demonstrating

the highly conserved C-terminus that contains two Fe/S cluster-binding motifs (CX₂CXC and a CX₂CX₇CX₂C in red), and the more divergent N-terminus with a methyltransferase domain.

Figure S11:

Phylogeny of the cytosolic Fe/S cluster assembly protein Dre2. The phylogenetic tree was inferred by Randomized Axelerated Maximum Likelihood (RAxML). A total of 78 taxa and 96 characters were used for the final phylogenetic analysis. Numerical values at the nodes represent branch support in the form of bootstrap (BP). Only bootstrap support values greater than 50 are shown.

Figure S12:

Phylogeny of the cytosolic Fe/S cluster assembly proteins Nar1 and iron hydrogenase. The iron hydrogenase protein sequences were used to root the tree. The phylogenetic tree was inferred by Randomized Axelerated Maximum Likelihood (RAxML). A total of 81 taxa and 233 characters were used for the final phylogenetic analysis. Numerical values at the nodes represent branch support in the form of bootstrap (BP). Only bootstrap support values greater than 50 are shown. Red line indicates the *Blastocystis* spp. homologues.

Figure S13:

Phylogeny of the cytosolic Fe/S cluster assembly protein Cia1. The phylogenetic tree was inferred by RAxML. A total of 59 taxa and 225 characters were used for the final

phylogenetic analysis. Numerical values at the nodes represent branch support in the form of bootstrap (BP). Only bootstrap support values greater than 50 are shown. Red line indicates the *Blastocystis* spp. homologues.

Figure S14:

Phylogeny of the cytosolic Fe/S cluster assembly protein MMS19. The phylogenetic tree was inferred by Randomized Axelerated Maximum Likelihood (RAxML). A total of 85 taxa and 53 characters were used for the final phylogenetic analysis. Numerical values at the nodes represent branch support in the form of bootstrap (BP). Only bootstrap support values greater than 50 are shown. Red line indicates the *Blastocystis* spp. homologues.

Figure S15:

Phylogeny of the cytosolic Fe/S cluster assembly protein Cia2. The phylogenetic tree was inferred by Randomized Axelerated Maximum Likelihood (RAxML). A total of 60 taxa and 98 characters were used for the final phylogenetic analysis. Numerical values at the nodes represent branch support in the form of bootstrap (BP). Only bootstrap support values greater than 50 are shown. The groups of the putative Cia2A and Cia2B are indicated. The *Homo sapiens* sequences are depicted in red and are the only Cia2A and Cia2B proteins that have been characterized to date.

Table S1

Distribution of homologues of the CIA machinery among publicly available genomes or expressed sequence tag data from eukaryote taxa. Table includes the accession numbers for all proteins reported in this manuscript.

Supplementary References:

1. **McGuffin, L. J., K. Bryson, and D. T. Jones.** 2000. The PSIPRED protein structure prediction server. *Bioinformatics* **16**:404-405.
2. **Netz, D. J., A. J. Pierik, M. Stumpfig, E. Bill, A. K. Sharma, L. J. Pallesen, W. E. Walden, and R. Lill.** 2012. A bridging [4Fe-4S] cluster and nucleotide binding are essential for function of the Cfd1-Nbp35 complex as a scaffold in iron-sulfur protein maturation. *J Biol Chem* **287**:12365-12378.
3. **Srinivasan, V., D. J. Netz, H. Webert, J. Mascarenhas, A. J. Pierik, H. Michel, and R. Lill.** 2007. Structure of the yeast WD40 domain protein Cia1, a component acting late in iron-sulfur protein biogenesis. *Structure* **15**:1246-1257.
4. **Tsaousis, A. D., S. Ollagnier de Choudens, E. Gentekaki, S. Long, D. Gaston, A. Stechmann, D. Vinella, B. Py, M. Fontecave, F. Barras, J. Lukes, and A. J. Roger.** 2012. Evolution of Fe/S cluster biogenesis in the anaerobic parasite *Blastocystis*. *Proc Natl Acad Sci U S A* **109**:10426-10431.

Figure S1

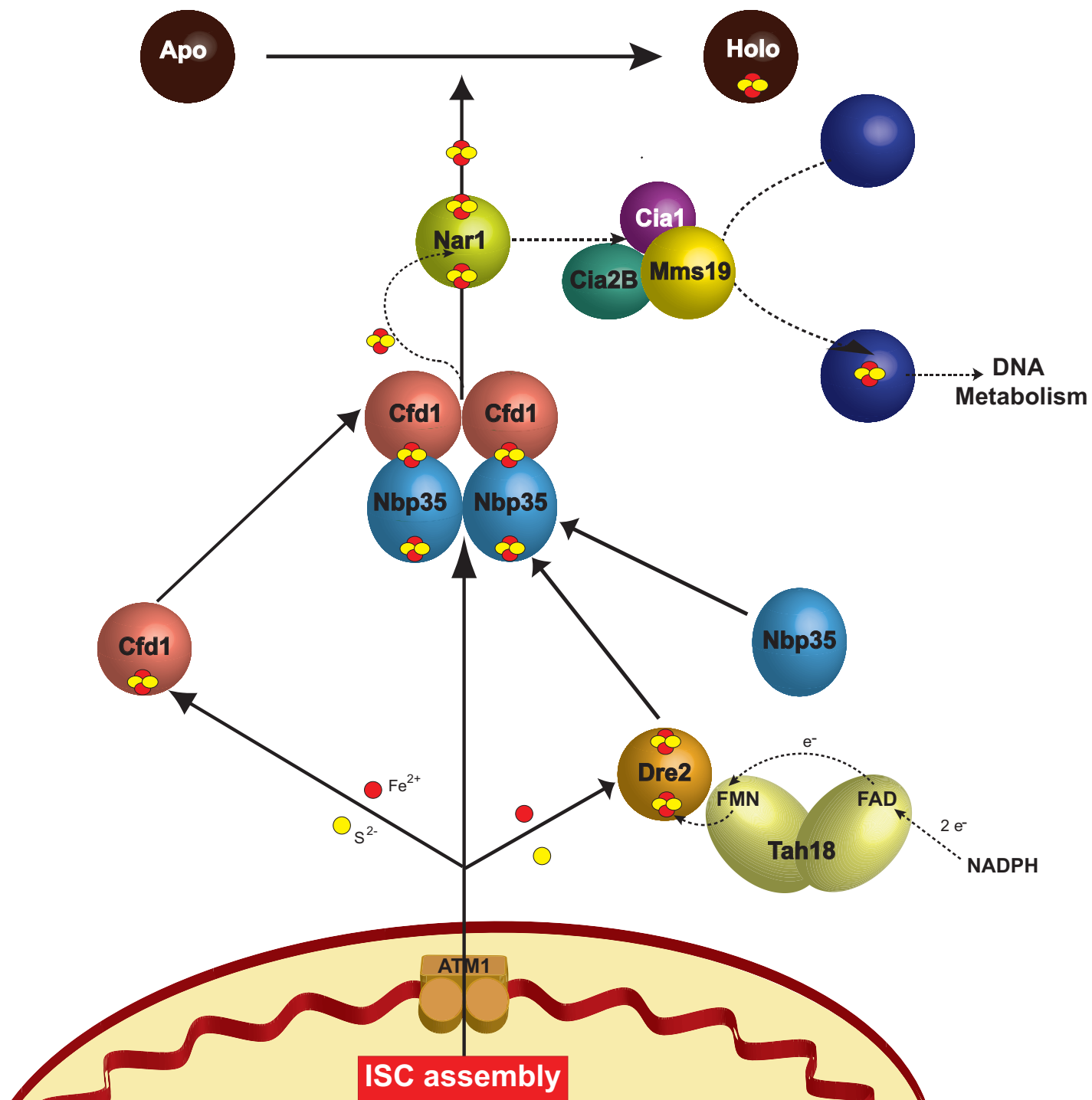


Figure S2

<i>Homo</i>	-----MKDS	β1	LVLGRV	----	PAHPDSR	β2	CWFLAW	NPAGT	β3	LLASC	β4	GDRR	IRIWGT	-----EG
<i>Saccharomyces</i>	-----MASI	NLIKSL	----	KLY	----	KEKI	WSFDF	SQG	----	ILATG	----	SDRK	IKLVSV	YDD-----
<i>Blastocystis S1</i>	-----MPS	VRLSSF	----	KAH	----	EGVA	WYVSW	SPDGN	----	ELLTC	----	SDRS	VKLWGR	-----MG
<i>Blastocystis S7</i>	-----MSRI	TLLSSE	----	QAH	----	DSRA	WYSSW	SPDGN	----	EILTC	----	GADKS	VKIWGR	-----MG
<i>Phytophthora</i>	-----	MGV	----	RRH	----				----		----			-----
<i>Phaeodactylum</i>	-----		----		----	EP	AWQVAF	SRDGR	----	YLAVCY	GAIEP	CVRIWQ	QSSPFH	-----ED
<i>Arabidopsis</i>	-----	MKVM	KNLGL	EEVQKI	----	EGH	----	TDR	VWVAV	NPAADG	VIASC	----	SADKT	VRIWEQSS
<i>Micromonas</i>	MLSFLGRGKPSDPPPEPSPAPDPPPASSAPPPSAPPD	FSDAKAVT	IELVAEL	----	EGH	DDR	VWGMQ	EPRGR	----	CLAST	----	SSDKT	CRLWSQSA	-----AG
<i>Cyanidioschyzon</i>	-----		----		----	MQW	EWLRVDV	----	----	ORE	----	HET	LWCATW	SPNGQ
<i>Dictyostelium</i>	-----	MTDT	TKNDKYN	LKLIDSMQ	KEAP	----	YDK	VWNLAW	HPNGE	----	ILATC	----	ANDKY	IQIWSKD
<i>Plasmodium</i>	-----	MT	VEFVVNL	----	ENH	----	KRR	IWSICW	SPDGN	----	FLASV	----	GADKY	ITIWV
<i>Cryptosporidium</i>	-----	MGSL	DKLGI	----	GVL	----	DSA	IWSVAS	HPKDR	----	IIASC	----	GS	----
<i>Tetrahymena</i>	-----	MIEEK	MEEQKE	FVKCIGOL	----	NGH	----	TDK	IWSVSW	HPTLD	----	IFATC	----	SDKT
<i>Naegleria</i>	-----	MT	NDL	LAALT	TOE	ISAVE	GNV	----	SDH	----	EES	VWSIAW	HPKYSN	LLATC
<i>Drosophila</i>	-----	MGR	----	LLEHTL	----	GGH	----	KGR	IWGVAV	HPKGN	----	VFASC	----	EDKA
<i>Giardia</i>	-----	MV	NDRV	SLLNHT	----	AAH	----	TD	IWR	LRA	SHTGE	----	LVASC	----
<i>Trypanosoma</i>	MLSS	YFSWME	DVDAR	VLRAGG	SDNH	MRPLY	LPEAP	LLRV	EPVCTL	----	----	----	RDH	----
<i>Entamoeba</i>	-----		----		----	MQL	----	VDSF	----	----	----	----	----	----
<i>Encephalitozoon</i>	-----		----		----	MK	----	YRITS	----	KKLGE	----	KILAV	----	HAGK

<i>Homo</i>	DS	----	WICKSVL	S-E-GH	QRT	VRKVAW	S-P-CG	NYLASAS	FDAT	TCI	WKK	KNQD	----	DF	ECV	TTL	EGHENE	-----	
<i>Saccharomyces</i>	DS	----	FTLIDVLD	ETAH	KKA	IRSVAV	R-P-HT	SLLAAGS	FDS	TVS	IWA	KEESADR	----	TF	EMD	LLAI	EGHENE	-----	
<i>Blastocystis S1</i>	DS	----	WAVIETH	D-G-IH	KKT	IRSCMS	P-DG	NYIAAAS	FDS	NSI	VYH	KTDG	----	EW	TF	TML	EGHVNE	-----	
<i>Blastocystis S7</i>	DS	----	WVLKETL	D-G-LH	SKS	IRSCMS	P-DG	KYIAAAS	FDS	NSI	VYR	XENE	----	EW	KL	TML	EGHVNE	-----	
<i>Phytophthora</i>	DS	----		----		SRGR	TD	PHHPSL	S-P-DG	RYLASV	FDGT	TVI	WEKQGS	----	SY	EVI	SSL	EGHESE	-----
<i>Phaeodactylum</i>	SG	----	WILDATL	T-G-IQ	TRT	IRSI	AFA	PIRTP	LLASAS	FDGT	VAV	WEHY	PATNGALV	TASAK	SPSG	----	VDE	WE	CTAQL
<i>Arabidopsis</i>	SG	----		----		TR	----	SWT	----	CKLGH	RLGS	FDGNT	TCV	WEN	FAT	----	DSE	VS	VL
<i>Micromonas</i>	GN	----	WVTVAEL	E-G-VH	NRT	VRQVSW	S-P-CG	RLATAS	FDAS	TAV	WT	QSGG	----	D	WE	CV	AVV	EGHENE	-----
<i>Cyanidioschyzon</i>	GN	----	GLLLGGLG	GD	FARS	VRRIDW	S-P-CG	LALAC	FDSK	VRVY	RLV	EGVSG	GNGLSL	EREPANA	----	VLA	AYL	REL	VATL
<i>Dictyostelium</i>	GK	----	WGLVOSLE	----	GHEKT	VRRVAV	S-P-CG	RFLAGAS	FDAS	TSI	WEK	SKD	----	EL	FT	HV	SSL	EGHYE	-----
<i>Plasmodium</i>	AGSIE	FDVY	DI	E-T-N	HEKS	LRHIEF	S-R-DG	SFFVVAS	FDSK	CS	IYK	KNNN	----	D	KV	FY	KSL	EGHEKE	-----
<i>Cryptosporidium</i>	NS	----	WVKAYEFG	SLE	HKRL	IRKIAW	S-P-CG	GMI	ISAS	FDSS	ISV	WEFVSR	----	D	IG	W	AC	ICKI	
<i>Tetrahymena</i>	NQ	----	YELKQTS	D-TH	ERT	IRTLAF	S-P-DG	M	LACGS	FDST	ISI	YANNG	----	S	FE	FV	SKL	EGHEHE	
<i>Naegleria</i>	SGR	LF	AKCIDVLE	N-QH	NRT	IRRV	SLP--	SGNAL	LACAS	FDGT	SSI	WILLQ	NHLQALEEESQ	NSKES	SPTT	SANL	GLLKCV	STL	
<i>Drosophila</i>	NT	----	WSTKTI	S-D-GH	KRT	IREIRW	S-P-CG	QY	LASAS	FDAT	TAI	WSK	SSG	----	E	FE	C	NATL	
<i>Giardia</i>	----		LHLVQRT	RP	GYDPVT	VRDCAF	S-A-ND	Q	HLVAA	YDGS	I	YVD	DLINEL	TQ	----	ND	PF	OL	
<i>Trypanosoma</i>	GT	----	WSCIYTL	E-G-EH	SRT	VRHVSW	S-P-SG	T	FI	AC	SDRT	ASV	WR	RASD	----	DP	NC	F	
<i>Entamoeba</i>	----		LKPL	T-PH	NRT	IRR	VKS	S-K-NG	LL	AC	SDST	VSL	WEL	NEN	----	TI	I	G	
<i>Encephalitozoon</i>	----		LVNQDTG	EV	MCR	CKKS	VR	SLA	----	S-HGR	Y	VCCGS	YDCT	AVL	----	DG	K	V	

<i>Homo</i>	----	VK	SVAV	APSG	----	N	LL	ATCS	RDKS	VW	WEV	----	DEE	----
<i>Saccharomyces</i>	----	VK	GVAV	SN	----	Y	LL	ATCS	RDKS	VW	I	WE	----	DE
<i>Blastocystis S1</i>	----	VK	NVCW	SHDS	----	Q	L	LASTS	SR	----		----		
<i>Blastocystis S7</i>	----	VK	NVCW	SRDG	----	K	L	ASCS	RD	RNI	I	V	----	NEE
<i>Phytophthora</i>	----	VK	SVAV	SPSG	----	S	Y	L	ATCS	RDKS	VW	I	WEA	----
<i>Phaeodactylum</i>	----	VK	C	VQ	----	S	L	ASCS	RD	KT	VW	I	WE	----
<i>Arabidopsis</i>	----	VK	SV	W	----	S	L	ATC	GR	D	K	S	VW	----
<i>Micromonas</i>	----	VK	SC	AW	----	T	L	ATC	GR	D	K	S	VW	----
<i>Cyanidioschyzon</i>	----	VK	A	AV	----	T	L	ATC	AR	D	K	T	VW	----
<i>Dictyostelium</i>	----	VK	SV	AW	----	T	L	ATCS	RDKS	I	W	I	W	----
<i>Plasmodium</i>	----	VK	C	AS	----	K	Y	I	V	T	C	GR	D	S

	----	K	Y	I	V	T	C	GR	D	S	I	W	I	H	A	K	T	E	E	V	S	R	K	E	D	D	Q	G	V	I	A	S	N	G	G	E	A	H	P	N	G	I	H	N	E	E	Q	T	K	E	S	Q	N	G	A	K	T	E	A	Q	C	G	N	P	G	E	T	H	P	N	A	E	P	Q	N	V	N	D	I	D	A	D	A	D	V	D
--	------	---	---	---	---	---	---	----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Cryptosporidium ---VKCVDWSPFN-----N FVAACRDRAIWFFSLDIGENRKLG-----T
Tetrahymena YKKVKCVAW DSEG-----KFLASC SRDKT VVVVDY-----ENG-----
Naegleria ---VKSVAV NYKSASLMDQSDHDGEDGDCG LLATCG RDKT VVWWEA-----IDK-----VG
Drosophila ---VKSVSW SRSG-----GLLATCS RDKS VVIWEV-----AGD-----
Giardia ---IKSIDI SKEG-----TVATCS RDRF VSFWRP-----CSDS-----
Trypanosoma ---VKCAAW GTDN-----TLATCS RDRT VVVWDR-----VDV-----
Entamoeba ---VKCVDW SFGS-----NMVATCS RDK SVWLWKS-----YSG-----
Encephalitozoon ---VKCVAF SEDG-----RYLAMAT RGRS VVVVKI-----D-----

β13 β14 β15 β16 β17 β18
Homo -DE-----YECVSV LNSHTQD VKHVV-----WHP-----SQE LLASAS YDDT VKLYRE-----EEDD NVCCATLE-----GHEST VWSLAF DPS
Saccharomyces -EE-----YECISV LQEHSDQ VKHVI-----WHP-----SEA LLASSS YDDT VRIWKD-----YDDD WECVAVLN-----GHEGT VWSSDF DKT
Blastocystis S1 -----
Blastocystis S7 -DDEMDSCSRD RCVSV LSGHTQD VKFVR-----FNP-----ANN NLYSCS YDD IIKVWLY-----D-GED FSNIKHA-----PHEGT VWGLAF NHR
Phytophthora -TD-----FECISV LH AHMQD VKFVA-----WHP-----KED LLVSAS YDDT IRIWAE-----NDDD NYCKETLT-----GHTAT VWGVAL SPQ
Phaeodactylum -GD-----FECIAV L HGHEGD VKCQVF TSSHDEWGD-----GDE ILLSSS YDNT IKCWAE-----DAGD NYCAASIED-----VHSST IWSLAM SPS
Arabidopsis -DE-----FDTI AVL TGHSED VKMVL-----WHP-----TMD VLFSCS YDNT IKIWC S-----EDEDGD Y NCVQTLSEL-----NNGHSST VWSISFNAA
Micromonas -ND-----FECVAV L NGHSDQ VKCVT-----WHP-----TED VLVSTS YDDT IKIWT E-----DPDGDW SCSKTL SKE-----DGGHEST VWCAST EPG
Cyanidioschyzon -LD-----FECVAV LA GHTQD VKSLV-----WHP-----RTE LLASAS YDNT IRLWCE-----DVIDGE NYCCAVLS-----GHEST VWSVAF APT
Dictyostelium -ND-----FECLSI N SGHGQD IKCVI-----WHP-----NEE LLASSS YDDT IKFWK D-----IDGD WECINTLT-----GHES IWDLAF NKD
Plasmodium DFD-----FSFDAY L TAHSEDI KFVA-----WCPL-----SENTF I SLS YDNS LKIWSK-----QISEWNCIQTLN-----EHTSV WCVAF NFD
Cryptosporidium PIE-----YDCIGV VTAHTND IKKIK-----WHPT-----IPMVL LSCS YDNT IIAWAPSSQLLGH-DEVKSLEWVKLYTLN-----GHSST VWDFTY SPN
Tetrahymena -FD-----FSCYSV IDAHTQD VKHVK-----WIP-----GTNNLASTE FDD KKLWEQ-----EDDD NKCSATYS-----NHSAT WCVVEF SKT
Naegleria FSD-----FDCNSV CSGHTQD VKFVA-----WHP LTRSNMPS LLYSAS YDNT IRIWKE--GGFEEDHDRQSDWKCVGILR-----GHTST VWGLAF EPQ
Drosophila -DE-----FECAAV L NPHTQD VKRVV-----WHP-----TKD VLASAS YDNT IKMFAE-----EPIDNDW DCTATLT-----SHTST VWGIDF DAD
Giardia -PD-----YDCIGLE NNHTEDI KCVR-----FNK-----NGNYLVSAS YDNNMCLYKRCIEMADELGEEIEES NVLVGATKSELDLNS CELNAPDSE ILLA SKSSCHT VVTAIF TAN
Trypanosoma -GE-----FECAGV LAGHAQD VKACA-----WLFPSGGVEKP LLFSCS YDNT VKVWTE-----SHKRDD NYCHQTLT-----RHDGT VWSVAV QPI
Entamoeba -ID-----YECDSV L TGHSGD VKTVL-----FHP-----SGT I LFSGS FDGT IKVVKG-----EETEWS ELOTLIQ-----AYGKT VWDLKIT KE
Encephalitozoon -GE-----IEIDGV I EDHLHD VKGCI-----FHG-----GLLFTY G YDNT VKVYDR-----FDYDSS WELVQ SID-----ERST WCVVIF HNG

β19 β20
Homo -G-----QRLA-----SCS DDRT VRIWRO Y-----
Saccharomyces EG-----VFRLC-----SGS DDST VRVVK-----
Blastocystis S1 -----
Blastocystis S7 -G-----QVRL VVRL SCGM DGH VVLWK-----
Phytophthora -G-----TEMA-----SVS DDTD VVIWQ-----
Phaeodactylum -G-----LRMI-----SGS DDQS LGIYKCY-----
Arabidopsis -G-----DKMV-----TCS DDLA VKIWKTD-----
Micromonas -G-----AHRVV-----TCS DDRT IAVWN-----
Cyanidioschyzon -A-DA-----GYEHLA-----SAGADGR LLLWQRC-----
Dictyostelium -G-----DKLV-----SCG EDKL VLFWKFD-----
Plasmodium -G-----SEFA-----TCS DDKS IRIWK SEKKI WYN KHKYPFL YQHIVKD ARKSSNESS LCSL INKSNGTLGGSAAATA VKQALA NDSRGYSNHNSKDSANGKGSNDR
Cryptosporidium -G-----EFLI-----SCS DDSS IVLWNSN-----
Tetrahymena -G-----QYMA-----SCG DDQK IKVYKKN-----
Naegleria LSDDP-----EYPQ YMV-----SVG DDKS LILWRED-----
Drosophila -G-----ERLV-----SCS DDTT IKIWKAY-----
Giardia -G-----NSIL-----TVD GNGC VRCYSIV-----
Trypanosoma EQPIDSLQMEEEER EGTALEYSPVVC-----CSS DDKT VTFWSRD-----
Entamoeba -G-----KFIV-----AGCANGV ILYE-----
Encephalitozoon -----RMV-----CTTEEGT VSIYALRS-----

Homo
Saccharomyces

Blastocystis S1 -----
Blastocystis S7 -----
Phytophthora -----
Phaeodactylum -----
Arabidopsis -----
Micromonas -----
Cyanidioschyzon -----
Dictyostelium -----
Plasmodium NFLQKTTKLFKRKIREPMTQGGGDKTSDAPSNMP **ILDY**SKDD**KFVQAALKAVVQ**NNFVPLYFDHGL**FKFVYK**YADVEGAFQAQKGSTEKREEQGGHTEKESPNGAHPIEKNEPRNDECISK
Cryptosporidium -----
Tetrahymena -----
Naegleria -----
Drosophila -----
Giardia -----
Trypanosoma -----
Entamoeba -----
Encephalitozoon -----

Homo -----LPGNEQGVACSGSDP-----SW**KCICT**
Saccharomyces **YMGDEDDQ**-----EW**VCEAI**
Blastocystis S1 -----
Blastocystis S7 -----
Phytophthora -----YDSNSK**EV**NEDGGSK-----QW**KLKFT**
Phaeodactylum -----TASEKRRHFPDEGKNRNG-----LW**KCVGH**
Arabidopsis -----ISRMQSGEGYV-----PW**THVCT**
Micromonas -----AAGAR**RVSS**SSLARFEDV**ASFSRFA**SSRGRS**LR**IAS**FVAGQTV**SSALHMRSRGPRGADLS**IACK**
Cyanidioschyzon -----EDGTR**PE**TIRHDAL**VEDTVSV**RTPTQPHSEKGVGAGTPQPTDEGVMAEAPSSARRESG**SW**SLDL**SSTRDA****LASAF**IEKDP**EEHARF**AEATQV
Dictyostelium -----KENE-----KW**INIFK**
Plasmodium VSKDGGDQPNTHSAEVEHKNDANDVSQLSNGEKTNKQSGCTT**QDEQ**QTSNPVSENNVSSMSAIPNLS**DK**EKDLKNVLFD-----DW**KIKHV**
Cryptosporidium -----QGNEN**KFKN**LNSVN**FAL**TD**FKMIF**YNTPTN**KRLSK****YIQID**QANS**FIN**NY-----
Tetrahymena -----ENGAFSS-----PY**IVETT**
Naegleria -----VVGNYIDM-----NV**TOVQT**
Drosophila -----HPGNTAGVATPEQQT-----VW**KCVCT**
Giardia -----DDG-----VK**QIGCT**
Trypanosoma -----GNG-----N**FRSVCT**
Entamoeba -----FKDN-----LL**VELDT**
Encephalitozoon -----G**WTLEMS**

Homo -----ISGFHSRT**IYDIA**WC-----QLT**GA**L**ATA**CGDDA**IRV**FQE-----DPNSDPQ-----QPT**FSLTAHL**HQAHSQ-D**VNCVA**
Saccharomyces -----LPDVHKRQ**VYNVA**WG-----FNGL**IAS**VGADGV**LAVYEE**-----VDG-----EW**KVFAKRAL**CHGVYE**INVVK**
Blastocystis S1 -----
Blastocystis S7 -----
Phytophthora -----ISNCHERT**IFSVD**WS-----KHGA**FLVT**GAADNA**IRV**FQ-----QPNDTPS-----SFDL**AICQ**KEAHAS-D**INCVR**
Phaeodactylum -----IPDAHLAS**IFSV**AYA-----PSRAGHG**RIATA**GADNR**IQIF**RE-----VSGSVSD-----QPL**FTVET**SATNELGD**VNCVS**
Arabidopsis -----ISGFHDRT**IYSV**HWS-----RDGVI**ASG**AGDDT**IQL**FVD-----SDSDSVD-----GPSY**KLLV**KK**KAHEM**-D**VNSVQ**
Micromonas -----FPCGHRDP**VLVNH**WG-----RNGL**IAA**AGGDNS**VRV**YAA-----M**KSD**SGSHGSGR**AWAEV**GA**IGEH**HE**DGHLD**-D**VNTVA**
Cyanidioschyzon DSSTPRNLD**PPLVSSVMS**G**KQR****NLVIKQVR**VNAGDAGFADEE**DES**V**YCV**DWS-----HDGR**LLATA**CADGH**IRV**YEA-----TEL**RILLD**I**PAAH**S**GAEVN**Y**VQ**
Dictyostelium -----FKNENSRP**IYSID**WS-----SLTNT**IVT**GSADDS**IIF**YEQ-----ESDDTPD-----KY**KILL**KK**NAHDS**-D**VNCTK**
Plasmodium -----TEGYHKRS**ISYLD**WN-----AYED**LIA**ASSFDNS**LKIF**QK-----NLD**TWNL**IENIE**NAHLS**-D**VNCVV**
Cryptosporidium -----DKELY**SYP****IYSI**EW**C**-----NYIN**CI**IVSSADK**SLHL**FSV-----TDSK-----RL**KHIC**E**K**PN**AHNS**-E**IN**SVS
Tetrahymena -----TKNAHART**IYSL**SFS-----EDAT**FLAS**V**GADNT**LV**YQ**KNMY**VTT**EGQD-----NNL**YELLE**KK**VNCH**FA-D**IN**CVA
Naegleria -----ISDVHTRT**IYV**D**WC**VYKHPS**TGQ**SIS**LVATA**GDDNT**IAI**YQF-----DT**TTR**-----QL**KLLT**K**IAN**AHDS-D**IN**CCI

Drosophila -----VSGQHSRAIYDVSWC-----KLTGLIATACGDDGIRIFKE---TSDSKPD-----EPTFEQITAEEGAHDQ-DVNSVQ
Giardia -----LHGQRPIYDISLVEPRNAPK-AFNTRYVATGQDGTICLSAI---TITTGVA---VPIVCIIGAHDG-EVNAVC
Trypanosoma -----ASGFAERTIYSVGWA---PCGSDVSPAIVACGSGDNKVTLLGV---YQSRGYE---EVHVSVVAEVPSAHEA-DVNTVA
Entamoeba -----LNNEKYRDIYSIDIN-----DNNVLVSGSDNAIRLFKI---NTIKK-----KLELEIEEKQDAHTN-DVNCVK
Encephalitozoon -----RKLVLPLIYSI-----CSVGENMAYVLMRSTI GIVDS-----NLNLVMSIENVHED-SINSIV

β27 β28
Homo -----WNP-----KEPGLLASCSDDGEVAFWKYQRPEG---L
Saccharomyces -----WTEL-----NGKTILATGGDDGI VNFWSLEKAA-----
Blastocystis S1 -----
Blastocystis S7 -----ELLTEVGEDDFI DALEQ-----
Phytophthora -----WSPQLLEDKG-----KKTFLLASAGDDALVRIWSMTI-----
Phaeodactylum -----WHP-----SDGSILATAGDDGSVCIWKFNL-----
Arabidopsis -----WAPD-----KESRLLASASDDKMVKIWKLASEP-----
Micromonas -----WHP-----TDPTCLASCSDDGLIKIWKVTPAEE-----
Cyanidioschyzon -----FOKRRDACGDWSIHLGKRFQDMSTETLQRSYLLASTGDDGRLRVWVLF-----
Dictyostelium -----WNP-----KFKNILASCGDDGFIKIWELQDK-----
Plasmodium -----WCPQKY-----QDYFLLATAGDDCVINIWKYTKG-----
Cryptosporidium -----WIND-----DKKGEFISAGDDGEIALWRFDPE-----
Tetrahymena -----FHP-----SKDILVTVSDDRQIKLWSVEINL-----
Naegleria -----WVK-----NEFGLSSCSDDGAVKFWKLRM-----
Drosophila -----WNP-----VVAGQLISCSDDGTIKIWKVSE-----
Giardia -----DVTQFAGA-----DGHIIVCSGGDDGYINMWNISIEERDLFS
Trypanosoma -----FSRSTNELWGDNSRG-----GEGLLASGGDDNI VRIWRVTAAL-----
Entamoeba -----WIN-----KTLISISVGGDDNMLKIWKIVN-----
Encephalitozoon -----YD-----EGRNRIVSGGDDGILNTIELL-----

Figure S3

	Ferredoxin-like domain	
<i>Homo</i>	-----MASPFSGALQLTDLDDDFIGPS-QE	CIKPVK-----VEKRAGSGVA-----KIRIED-----DGSYFQINQDGGTRR-----LEKA
<i>Saccharomyces</i>	-----MSALLSESDLNDFISPA-LA	CVKPTQ----VSGGKKNVNMNGE----Y--EVSTEP-----DQ-----LEKV
<i>Blastocystis S1</i>	-----MSYFSGTVLLGELDDFIAPA-QA	CSYGVF-----GDATSTSGGKM-----QLVMED-----DLGDDYGSATSSVIQTNV-----AKVA
<i>Blastocystis S7</i>	-----MSYFSGKVLGDLDDFIAPA-QA	CSYGLF----SDSGSSTHRGA-----QLVMED-----DLGESFSKA--SVIRSSA-----AKIA
<i>Phaeodactylum</i>	-----MSGVFLSNVDDYLAPS-QA	CVNPLF----STDKKKDDKESG-----VVGTLNSGNHANDDPNSFDTAAASENPAIVPRKRVRRLPAAITASSDWTTPRVPKDVP
<i>Phytophthora</i>	-----MASVFLGDLNDYIQPS-QA	CVNPLF----TSDKSENGSSNGL-----A--KITLET-----ELSAADFAVPQPVKPNIIIRTTT-----QEKA
<i>Arabidopsis</i>	-----MSEKFSPTLRLGDLNDFIAPS-QA	CVTSLKDSKPIVKKSDRP-----QVVVAP-----KQQ-----LEPV
<i>Micromonas</i>	-----FSGAVKLGDLNDFINPS-QN	CVVALTAGEVILQRRGAPPS-----DLPPGA-----PTIYGDPS-----APTQ
<i>Dictyostelium</i>	-----MAEKFSVLLKTELDLDFITPS-QE	CIKPMI----IDKKNST-----QFTIES-----DGSYVETSDGKQVQ-----MEKA
<i>Tetrahymena</i>	-----MFSGTIKIASLDDYISPS-QE	CILPIF----DKNSKLTTE-----DPTVKA-----YGMIPQKPDLIKTKTA-----KQTA
<i>Naegleria</i>	MSSYIPQQEPTMSTALMLADLDDFLPMNRQD	CIKPFLL----MTNSKSDSTTSSKNITSENGTAAVSVSLDN-----DDSVGGVSNRNPTRRIGRSNNRGTINLNSQQAEPKkia
<i>Encephalitozoon</i>	-----MDALIRPPMSFFADLP-KDN-KK	CIKI-----KIKI-----DSFEESESGLNK-----GSPIL
<i>Drosophila</i>	-----MSRLSTALQLTDIDDFITPS-QI	CIKPVQ----IDKARSKTGA-----KIKIKG-----DSFEESESGLNK-----LNKV
<i>Giardia</i>	-----MSLKVVASDLNLTLP-EE	CVVPLK----PADGPSTGTV-----KLRLKA-----CDPVPIG-----STPV
<i>Entamoeba</i>	-----MSLSVGLQIAGVDDYIQQN-LV	CVMPLK----ETPPQEHKGA--KISLGG-----PEEGNELPK-----LTKV
<i>Cryptosporidium</i>	-----MFSTAVKLANLDDYLESS-QD	CIIVSLL--SDKDDTKP-----KIAVMR-----PAKAQDNKDDKKSQT-----SDKA
<i>Cyanidioschyzon</i>	MPRDT--QSRFSTGLRVLDLDFLPS-TA	CIKPLQ----GGLPPGSAV-----APVYSH-----NAHDANGETN-----TPVA
<i>Trichomonas</i>	-----MSADPAASTS-FD	CLHPVS----IEERG-----RVKADD-----EATF
<i>Trypanosoma</i>	---M--SANNFASLMLAGM-DYIAPS-EA	CIILPTK----LQGGTSSD-----SVKRHG-----AG-----NEAV
<i>Homo</i>	KVSLNDCLACSGC-ITS	AETVLIITQSSHEELKVKLDANKM-----AAPSQQ--RLVVVSVSPQSRASLAARF-----QLNPT-DTARKLTSF
<i>Saccharomyces</i>	SITLSDCLACSGC-ITS	SEEILLSSQSHSVFLKNWGKLSQ-----QQD--KFLVVSVSPQCRLSLAQYY-----GLTLE-AADLC LMNF
<i>Blastocystis S1</i>	SVSLADCLACTGC-VTS	AETVLIQQSLQTFDLDELKAKAH-----KLYVAIVSRPSCVSLANKL-----GIPID-EAFERIRSV
<i>Blastocystis S7</i>	TVTLEDCLACSGC-VTS	AETILIQQSTSTFVEELKERKH-----KLYIVIVSRPSCVSIADAL-----SISPD-EAYQHITSV
<i>Phaeodactylum</i>	QASIADECLACSGC-VTT	AETVLLTQHSVVALKELIAKKE-----NDR--PKIVATISPAAWDLHRHLSREFNCSFSLSLA-QQRWITLL
<i>Phytophthora</i>	TISLDDCLACSGC-VTS	AETVLIQQSFREMLDVLATKEH-----KRVVVTLSPPQSRASLAHF-----EMPVV-AVHRKLVTL
<i>Arabidopsis</i>	KISLKDCLACSGC-ITS	AETVMLEKQSLDEFLSALSkg-----KDVVSVSPQSRASLAVHY-----DISPL-QVFKKLTTF
<i>Micromonas</i>	KVSLSDCLACSGC-VTS	AETVLLAQSADEFRRVRAAAM-----SGG--RMVVV-----
<i>Dictyostelium</i>	TITLNDCLACSGC-ITS	AESVLIQAQSTVEFSNVLKSIAE-----SKPD--SIVVVISPPQSRASLANHF-----GIDSM-QLHRKLVTF
<i>Tetrahymena</i>	KVTLSDECLACSGC-VTT	AETILIQQQSVEEFLKLQSY-----KHAVVGISQQARASMAHYF-----GLSEE-HIQRALTYF
<i>Naegleria</i>	KIELADCLACSGC-VTT	AESVLVNVQSQVEQFLTSLKEMKLFSPFIEKIQDKVNVSDLEDDLDSILLSGSKRVVKPQPNISFVITISQQSAASLSSYY-----QCSSVRECLSRLSYL
<i>Encephalitozoon</i>	ALSLSDECLACSGC-VS	ADEAGALSSEDLS--FVLDLS-----PQTSFVLSPPQSKINIFNLY-----REDGMEYR-EFEAVLSSF
<i>Drosophila</i>	DISLQDECLACSGC-ITS	AEEVLIITQSQEELLKVLQENSK-----NKASEDWDNRTIVITLATQPILSLAHRY-----QIGVE-DAARHLNGY
<i>Giardia</i>	KITINDCLMCSGC-VTS	AEEVFFRELNTALQNAITSGPK-----AG--RPIVLSLSQSAIILSLSSVL-LNITSVEPCTVD-TLQKLEYA
<i>Entamoeba</i>	TVRLEDCLACSGC-ITS	AETVLIQQGLPEFRKNIKELSQ-----R--KKVICITIADECIASMSVVH-----NQPFN-VVWTRVEKA
<i>Cryptosporidium</i>	TVNVADCLACSGC-VTS	AEAQLLEDQNVSEFMNILLKQK-----RLTVVVISNSQCSSFACHL-----NCDLI-TIQKRLSGL
<i>Cyanidioschyzon</i>	RVTLSDECLACSGC-ITS	AETVLLATHSDVNFVTAQCSQAD-----AFGAVVVAPAVVASLASVW-----KLAEIESVLERIQT
<i>Trichomonas</i>	KVTLQDECLACSGCAIT	KDEITIIISEQNTSRIFEKLDDEV-----KDYIVLVATHVVANLAAVR-----NWSAA-KAFSTIKQL
<i>Trypanosoma</i>	KITLQDECLACSGC-VTT	AETILITISQSRELLKDRALDPT-----RPFVVTISDQSAASIAAFL-----KTDVQ-KAFHIVSGF
<i>Homo</i>	FK-KIG-----VHFVFDTA-FSR	HFSLLESQREFVRRFRGQAD-----
<i>Saccharomyces</i>	FQKHQ-----CKYMGTE-MGRI	ISISKTVEKIIAHKKQKENTG-----
<i>Blastocystis S1</i>	LY-SHG-----VSVVVRQD-LGE	VIAMLESITEFDHRRRSTSSPSSISQDPSYADSTRSV-----YPKEGQLFSSLDIEHLSSTNDHYTLLSNPS-----
<i>Blastocystis S7</i>	L-KIG-----VDIVVRQD-IGC	ISLVESIAEYTKRASSAAVASVLTQPSYAVNQDLSMLLHLSILLHSHYKPNQVLTSIDLDHITESTDDYDVLSSQA-----
<i>Phaeodactylum</i>	WR-ALK-----ISSVLDGN-I	PLAWSLEEALEFCRAYKRKQT--TNDPDAMAVDVPQDELWQQQLIPSFASERSQSQYVNGET-----
<i>Phytophthora</i>	FR-NLG-----VTLVIDST-C	SDFALLESRAEFLHRYRNHQTIWARPPSSVAVSSAKTEFLE-----
<i>Arabidopsis</i>	LK-SLG-----VKAVFDTCS	RDLVLIESCNEFVSRYKQANS-----
<i>Micromonas</i>	-----VSVVLDTT-ASR	DLLESCEEFVSRYRNAKH-----
<i>Dictyostelium</i>	LK-SIG-----VNHVFDTS-FS	REFALIESAEFIARYKQTYD-----
<i>Tetrahymena</i>	FQDQLN-----	
<i>Naegleria</i>	FKVKFG-----AVAVFETSTLAR	LVSHLECEDFLNRYKEGK-----
<i>Encephalitozoon</i>	LRSKFN-----IHRIVDTS-YL	RSKIYEETIREYMAT-----
<i>Drosophila</i>	FR-SLG-----ADYVLSTK-V	ADDIALLECQRQFVDYRENE-----
<i>Giardia</i>	LRTRVADLRHCAYED	APPVYIVSEA-QHQEQSVLMNVRQISLLMQSSEP-----
<i>Entamoeba</i>	LK-KEG-----VDELRLDS-QA	QDISLFGIYDEFKEYQKMN-----
<i>Cryptosporidium</i>	FK-HIG-----ARFVMNST-I	SEYISLLETKEYEFISRYKAKSD-----

Cyanidioschyzon FD-KFG-----AFAVVLNS-VGR**C**LSVLET**CA**QAVERLQPDQNDG-----
Trichomonas FLSKGA-----QKVVLDTD--IQLVFRRLVVKFEFIENQTL-----
Trypanosoma FRAVLN-----ARYVSDLH-WALRISVEKTAEEY**CRRVR**CERE-----
H
Homo -----CRQALPLLASA**C**PGW**C**YAEKTHG**S**FILPH**I**STAR**S**PQ**Q**VM**G**SL**V**K**D**F-----F-----
Saccharomyces -----ADRKPLLSAV**C**PGFL**I**Y**E**TK**T**K**P**Q**L**V-PMLLN**V**K**S**P**Q**Q**I**T**G**SL**I**R**A**T-----
Blastocystis S1 YFQNESSHPSITANHPFIT**S**N**C**PGW**V****C**YAEK**K**TH**A**L**V**-PFL**S**T**T**K**S**A**Q**Q**I**M**G**S**L**L**R**K**M**-----L-----
Blastocystis S7 YHALPETS**P**HIGAN**A**PL**I**T**S**N**C**PGW**V****C**YAEK**K**V**H**T**L**V-PFL**S**S**V**R**S**A**P**Q**V**T**A**ALL**R**RL-----L-----
Phaeodactylum KTVYHDGGAQ**Q**AG**S**L**P**LL**S**G**S****C**PAV**V****C**L**V**E**K**S**T**H**K**A**V**-P**H**L**A**T**T**K**S**P**L**A**L**A**G**E**F**W**K**R**Q**H**F**-----
Phytophthora PSTTAN**P**LQ**D**PL**R**AMP**L**ASS**C**PGW**I****C**YAE**K**S**Q**P**N**A**I**-P**F**I**D**T**T**K**S**P**Q**Q**I**A**G**S**I**I**K**R**F**-----V-----
Arabidopsis -----DDGEN**S**Q**S**PL**P**V**L**S**S**A**C**PGW**I****C**YAE**K**Q**L**G**S**Y**V**L**P**Y**V**S**S**V**K**S**P**Q**Q**A**I**G**A**A**I**K**H**H-----L-----
Micromonas AAVTNT**P**PP**S**PA**D**V**L**P**V**L**T**S**A****C**PGW**V****C**YAE**K**T**H**G**G**A**V**L**N**H**V**S**A**V**K**S**P**Q**Q**V**M**G**C**I**V**K**R**K-----I-----
Dictyostelium -----K**L**P**L**M**L**S**A**S**C**PGW**I****C**YAE**K**T**H**G**D**Y**V**L**P**Y**I**S**T**T**K**S**P**Q**Q**I**M**G**T**L**V**K**Y**Y-----L-----
Tetrahymena -----V**Q**Q**S**V**P**V**L**S**S**E**C**PGW**A****C**YAE**K**A**V**G**E**F**V**I**P**Y**M**S**Q**V**K**S**P**Q**Q**V**M**G**S**L**V**K**N**S-----L-----
Naegleria -----P**V**F**A**S**A****C**PGW**I****C**YAE**K**T**Q**P**E**I**I**-P**S**I**S**T**V**K**S**P**Q**Q**I**M**G**T**F**V**K**K**F**I**T**S**N**M**K**Q**L**-----
Encephalitozoon -----N**H**L**I**V**S**A**C**PGV**V**T**Y**I**E**R**T**A**P**Y**L**I-G**Y**L**S**R**V**K**S**P**Q**Q**M**A**F**S**L**V**K**G**S**-----
Drosophila -----L**T**M**L**S**S**S**C**PGW**V****C**YAE**K**T**H**G**N**F**I**L**P**Y**V**S**T**R**S**P**Q**Q**I**M**G**V**L**V**K**Q**I**-----L-----
Giardia -----R**S**N**I**A**I**I**T**H**C**PAV**R**L**F**I**T**K**R**N**R**E**L**I-P**Y**I**V**S**T**A**S**P**M**E**L**F**G**A**S**Y**C**N**I**-----
Entamoeba -----K**V**L**L**T**S**T**C**PGW**V****C**Y**E**K**M**Q**G**K**M**F**E**Y**M**S**K**V**A**S**M**T**I**A**G**M**I**M**K**K**Q**-----
Cryptosporidium -----L**P**M**I**I**S**H**C**PGW**I****C**Y**E**K**S**L**N**S**S**V**L**P**L**L**S**K**V**R**S**A**Q**Q**L**Q**G**I**L**I**K**T**L**T**L**E**I**Y**N**Q**L**L**F**L**Y**K**F**R**L**S**N**S**Y**G**T**N**M**N**I**K**T**T-----
Cyanidioschyzon -----S**G**A**Q**Q**P**L**F**A**S**A**C**PGW**T**F**Y**E**K**T**Q**P**H**L**V**-S**C**L**A**T**A**K**S**P**Q**A**M**M**Q**L**L**V**R**H**E**-----
Trichomonas -----S**P**F**M**I**S**R**C**A**G**S**V**V**Y**E**R**K**T**S**Y**A--D**H**L**A**Q**I**K**P**Y**P**Q**L**Y**A**M**Y**E**K**K**I**-----
Trypanosoma -----R**L**P**L**I**V**S**A****C**PGW**V****C****C**E**K**Q**G**A**A**I**L**-P**L**L**C**P**V**M**S**P**Q**I**A**G**C**Y**S**K**T**L-----
H
Homo -A**Q**Q**Q**--H**L**T**P**D**K**I**Y**H**V**T**V**M**P****C**Y**D**K**K**L**E**A**S**R**P**D**F**F**N**Q**E**-----H**Q**T**R**D**V**D**C**V**L**T**T**G**E**V**F**R**L**L**E**E**E**G--V**S**L**P**D-----
Saccharomyces -F**E**S**L**--A**I**A**R**E**S**F**Y**H**L**S**L**M**P****C**F**D**K**K**L**E**A**S**R**P**E**S**L**D**-----D**G**I**D****C**V**I**T**P**R**E**I**V**T**M**L**Q**E**L**N**L**D**F**K**S**F-----
Blastocystis S1 -P**T**L**T**N**E**P**I**Q**S**D**I**Y**I**A**S**V**S****C**Y**D**R**K**L**E**A**S**R**R**D**F**M**D**P-----A**G**I**H**E**I**D**C**V**L**A**T**Q**E**I**A**E**L**L**E**K**P**A**P**S**A**P**S**A-----
Blastocystis S7 -P**A**H**V**V**P**A**L**A**N**G**D**I**F**I**A****C**V**A**S**C**Y**D**R**K**L**E**A**S**R**R**D**F**L**D**A-----D**G**T**R**H**V**N**C**V**L**S**S**Q**E**L**F**D**L**I**Q**N**P**P**D**F**R**H**A**S-----
Phaeodactylum -D**K**H**T**--S**L**P**R**Q**E**Y**Y**H**V**A**I**M**P****C**H**D**K**K**L**E**A**S**R**K**D**F**E-----D-----E**S**G**D**K**D**V**D**I**V**I**T**T**Q**E**C**M**R**L**I**Q**E**L**L**D**V**S**I**D**D**I**V**K**C**F**R**E**L**P**L**A**T**L**S**D**C**
Phytophthora -S**G**E**H**--G**V**K**P**S**E**V**Y**H**V**A**V**M**P****C**F**D**K**K**L**E**A**S**R**K**D**F**Q**D**A**E**-----D**A**T**K**D**V**D**C**V**L**A**T**E**I**E**L**I**E**S**L**N**V**D**F**A**S**L-----E
Arabidopsis -**C**Q**A**L--G**L**R**L**H**E**V**Y**H**V**T**V**M**P****C**Y**D**K**K**L**E**A**A**R**D**D**F**V**D**D**G**T**Q**D**N**-----G**D**--L**K**L**T**E**V**D**S**V**L**T**T**G**E**I**M**D**L**I**K**L**G**-V**D**F**K**D-----
Micromonas -A**A**E**L**--G**V**P**A**S**A**V**F**H**A**T**V**M**P****C**F**D**K**K**L**E**A**S**R**D**F**A**M**D**D**L**G-----E**N**V**R**E**V**D**C**V**L**T**T**G**E**V**A**E**M**I**A**V**A**G**L**G**A**P**S**-----S
Dictyostelium -S**K**K**I**--N**T**L**P**S**N**I**Y**H**V**T**I**M**P****C**Y**D**K**K**L**E**A**S**R**S**D**F**Y**N**D**V**-----F**K**T**K**D**V**D**C**V**L**S**T**S**E**V**L**L**L**K**E**H**G**D**V**D**L**L**K**-----
Tetrahymena -A**S**K**M**--G**I**E**S**K**D**I**L**F**V**S**V**M**P****C**Y**D**K**K**V**E**S**A**R**K**E**F**E**R**-----N**G**I**K**D**V**D**V**L**T**A**Q**E**I**M**D**L**L**K**K**V**E**G**Q**K**L**V**D**-----
Naegleria -S**N**E**I**--S**L**E**N**L**K**V**Y**H**T**T**V**M**P****C**F**D**K**K**L**E**A**S**R**P**D**F**T**N**-----D**P**F**D**K**V**D**M**V**L**T**S**S**E**I**T**E**L**L**Q**K**E**L**Q**I**E**T**P**E-----D**F**I**R**N**T**E**N**F**I**
Encephalitozoon -----R**T**V**S**V**M**P**C**Q**D**K**K**L**E**N**G**R**D**G**V**K-----F**D**F**I**L**T**T**R**G**F****C**K**A**L**D**S**L**G**F**R**R**P**A**R-----D**F**I**R**N**T**E**N**F**I**
Drosophila -A**D**K**I**--N**V**P**A**S**R**I**Y**H**V**T**V**M**P****C**Y**D**K**K**L**E**A**S**R**E**D**F**F**S**K**A**-----N**N**S**R**D**V**D**C**L**R**W**N**N**C**S**V**R**L**-----
Giardia -----D**A**A**P**L**L**V**S**I**Q**P**C**Q**D**R**K**L**E**Q**F**R**G**-----A**A**V**D**V**C**L**T**A**Q**E**V**H**G**F**L**A**E**T**P**Q**G**P**A**P-----A**F**
Entamoeba -----N**S**E**I**Y**H**V**S**I**Q**M**C**F**D**K**K**L**E**A**T**K**T**Y-----N**N**I**H**V**I**D**C**V**L**T**T**S**E**I**D**S**I**I**D**W**N**E**P**I**N**E**I**T-----
Cryptosporidium -F**T**Q**N**D**D**F**V**E**Q**S**D**I**F**H**V**A**I**M**P****C**H**D**K**K**L**E**S**T**R**S**S**L**S**L**K**S**-----S**D**K**N**S**S****C**P**E**V**D**I**V**L**A**T**S**E**V**G**E**I**I**K**L**A**G**F**N**S**L**L**D**-----
Cyanidioschyzon -----L**G**E**H**A**W**T**V**S**I**A**P****C**Y**D**K**K**L**E**S**Q**R**D**A**K**D-----S**E**H**V**F**V**L**T**A**T**E**V**L**E**L**A**E**R**A**P**H**M**P**N**D-----
Trichomonas -----L**Q**S**T**N**Y**V**L**Y**I**G**P****C**Y**D**R**K**L**E**A**A**R**F**E-----E**D**V**D**A**V**L**T**I**A**E**I**N**D**H**I**T**E**P**T**E**E**I**P**V**K**-----
Trypanosoma -----I**P**Q**M****C**H**V**S**V**Q**P****C**F**D**R**K**L**E**A**A**R**D**G**S**S**V**-----S**G**E**R**Y**T**D**F**V**L**S**T**Q**E**L**L**D**W**M**L**E**V**D**P**S**L**P**W**Q-----
*** * ***
Homo -L**E**P**A**P**L**D**S**L**C**-----S**G**A**S**A**E**-----E**P**T**S**H**R**G**G**-----G**S**G**G**Y**L**E**H**V**F**-----R**H**A**A**R**E**L**F**G**I**H**V**-----
Saccharomyces -----L**T**E**D**T**S**L-----Y**G**R**L**S**P**P**G**W**D**-----P**R**V**H**W**A**S**N**L**G**G-----T**C**G**G**Y**A**Y**Q**Y**V**-----T**A**V**Q**R**L**H**P**-----
Blastocystis S1 -----P**V**R**V**W**W**D**E**K**R**W**E**A**M**E**S**G-----C**R**T**A**F**D**D**V**L**M**-----S**S**G**G**V**L**Q**S**L**L**K**F**E**L**L**S**R**P**A**A**R**L**V**F**A**L**F**V**E-----
Blastocystis S7 -----S**V**V**W**N**E**A**W**S**A**L**E**A**P**-----L**L**T**D**F**D**A**A**L**F**-----S**S**G**G**V**L**E**G**V**V**A**S**V**V**A**S**R**P**S**A**N**V**E**W**V-----
Phaeodactylum T**S**F**T**K**A**E**P**V**L**-----I**A**D**S**N**S**H**C**I**T**T**L**T**T**E**D**-----A**E**I**S**S**N**A**A**F**T**L-----G**S**G**G**Y**A**S**F**I**F**-----A**Y**A**A**K**R**L**F**G**V**Q**L**D**A**H**E**L**P**
Phytophthora L**A**T**L**T**P**E**E**V**M**L-----S**G**V**S**E**D**-----G**S**S**V**L**G**S**S**Q**N**A-----S**S**G**H**L**E**H**I**F-----R**Y**A**A**K**E**L**F**N**V**D**V**-----
Arabidopsis -L**E**E**S**P**L**D**R**V**L**-----T**N**V**T**-----E**E**G**D**L**Y**G**V**A**G**-----S**S**G**G**Y**A**E**T**I**F**-----R**H**A**A**K**A**L**F**G**Q**T**I**-----
Micromonas A**L**A**F**N**P**R**P**R**C**L-----S**T**P**P**-----D**A**F**Q**L**H**P**V**V**R**S-----G**S**G**G**Y**L**D**A**V**F**-----R**H**A**A**K**V**L**H**G**V**D**V**-----
Dictyostelium -L**E**E**A**T**L**D**N**S**I**-----F**N**N**V**I**F**N**Q**Q**T**G-----Q**P**E**K**F**L**G**V**T**G**-----S**T**G**G**Y**F**E**Y**L**F**-----R**R**A**A**K**E**L**F**G**K**E**I**-----
Tetrahymena -V**Q**E**Q**I**K**Q**A**T**N**V**N**S**N**D**E**S**K**N**E**L**I**S**K**E**R**E**Y**I**E**N**N**I**F**Y**E**N**L**D**I**S**R**I**V**S**N**I**E**Y**L**F**G**D**S**Q**T**F**D**Q**L**L**P**I**L**S**N**I**F**D**S**T**I**-----G**S**N**D**Y**L**D**Y**I**I**-----R**R**A**A**S**D**I**H**L**N**P**D**-----
Naegleria K**N**E**Y**Q**L**D**S**I**F**-----E**I**L**S**Q**S**T**H**H**S**T**Q**D**I**S**E**D-----A**T**L**L**E**W**L**G**S**E**S**D**A**T**G**S**G**G**Y**C**E**I**V**F**-----K**Y**A**A**K**L**F**G**I**D**L**R**-----
Encephalitozoon -----A**S**G**K**S**L**C**S**M**E**E**A**-----E**T**T**Q**W**N**I**G**T-----S**S**G**G**Y**A**E**F**I**L**-----

Drosophila -----
Giardia CSSYTPSPTSF---WQYALGPLLVLYLKAERWISDESL-----SHLLV-----PHAGGIDLRW-----AKVGNELFSCTIEL-----
Entamoeba -----STRMKGFIS-----SPAQYIALME-----
Cryptosporidium -VPEAPLDNLWLNQNFQITKKHNSLLITENYVSNQI-----LNQFPWLIPSYFNNSGGFCHEYII-----RSAIKELAGDNID-----
Cyanidioschyzon -----ARRQLRSQQVS-----ALTKFRGAWWATFGGLSGGYAVEVF-----RYVAQYVFKHRVS-----
Trichomonas ----FPADTDL-----NAISQKLGQIKDS-----LNSDSIYQLI-----AEIEPTLNEEEEIN-----
Trypanosoma ----APLSDSL-----EPLPLPPEE-----PKRSFAATME-----GSGGYHRYAM-----HRAARELHGLELA-----

Homo -----AEVTYKP--LRNKDFQEVTL-----EKEGOVLLHFAMAYGFRNIQNLVQRLKR-----GRC-----PYH
Saccharomyces -----GSQMIVLE--GRNSDIVEYRL-----LHDDRIIAAASELSGFRNIQNLVVRKLTSGSGSERKRNITALRKRRTGPKANSREMAAATAATADPYHSD
Blastocystis S1 -----NSMKRTRT--VRNSDFVESSV-----CVGEEIVFRGAFIYGFRNIQNLAMKIKR-----GKC-----VYD
Blastocystis S7 -----RTRT--VRNSDFVEVTV-----VDGEERVFSGAFVYGFNRNIQNLVMKTKV-----GKC-----AYD
Phaeodactylum WEPVGPDQAGRVSARVAAST--QRRRDYHYVALYRSQDGNFTT---NANLSSDSKPIHLFAIAYGMQTLQRVLKPYTSEHLQ---SGI-----GYD
Phytophthora -----NGPL-----DRS-----
Arabidopsis -----EGPLEFKT--LRNSDFREVTL-----QLEGKTVLKFALCYGFQNLQNI VRRVKT-----RKC-----DYQ
Micromonas -----TGPLKYATPSSRNLDLKEVTL-----EVDGAVVLRFAAAYGFRNIQNI VRKCKAGSVA-----TGD-----AYD
Dictyostelium -----EGEIEYKV--GRNTDFKEASL-----EVDGKVLVSFAKAYGFRNIQNI VRKIKTTASSK-----KEP-----QYH
Tetrahymena -----QYEIITKQ--GRNSDFNEIFL-----VKDGANILSFARVYGLRNIQNI IRNLRQ-----NKC-----KYD
Naegleria -----DKTLIFES--KRNSDYRETVL-----VDPTDSSKILLRFVIANGFRNIQNLVRRMKQAD-----NSE-----PFH
Encephalitozoon -----GKHCVVETREIRN-GIKEHLL-----DDGRTI---SQTGLENSINYFKSSKT-----KGP-----RHK
Drosophila -----SVRYRSTISLI-----
Giardia -----SQNQSPYS CVV-----IYKSAGYHNLQNLVRRVHALCP-----NKD-----ALY
Entamoeba -----QKKEPFKV--TRNKDFLENDG-----IAIANGFRNIQNVVRFVK-----SKT-----KLQ
Cryptosporidium -----NKVQLPFN--KLKNDILEAKY-----IKNNVELNYCLVYGFRAIQSISRKLNQKNASQHTQYKQNVVNHV-----NYH
Plasmodium -----
Cyanidioschyzon -----DAVLEPAAVRERNPDLRQLLLYQHRSSGDFVFNFRPVIADDDLQLRYSVATAYGFRNIQNI VRQYKHT-----GRC-----NWN
Trichomonas -----SLISEL--PSRFDLEISTN-----SFDGETLNKRLT---KTLDMSSGKKV-----PKP-----APR
Trypanosoma -----PRDIHYEM--KRNANHHLTT-----PSNPGEV--FCVAYGFQQIQNTVRGIKRKL-----SVA-----SYT

H-cluster

Homo YVEVMA **CP**SGCLNNGGQLQAPDRPS-----RELLQHVERLYGMVRAEAPEDAPGVQ-----
Saccharomyces YIEVNA **CP**GACMNGGGLLNGEQNSL-----KRKQLVQTLNKRHGEELAMVDPLTLGPKLEEAAARPL-----SLEYVFPVVKQAVEKDLVSVG---
Blastocystis S1 AVEVMA **CP**SGCLNNGGQIRPEKRENMMKIAEELGVAVTKWKRRI G-----MFERDLRSRLNDYWRVESK **C**MMYW
Blastocystis S7 VVEVMA **CP**SGCLNNGGQIKPEKREDMML-----
Phaeodactylum YVEAMA **CP**SGCLVNGGGQIRTSARETP---TETRFVGTQTTLRVPQMNESSGRTQLGAGS-----SLHTRYHIVPPLQHS LGAA---
Phytophthora -----
Arabidopsis YVEIMA **CP**AGCLNNGGQIKPKTGQS-----QKELIHSLEATYMNDDTTLNTDPYQNPT-----AKRLFEEWLKEPGSNEAKK-----YLHTQYHPVVKSVTSQL-----
Micromonas FVEIMA **CP**SGCLNNGGQLPP-----
Dictyostelium FVEVMA **CP**SGCLNNGGQIKAAASGSL-----REQKLLIEQSEQKYQDLSIQSYEQLPST-----LEIKDIYDQWMNGPFS LDAKT-----KLHTQYHMIEKSNALS-----
Tetrahymena YIEIMA **CP**SGCLNNGGQLKSLDENM-----KTKDLIEKIQEILRKKSVISNDQYNNLAFQQIIRDSTNI-----YFETIFKHIEKLNTOSS-----
Naegleria FVEVMA **CP**IGCLNNGGQVVKQPKITLDIDSPSSVTITPKQH LAQTEKLYRDDISNIKSL **C**ENNSNLVKEGLNNLYNWIKGSVGSVQAKS-----ILHTQYHAVEKLNIVTPTA-----
Encephalitozoon MTEIFL **CK**NGCIGGPGQERVNDVEMDIREYDRNGREQPRIFYSSPGL-----EKRVFREVAKRVDLR-----
Drosophila -----
Giardia ILDLHA **CP**YGCY--GG **C**IAGDDRHPVSSVASASHAAVVSADKSVLSHILAAGT **CP**GLLEELVQAVDGITPQET-----VIRTDGEVVSPEEIAARKGIQGGI
Entamoeba FIEVEA **CP**GGCICGGGQIK **C**SP-----KEKDERVKKMMEILEPKVVDEKNKSIYESIKD-----SIKLTFDIKESAQENALH-----
Cryptosporidium LIEAMA **CPT**GCVSGGGQIILSQNDQNDNSDLNKRKNIKFIDEVQEAALYKGINLNKNQEIILPDE---IPIVNILYEYLIHIDKQIDRSSGLKLPFLRNDFVSI NEVPTASS-----
Cyanidioschyzon FVEVMA **CP**GGCNGNGLWVVDPTSTPVSASVAVANRVFAQRTRQERI QEMERSLQRLP **C**YHVDELPLQRELWATLREAEFQ-----SLRFEF **C**GFNDRQQQEATYRQQDG
Trichomonas LAQIDF **CK**GGCLVGGGQIRGNSP-----AQRALIAATQEVHTQNESTNISFPTELYNELIKF-----GYKTHYESLPQEEEKDQ-----
Trypanosoma FIELMA **CPE**GC L NNGGQARNGTTQSHVETTAAAKSAFISYISGSQPTMEVRRGGMDGTSVANSNRVDVGDGPPFPFLSAAFAEVEKQVGS S LWS **C**TFT--DRQREFEATLNTGGVHS-----

Homo -----
Saccharomyces -----STW
Blastocystis S1 VLREDRRKHMRCFTLDSM
Blastocystis S7 -----
Phaeodactylum -----AGVPVKDTQW
Phytophthora -----
Arabidopsis -----NNW

<i>Micromonas</i>	-----
<i>Dictyostelium</i>	-----IKW
<i>Tetrahymena</i>	-----LNW
<i>Naegleria</i>	-----LKW
<i>Encephalitozoon</i>	-----VDW
<i>Drosophila</i>	-----
<i>Giardia</i>	RIQD-----LAW
<i>Entamoeba</i>	-----LNW
<i>Cryptosporidium</i>	-----LKW
<i>Cyanidioschyzon</i>	DIKSQVTVSERHRLVHQW
<i>Trichomonas</i>	-----FAW
<i>Trypanosoma</i>	-----LKW

Figure S4

Homo sapiens-nbp35
Saccharomyces cerevisiae-nbp35
Dictyostelium purpureum-nbp35
Entamoeba dispar-nbp35
Naegleria gruberi-nbp35
Guillardia theta-nbp35
Encephalitozoon cuniculi-nbp35
Trypanosoma vivax
Giardia lamblia-nbp35
Blastocystis-nbp35
Blastocystis sp. NandII-nbp35
Phytophthora infestans-nbp35
Phaeodactylum tricornutum-nbp35
Arabidopsis thaliana-nbp35
Micromonas sp.-nbp35
Cyanidioschyzon merolae-nbp35
Cryptosporidium hominis-nbp35
Tetrahymena thermophila-nbp35
Metallospira cuprina
Pyrococcus furiosus
Methanococcus maripaludis
Homo sapiens-cfd1
Saccharomyces cerevisiae-cfd1
Dictyostelium purpureum-cfd1
Entamoeba dispar-cfd1
Naegleria gruberi-cfd1
Guillardia theta-cfd1
Encephalitozoon cuniculi-cfd1
Trypanosoma vivax-cfd1

Homo sapiens-nbp35
Saccharomyces cerevisiae-nbp35
Dictyostelium purpureum-nbp35
Entamoeba dispar-nbp35
Naegleria gruberi-nbp35
Guillardia theta-nbp35
Encephalitozoon cuniculi-nbp35
Trypanosoma vivax
Giardia lamblia-nbp35
Blastocystis-nbp35
Blastocystis sp. NandII-nbp35
Phytophthora infestans-nbp35
Phaeodactylum tricornutum-nbp35
Arabidopsis thaliana-nbp35
Micromonas sp.-nbp35
Cyanidioschyzon merolae-nbp35
Cryptosporidium hominis-nbp35
Tetrahymena thermophila-nbp35
Metallospira cuprina
Pyrococcus furiosus
Methanococcus maripaludis
Homo sapiens-cfd1
Saccharomyces cerevisiae-cfd1
Dictyostelium purpureum-cfd1
Entamoeba dispar-cfd1
Naegleria gruberi-cfd1
Guillardia theta-cfd1
Encephalitozoon cuniculi-cfd1
Trypanosoma vivax-cfd1

MEKERIVPEGAPDSCPGVKSQLAGKVEPCSGCPHQPNCALGNPNLDGRIIP EEAPSQCPTQSNLAGHTVVCSGCCPHQPNCSKALNVHYKLVDPNAPTCCPGTLSQDAGKVNSCSGCPN

---MKIASNRRIAFFGGLISGIIFKDYFKTF
QQKCNSSGQSNITIQEHKKLVDPNAPQNC PGVDSQNAKTEVCSVCSPHQPNCALGKLVGHKVVVPVDAPEQCPTQAQNAKTPVCSGCCPHQPNCSAAGKQANI TADKKIVPVNANTGCVGTQ

	N-terminal	CX13CX2CX5C motif	Walker A
<i>Homo sapiens-nbp35</i>	-----MEEVPHD	CG	ADSAQAGRGASCGGCPNQR
<i>Saccharomyces cerevisiae-nbp35</i>	-----MTEILPHVNDVLP	PG	PESDMAGKSDACGGCANKEI
<i>Dictyostelium purpureum-nbp35</i>	-----MSDQLVAPPPEHC	PG	TQSEMSGKSAACAGCPNQI
<i>Entamoeba dispar-nbp35</i>	-----MSCDKNCATCP	DS	BQAGS SGC THNCATC SHKGS
<i>Naegleria gruberi-nbp35</i>	-----MENKPSSEAVPENANEHC	PG	VESQNAGKSSSCAGCPNQSL
<i>Guillardia theta-nbp35</i>	-----	CG	MAGKSSTCVGC
<i>Encephalitozoon cuniculi-nbp35</i>	-----MGESC	PG	VSSKDAGKAECKGCPNVGYC
<i>Trypanosoma vivax</i>	-----MEGANSTC	VG	PESEAGVAPSCQGC
<i>Giardia lamblia-nbp35</i>	-----MTADHKNCRSRA	VG	PVSAASNCQPDKAGC
<i>Blastocystis-nbp35</i>	-----MSSVPNANPSC	VG	PQSEQAGRAAGCQGC
<i>Blastocystis sp. NandII-nbp35</i>	-----MSSVPNANPSC	VG	PQSEQAGRAAGCQGC
<i>Phytophthora infestans-nbp35</i>	-----MATSAPANANAEC	VG	PQSEQAGKAGCQGC
<i>Phaeodactylum tricornutum-nbp35</i>	-----MTEASENAGDAPANANTC	VG	PTSETAGKASACDGC
<i>Arabidopsis thaliana-nbp35</i>	-----MENGDIPEANNEHC	PG	PQSESAGKSDSCAGC
<i>Micromonas sp. -nbp35</i>	-----MDQVTIPEDANTNC	PG	TNADAAGKVPACTGCPNQGA
<i>Cyanidioschyzon merolae-nbp35</i>	-----MTETSVPVSC	PG	TESEQAGRAAACGCPNQTA
<i>Cryptosporidium hominis-nbp35</i>	SRLTINYLKSIFISYIKTDSVISKYQNHFNQIDNC	VG	VDS PDAG IADSCAGC
<i>Tetrahymena thermophila-nbp35</i>	SQQAGTASSCAGCPNKAKCSSQNTTEENKVVDPNANEHC	PG	TQNKEAGKMSACGCPNQIC
<i>Metallospira cuprina</i>	-----MSSNPFRLQSPQAKQPRDLRKAPAFQ	CG	VQGADLKIQRMRNKVYKVA
<i>Pyrococcus furiosus</i>	-----MTIKAPTLNVPGLGVDPLTQRIKEKEKWKYKIAVL	CG	SGNCDSCGSSSDCSDTKK
<i>Methanococcus maripaludis</i>	-----MAEEC	CG	MMEQNAQIRDNMSKIKYIAVMS
<i>Homo sapiens-cfd1</i>	-----	CG	MEAAAEPGNLAGVRHII
<i>Saccharomyces cerevisiae-cfd1</i>	-----	CG	MEEQIGVPAASLAGIKHII
<i>Dictyostelium purpureum-cfd1</i>	-----	CG	MDKIKHII
<i>Entamoeba dispar-cfd1</i>	-----	CG	MTELNADRNFVGDHVKNV
<i>Naegleria gruberi-cfd1</i>	-----	CG	MQEQLFSEKLSHIKNI
<i>Guillardia theta-cfd1</i>	-----	CG	MLQEGTKAVVLS
<i>Encephalitozoon cuniculi-cfd1</i>	-----	CG	MSRIAVMS
<i>Trypanosoma vivax-cfd1</i>	-----	CG	MTSALHGVKHI

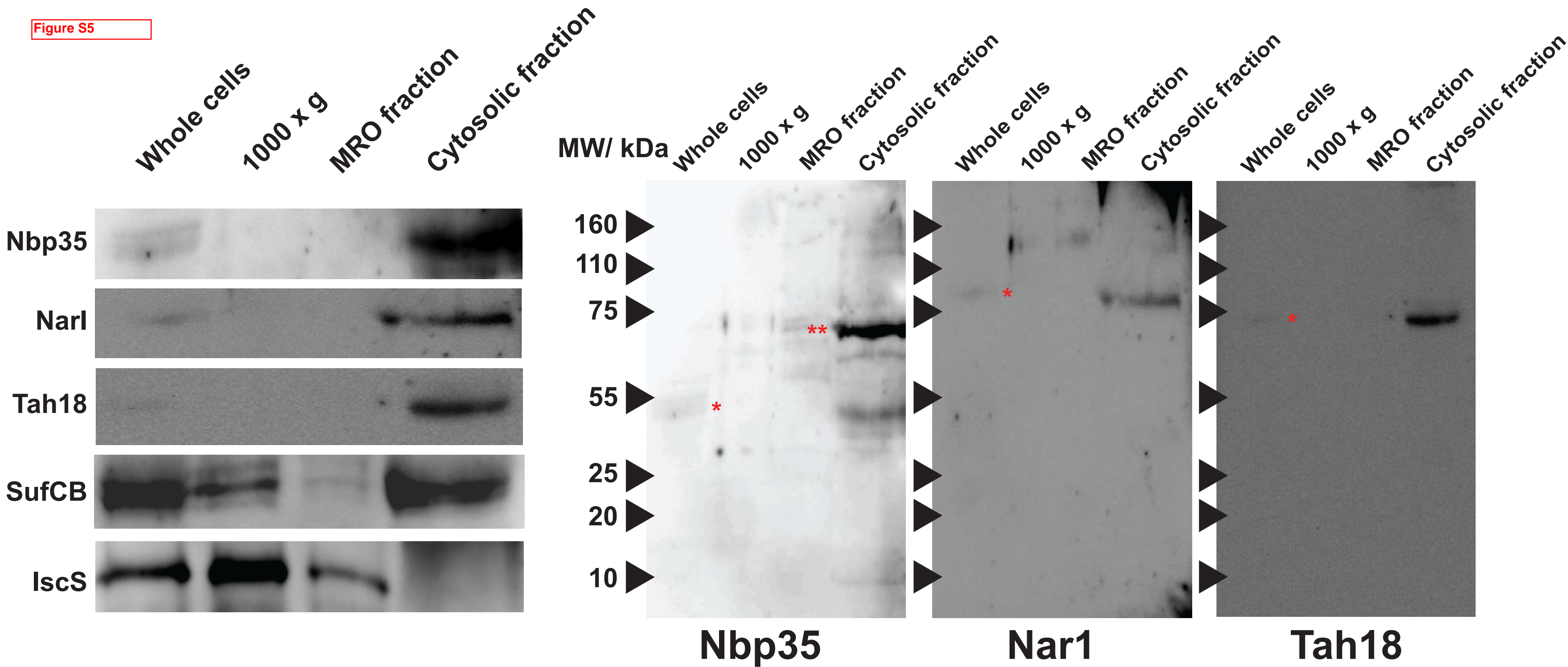
<i>Homo sapiens-nbp35</i>	SAHLAHGLA	CG	EDENTQIALLDIDICG	SP	GLEGEVHQSG	SG	WSPVY	VE	DNLGVM
<i>Saccharomyces cerevisiae-nbp35</i>	AAMLSWALS	CG	ADELQVGAMDLDICG	SP	GKETVHESN	SG	WTPVY	VT	DNLATM
<i>Dictyostelium purpureum-nbp35</i>	SSQLSFALA	CG	MNTEEQVGLLDIDICG	SP	GLEGEVHISG	SG	WDPVY	VE	DNLAVM
<i>Entamoeba dispar-nbp35</i>	TTQFGWVLS	CG	EDKQVGI	CD	LDDICG	SP	IQMF	VT	ENLCTM
<i>Naegleria gruberi-nbp35</i>	SSQLALTMA	CG	LCEKDSVPQVGLD	V	DICG	SP	ITMF	YE	DNLAVV
<i>Guillardia theta-nbp35</i>	ASQVAMALS	CG	GSGNVTGLLDIDV	CG	SP	APRML	GL	VD	ENL
<i>Encephalitozoon cuniculi-nbp35</i>	TRNIAELMS	CG	SRGIAT	C	LDDLDL	SG	SP	VT	GTDG
<i>Trypanosoma vivax</i>	TKELAFALG	CG	RHGLEVA	V	VDL	V	CG	ID	GVR
<i>Giardia lamblia-nbp35</i>	STQLGFYLA	CG	ENMEKNVGLMD	V	DICG	SP	ITMT	VL	FNMA
<i>Blastocystis-nbp35</i>	SSQIAFSLA	CG	SQGFQVGLD	I	DICG	SP	PRMM	VD	DNL
<i>Blastocystis sp. NandII-nbp35</i>	SSQIAFSLA	CG	SQGFQVGLD	I	DICG	SP	PRMM	VD	DNL
<i>Phytophthora infestans-nbp35</i>	ACQLAFALA	CG	GKGFQVGLD	V	DITG	SP	PRML	VD	DNL
<i>Phaeodactylum tricornutum-nbp35</i>	AAQLSHTLS	CG	NQGYAVGLD	V	DICG	SP	APRML	AS	ANL
<i>Arabidopsis thaliana-nbp35</i>	SAQLSFALA	CG	GMDHQVGLMD	I	DICG	SP	PKML	VE	DNL
<i>Micromonas sp. -nbp35</i>	AAQLAYALS	CG	SRNYRVGLLD	V	DICG	SP	PLLF	VT	DNL
<i>Cyanidioschyzon merolae-nbp35</i>	AAQLAFAFASGVA	CG	EAVDRELEVGLL	D	ADVTG	SP	PHLL	VR	GNL
<i>Cryptosporidium hominis-nbp35</i>	SSQISWCLS	CG	SKFFNVGLLD	I	DICG	SP	APKMM	VN	DNL
<i>Tetrahymena thermophila-nbp35</i>	SSQLAFYLA	CG	NLGYEVGLLD	I	DICG	SP	PRML	VE	DNL
<i>Metallospira cuprina</i>	SSNLAMALA	CG	AAGKSVGI	I	DVDFH	GP	SVKML	GP	FGI
<i>Pyrococcus furiosus</i>	AVNLTAALA	CG	KMGYFVGI	L	DADIH	GP	NVAKMF	VD	FMG
<i>Methanococcus maripaludis</i>	TVNLAATLN	CG	MMGYKVG	V	LDGD	I	HGPNIPQML	TP	QGK
<i>Homo sapiens-cfd1</i>	STELALALR	CG	HAGKKG	V	GLD	V	DICG	LD	REQ
<i>Saccharomyces cerevisiae-cfd1</i>	TTQTALTL	CG	SMGFK	V	GLD	I	DITG	VE	TNST
<i>Dictyostelium purpureum-cfd1</i>	SSQLAFYLA	CG	HTGNK	V	GLLD	V	DICG	TD	ETQ
<i>Entamoeba dispar-cfd1</i>	ATALARSFA	CG	LVGKKT	G	I	DICG	SP	SQ	IGD
<i>Naegleria gruberi-cfd1</i>	SCQLALTLA	CG	NMKYK	V	GLD	V	DICG	VE	DEL
<i>Guillardia theta-cfd1</i>	ATGIAISLS	CG	EEGK	V	GLD	V	DICG	TD	GEN
<i>Encephalitozoon cuniculi-cfd1</i>	SIMLSTVLS	CG	EKG	R	L	L	D	VS	KNL
<i>Trypanosoma vivax-cfd1</i>	ACQLALALA	CG	FKHGK	V	GLLD	V	DICG	CG	GLT

	Mrp family Signature			
<i>Homo sapiens-nbp35</i>	QFLRDVDN	C	V	VDYLVVDTPPGTSEHLSVVRVYLAH-I-----DGAVIITTPQEVSLQDVRKEINFCRKVKLP IIGVVENMS--GF-----
<i>Saccharomyces cerevisiae-nbp35</i>	KFLKDVDN	D	L	LDYLVVDTPPGTSEHIS INKYMRESG-I-----DGALVVTT PQEVALDVRKEIDFCRKAGINILGLVENMS--GF-----
<i>Dictyostelium purpureum-nbp35</i>	QFLKDVMN	N	L	LDYLVVDTPPGTSEHLS IVQYKTSN-I-----DGAVIITSPQDVALIDVRKEINFCRKVGVP IIGVVENMS--GF-----
<i>Entamoeba dispar-nbp35</i>	QFLHDVEM	C	L	LDYLVVDTPPGTSEHLS IVSILTKSN-V-----DGAIIVTTPQDVSLIDVRKEINFCKKIGLPI IIGVVENMS--GF-----
<i>Naegleria gruberi-nbp35</i>	QFLRDVMN	C	L	LDYLVVDTPPGTSEHIS IVQYKLVND-I-----DGAIIVTTPQDVSCNDVRRINFCRKVGVP IIGI IENMS--GF-----
<i>Guillardia theta-nbp35</i>	QFLKDVNM	C	L	LDYLVVDAPPGTSEHITIAQCLNQTS-N-----VGAVIVTTPQEVALDVRKEINFCRKAGVKILGV IENMA--GF-----
<i>Encephalitozoon cuniculi-nbp35</i>	KLLKWCXY	E	T	TDVLLLDTPPNVTDEHLMGNVFIRP-----RFGIVVTT PQKFSLQDVARQVDFCRKARIEVLGI IENMK--RF-----
<i>Trypanosoma vivax</i>	MFFKDVMN	C	I	IDIMLIDTPPGTSEHITASSLLQCCGGV-----TGAIIVTTPQLVAEADVRREVNFCQKAKIP IIGI IENMS--SF-----
<i>Giardia lamblia-nbp35</i>	NFLKDVHM	F	D	SEKI DNYLII DTPPGTSEHLSVINMLSAAMRVLNKEKETDPSVHTPTFFAVVSTPQEVALADVRKEINFCQKIKVDVKGV IENMS--GF-----
<i>Blastocystis-nbp35</i>	QFLSEINM	C	L	LDYLVVDTPPGTSEHIS IVNFLRDVG-I-----DGAVIVTTPQEVALSDVRKEIRFCQKQSGIRI IIGI IENMSEVNF-----
<i>Blastocystis sp. NandII-nbp35</i>	QFLSEINM	C	L	LDYLVVDTPPGTSEHIS IVNFLRDVG-I-----DGAVIVTTPQEVALSDVRKEIRFCQKQSGIRI IIGI IENMSEVNF-----
<i>Phytophthora infestans-nbp35</i>	QFLVDVQN	C	L	LDYLVVDTPPGTSEHIS IVQYMKDAD-I-----DGAVVVTT PQEVALSDVRKELNFCRKTNINVLGVVENMS--GV-----
<i>Phaeodactylum tricornutum-nbp35</i>	QFLTEVDWTGDT	C	X	LDYLVVDTPPGTSEHIS IVQYKASAV-----SGAVVVTT PEEVSLADVRKELSFCKRTDVPVGLGI IENMG--SYQTRLRQMEFSKD
<i>Arabidopsis thaliana-nbp35</i>	QFLKDVMN	C	L	LDYLVVDAPPGTSEHIS IVQYLLPTG-I-----DGAIIVTTPQEVSLIDVRKEVFCRKVGVPVGLGVVENMS--GL-----
<i>Micromonas sp.-nbp35</i>	QFLKDTVM	C	L	LDYLVVDAPPGTSEHLSVQVHMKLAG-I-----DGAIIVTTPQEMALADVRKEINFCRKVGINILGVVENMS--GL-----
<i>Cyanidioschyzon merolae-nbp35</i>	QFLRDVDN	C	L	LDYLVVDTPPGTSEHITLTQALAGIA-N-----TAALIVATPQEVALLDVRKQVRFCEKAGVPIVGV IENMS--YY-----
<i>Cryptosporidium hominis-nbp35</i>	QFLSDVMN	C	L	LDYLVVDTPPGTSEHLS IVSVLNLSN-V-----NGALIVTTPQEVSLQDVRKEINFCRKVGINILGVVENMG-----
<i>Tetrahymena thermophila-nbp35</i>	QFLTDVMN	C	L	LDYLVVDTPPGTSEHIS CVQYLPQGE-G-----DGAVIVTTPQEVSLQDVRKELSFCKQKTKTNILGVVENMS--GF-----
<i>Metallosphaera cuprina</i>	QFLGDVMN	C	L	LDYLVVDTPPGTSEHMAISVAQVLPN--I-----TGFIIVTTPQEVSLADVRKELSFCKRTDVPVGLGI IENMG--YF-----
<i>Pyrococcus furiosus</i>	QLLGDVMN	C	L	LDYLVVDTPPGTSEHIS IVQYKASAV-----DAAVIVTTPQEVALLDVGKAVNMMKKMEVPI IAV IENMS--YL-----
<i>Methanococcus maripaludis</i>	QFLSDVMN	C	L	LDYLVVDTPPGTSEHIS IVQYKASAV-----DGAVIVTTPQEVSLADVRKELSFCKRTDVPVGLGI IENMG--YF-----
<i>Homo sapiens-cfd1</i>	QFVSDVAN	C	L	LDYLVVDTPPGTSEHMAISVAQVLPN--I-----LQALIVATPQEVALLDVRKQVRFCEKAGVPIVGV IENMS--YY-----
<i>Saccharomyces cerevisiae-cfd1</i>	QFISDVAN	C	L	LDYLVVDTPPGTSEHIS IVQYKASAV-----DGAIIVTTPQEVSLADVRKELSFCKRTDVPVGLGI IENMG--YF-----
<i>Dictyostelium purpureum-cfd1</i>	QFIDVVMN	C	L	LDYLVVDTPPGTSEHIS IVQYKASAV-----DGAIIVTTPQEVSLADVRKELSFCKRTDVPVGLGI IENMG--YF-----
<i>Entamoeba dispar-cfd1</i>	QFLNDVEM	C	L	LDYLVVDTPPGTSEHIS IVQYKASAV-----DGAIIVTTPQEVSLADVRKELSFCKRTDVPVGLGI IENMG--YF-----
<i>Naegleria gruberi-cfd1</i>	QFIQDVCM	K	L	LDYLVVDTPPGTSEHIS IVQYKASAV-----DGAIIVTTPQEVSLADVRKELSFCKRTDVPVGLGI IENMG--YF-----
<i>Guillardia theta-cfd1</i>	QFFTVMN	C	L	LDYLVVDTPPGTSEHIS IVQYKASAV-----DGAIIVTTPQEVSLADVRKELSFCKRTDVPVGLGI IENMG--YF-----
<i>Encephalitozoon cuniculi-cfd1</i>	MFYESID	D	L	LDYLVVDTPPGTSEHIS IVQYKASAV-----DGAIIVTTPQEVSLADVRKELSFCKRTDVPVGLGI IENMG--YF-----
<i>Trypanosoma vivax-cfd1</i>	QFISDVNM	C	L	LDYLVVDTPPGTSEHIS IVQYKASAV-----DGAIIVTTPQEVSLADVRKELSFCKRTDVPVGLGI IENMG--YF-----

	C-terminal CPXC motif			
<i>Homo sapiens-nbp35</i>	-----I	CP	KCKKE	-----SQIFPP--TTG-----GAEMLCQD
<i>Saccharomyces cerevisiae-nbp35</i>	-----V	CP	NCKGE	-----SQIFKA--TTG-----GGEALCKE
<i>Dictyostelium purpureum-nbp35</i>	-----V	CP	KCNKE	-----SQIFIP--TTG-----GAEQMSKD
<i>Entamoeba dispar-nbp35</i>	-----V	CP	CKHKE	-----STIFPP--TNG-----GAQKMCCE
<i>Naegleria gruberi-nbp35</i>	-----V	CP	NCKNK	-----AMIFKP--TSG-----GGQQLAID
<i>Guillardia theta-nbp35</i>	-----V	CP	HCGNS	-----SEIFVP--SEADSVQGPSAVLRGAEMAKK
<i>Encephalitozoon cuniculi-nbp35</i>	-----T	CP	KCGHS	-----KSIFF--RSV-----GVESYCMS
<i>Trypanosoma vivax</i>	-----V	CP	GCCKS	-----SVIFPSAGSCG-----AGERLSAE
<i>Giardia lamblia-nbp35</i>	-----V	CP	CCKNE	-----TQIFNP--SSG-----GVKQLCAD
<i>Blastocystis-nbp35</i>	-----E	MT	ECRYR	-----DYFGNDITDAVAKLNEAFPEFAHLIGSICIFPP--SNN-----GGEGLAQW
<i>Blastocystis sp. NandII-nbp35</i>	-----E	MT	ECRYR	-----DYFGNDITDAVAKLNEAFPEFAHLIGSICIFPP--SNN-----GGEGLAQW
<i>Phytophthora infestans-nbp35</i>	-----Q	RPLSDVKFVGA	DGNDE	-----TSAFMKLLQEKAPPELLKHSVQMEVFFPA--STG-----GGEAMAKK
<i>Phaeodactylum tricornutum-nbp35</i>	GQDCTAQMLAVLREK	CP	EVLDVAA	-----SNLFSV--NAG-----GAEQMATD
<i>Arabidopsis thaliana-nbp35</i>	-----S	QLPKDVKFMKLATETGSSINVTEDVIA	CLRNAP	-----ELLDIVACSEVFDS--SGG-----GAERMCRE
<i>Micromonas sp.-nbp35</i>	-----N	LP	NCAQVMFTSAENGSDITLKRDAIQKPNPSGFTASSANVHFEPSIRIQIDVFFPA--SRG-----GALMKCFE	
<i>Cyanidioschyzon merolae-nbp35</i>	-----T	CR	HCGHQ	-----EEIFTP--TTG-----GVQALCND
<i>Cryptosporidium hominis-nbp35</i>	-----I	CP	GCKCE	-----MIFKNAEHS-----SVKDMCDN
<i>Tetrahymena thermophila-nbp35</i>	-----V	CP	SESKN	-----SQIFPP--VTG-----GAARKCQD
<i>Metallosphaera cuprina</i>	-----V	CP	SESKN	-----YYIF--GQD-----KGGKMAEE
<i>Pyrococcus furiosus</i>	-----I	CP	HCGNK	-----IDIF--GEG-----GGEKLAEK
<i>Methanococcus maripaludis</i>	-----V	CP	ECDKV	-----IDIF--GKG-----GGEKAAKE
<i>Homo sapiens-cfd1</i>	-----T	CP	HCTEC	-----TSVF--SRG-----GGELAQL
<i>Saccharomyces cerevisiae-cfd1</i>	-----V	CP	HCAEC	-----TNIF--SSG-----GGKRLSEQ
<i>Dictyostelium purpureum-cfd1</i>	-----V	CP	HCEC	-----TNIF--SSE-----GGKLLAEQ
<i>Entamoeba dispar-cfd1</i>	-----L	CP	HCASTV	-----TNIF--SSN-----GGKELADK
<i>Naegleria gruberi-cfd1</i>	-----V	CP	CKKEI	-----TFIF--GSG-----GGQKLSNE
<i>Guillardia theta-cfd1</i>	-----Q	CP	CGEV	-----TDIF--SRG-----GGEKLAIE
<i>Encephalitozoon cuniculi-cfd1</i>	-----C	CP	CGSS	-----VNIFF--GSK-----GGERLAEE
<i>Trypanosoma vivax-cfd1</i>	-----V	CP	HCAHC	-----TDIF--STG-----GGRKLAEM

Homo sapiens-nbp35 LEVPLLRVPLDPLIGKNCCKGQS---FFIDAPDS-----PATLAYRSIIQRIQEF----CNLHQSKREENLISS
Saccharomyces cerevisiae-nbp35 LGIKFLGSLDPRIGKSCDMGES---FLDNYPDS-----PASSAVLNVVEALRDA----VGDV-----
Dictyostelium purpureum-nbp35 MNVPFLGRIPIDPLIARSCDEGKS---YLISHPDS-----EATKQYNIVFNSKYRY----RKKKKTTNY----
Entamoeba dispar-nbp35 MSVKFLGRIPLDPVIARSCDVGPS---YFIEHPDA-----DATKSFKEIYKQIIAN---I-----
Naegleria gruberi-nbp35 YEIPFLGSIPLDPMVMQSCETGKS---IVKDHPE-----PASQAMKEIVQKIITH----KKNQ-----
Guillardia theta-nbp35 HGIPFLGRIPLDPKLGIASERGEA-----LKN-----LALQSLQSAVQNLRLN----CKGNT-----
Encephalitozoon cuniculi-nbp35 NGIAYLGSIDLKQDIARSDSGDT---IEEE-----VLGKIVDAIMVV---CSSKA-----
Trypanosoma vivax FGIPLWGKIPLDQRLMKACEEGVS---LVNCTDEDT-----LTLKTLKLSAKLISA----AGLE-----
Giardia lamblia-nbp35 YKVKFLGRVPLDPQLTKASESGQA---WKKAVEEG-----TVSKGMEFMFYEVVVKGI----LSE-----
Blastocystis-nbp35 ANVPFLGRIPIFTSLEKAGEMGQG-----APAIAGGHVFFVVVSKIVNGMARR-----
Blastocystis sp. NandII-nbp35 ANVPFLGRIPIFTSLEKAGEMGQG-----APAIAGGHVFFVVVSKIVN-----
Phytophthora infestans-nbp35 FNVFPLGRPLDNDKMTGACEEGVS---FLEEYPDS-----VAAPAFGKIVQDLVAA----VEK-----
Phaeodactylum tricornutum-nbp35 YGVFFMGRPLDPELLKACEQGKS---FVQTHPNA-----NAAVALKQFARQLNKV----LPVNMDE-----
Arabidopsis thaliana-nbp35 MGVPFLGKVPMDPQLCKAAEQGKS---CFEDNKCL-----ISAPALKSIQKVVPS----TVMTE-----
Micromonas sp.-nbp35 AGVEYLGSIPLDPAIVASEKGLS---LFSKAEVSDKDLSDAPIQDFASLTSLSNSIIDALLLQ----LNEGKNR-----
Cyanidioschyzon merolae-nbp35 SGVRFLGRIPLDPALMRASDLGEA---ILNY--DC-----PARDAI FELARHIGKA----FSPKTIASLEVESS
Cryptosporidium hominis-nbp35 MEVEYLNKI PWDELLYVCDLGLS---ICEKFPQS-----PSSIGIKKLVDI IYQ----SKIN-----
Tetrahymena thermophila-nbp35 YKIDLLGKVPLEPKVLICTEKGKS---IVKEHPDS-----VAAKVYQHTAERTVQT----LKVELPK-----
Metallospira cuprina LGVPLLGGVPLDPRIAESNDLGEF---FFLKYLD-----PASKEFLKIADEVIEQ----VENQKLSDLNLK--
Pyrococcus furiosus EGVDFLGVPI DLKARASDLGIP---IVL-YGDT-----PAAKAFMEIAEKL VNK----LKEIKGDGREK--
Methanococcus maripaludis LNVFPLGRIPLDIKARVASDRGVP---MVTM--DC-----KASEEFKVVNTVLER----IKKE-----
Homo sapiens-cfd1 AGVPFLGSLDPLMRMTLEEGHD---FIQEFPGS-----PAFAALTSIAQKILDA----TPACLP-----
Saccharomyces cerevisiae-cfd1 FSVPYLGNVPI DPKFVEMINQVSSKTLVEMYRES-----SLCPI FEEIMKKLRQDTTTTPVVDKHEQPQIESPK
Dictyostelium purpureum-cfd1 CNIKFLGKLPIDPNLSICSERGIN---YFKEYPNS-----STLSALKTFAEALNN-----
Entamoeba dispar-cfd1 YQLKFGVGAIPIEPKICLAGETGVN---PFADEPFA-----NALKPITDFVANLAKTFA-----
Naegleria gruberi-cfd1 YNIPFLGSIPIEPELANAEDNGIN---YIKNFSNS-----VTSMQFTNIVNIILN-----
Guillardia theta-cfd1 MGMKFLGRIPIELKWSEMDRGM---TIQGEERK-----AVMLPFKQVIHGIIASA-----
Encephalitozoon cuniculi-cfd1 TGIFFVCRLPIDSLLEALDEGR---FVERCGSI-----EAYMKFRKAVLGLAD-----
Trypanosoma vivax-cfd1 YEVEFLGAIPI DRLSLAEDNGQC---FITTSACPEGSPG-----DGVTVAVTNVINAIVNQ----AEKVLLQRSNA--

Figure S5



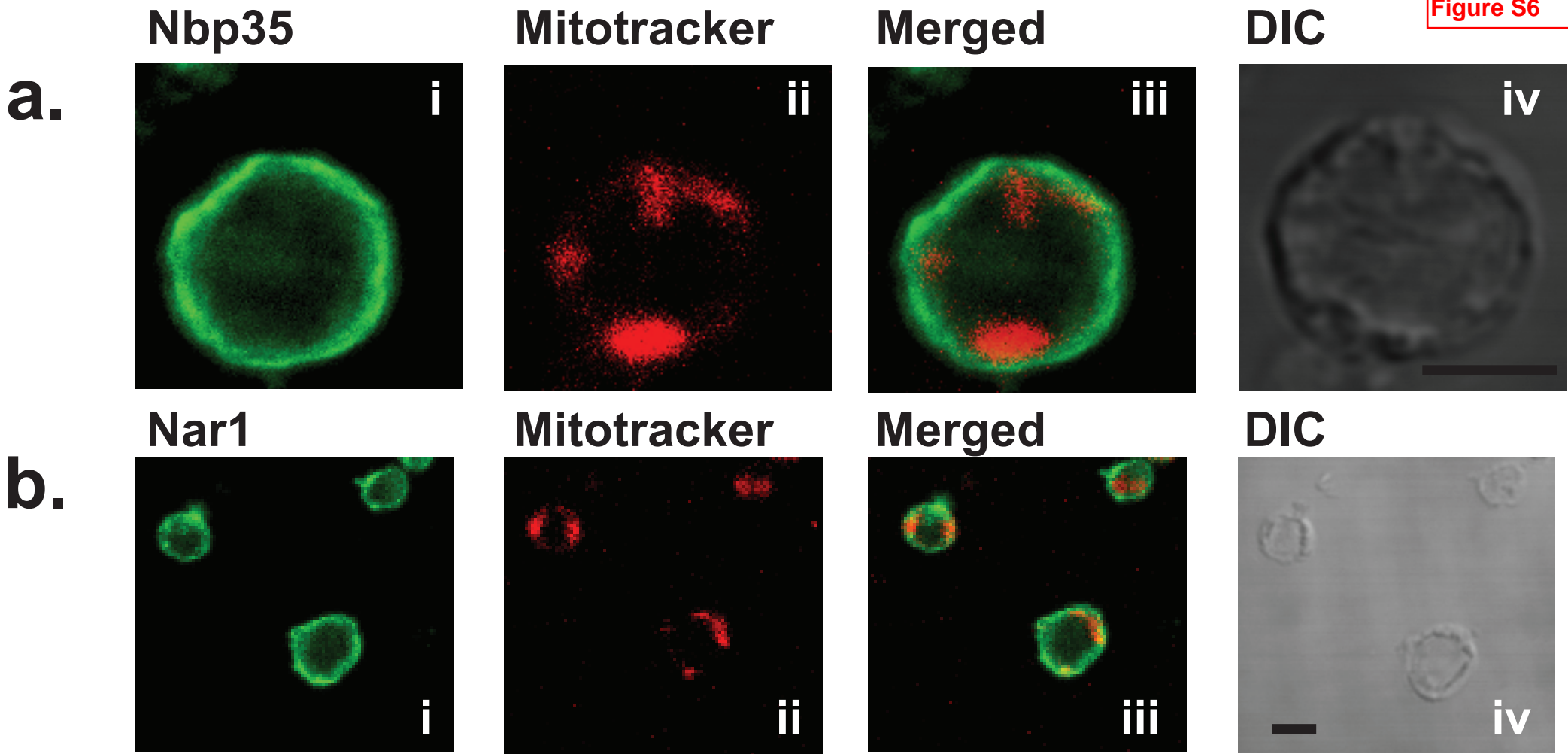


Figure S7 - Localization of Nbp35

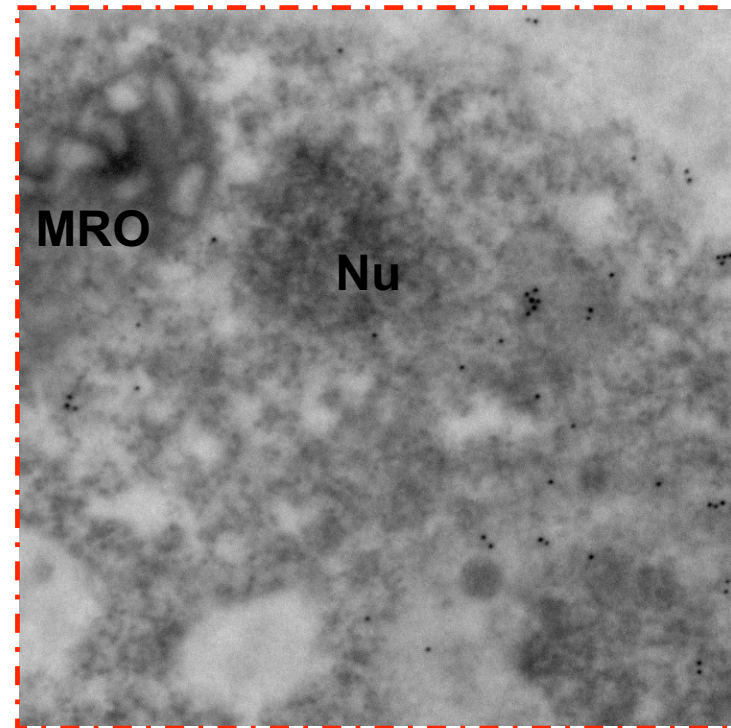
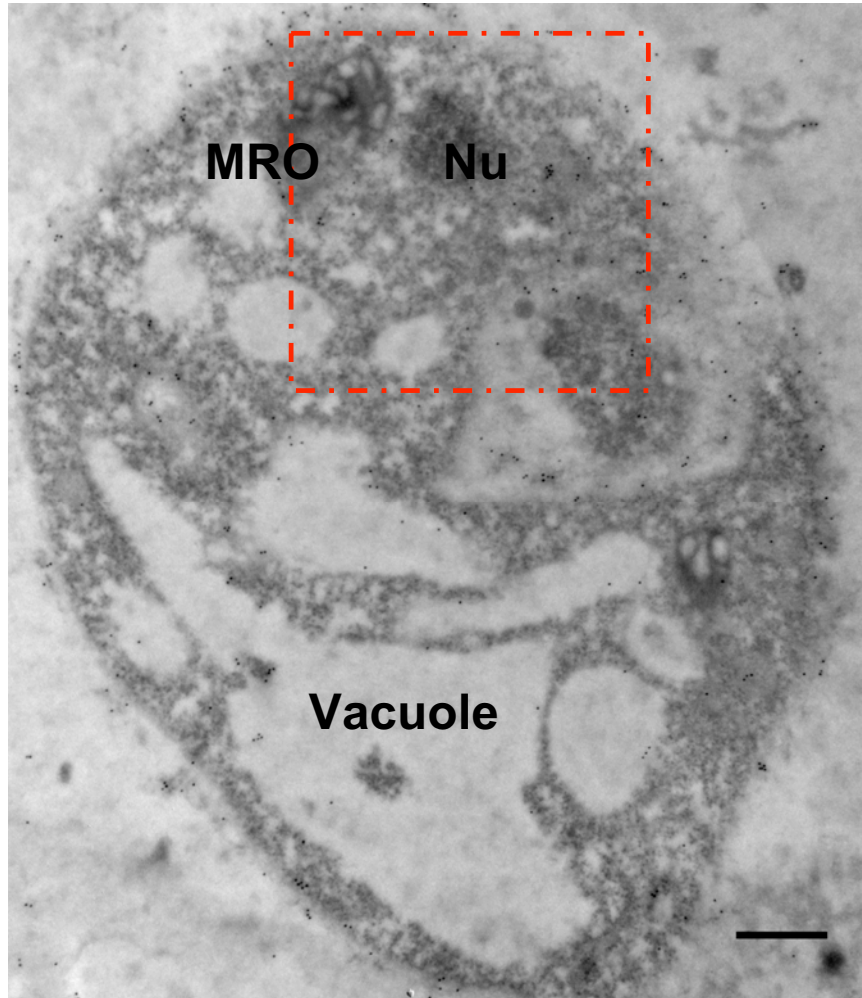


Figure S8 - Localization of Nar1

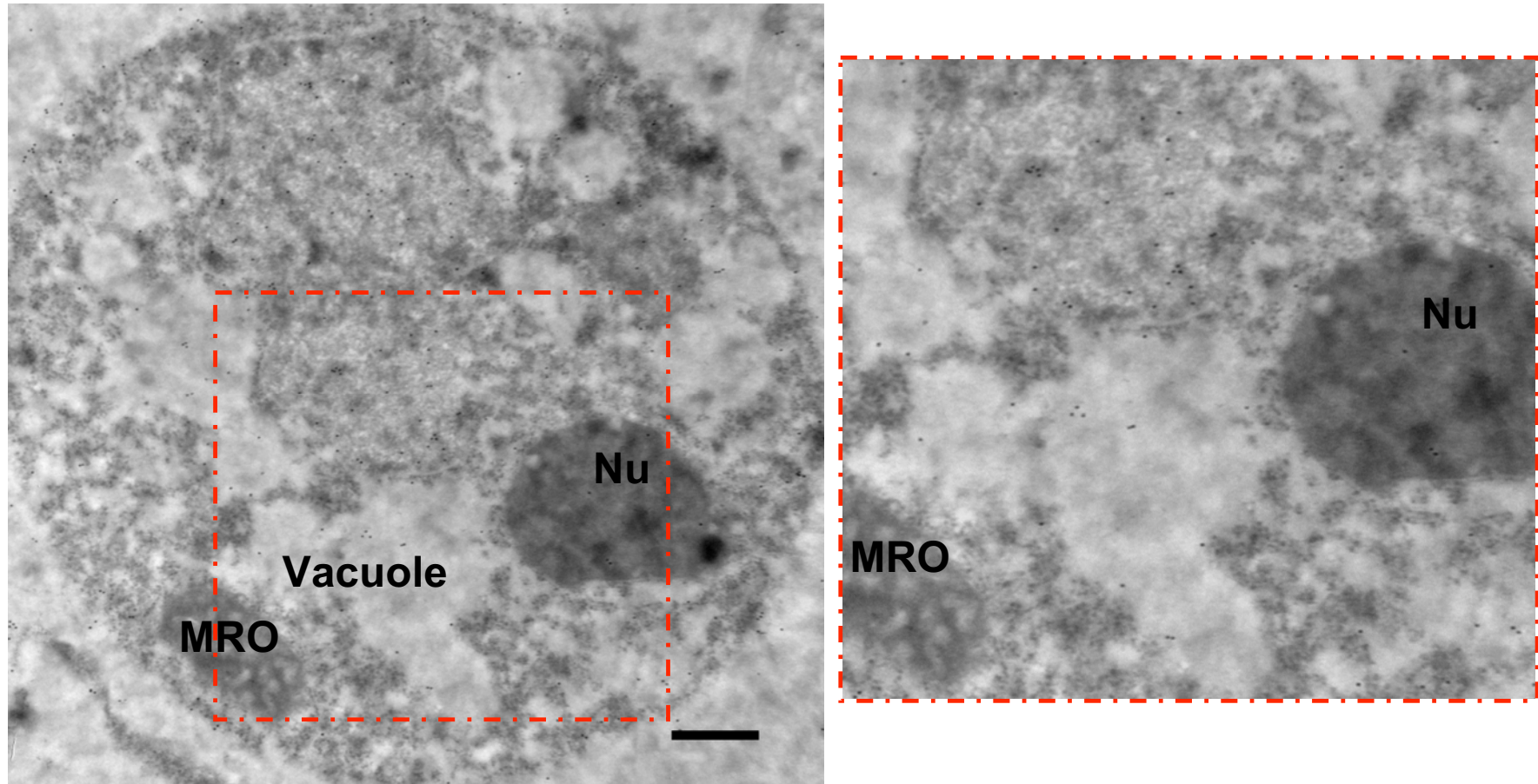


Figure S9 - Localization of Tah18

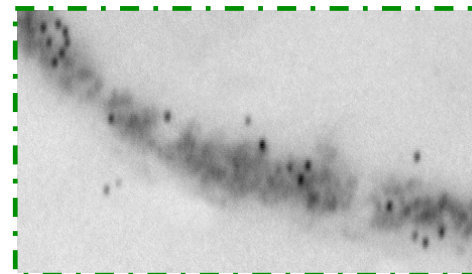
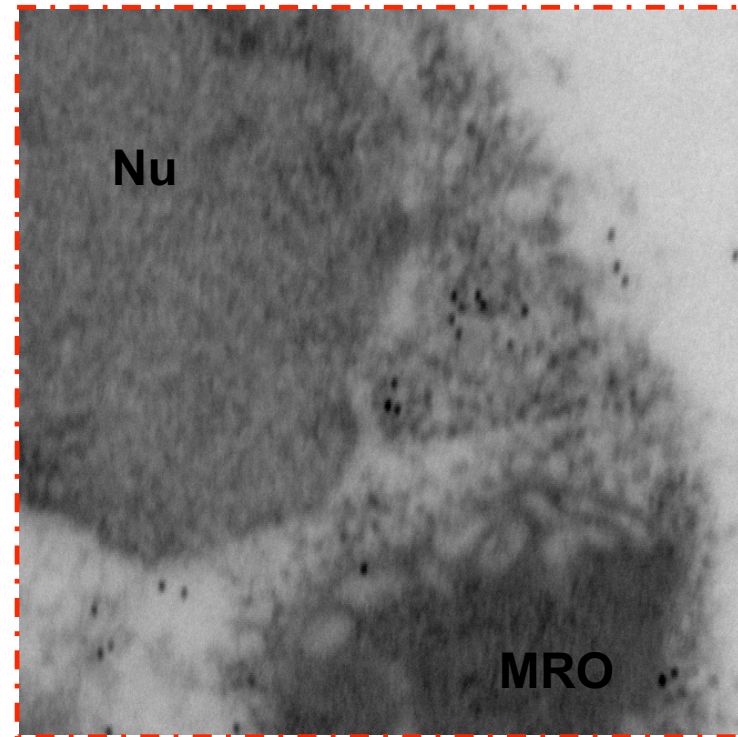
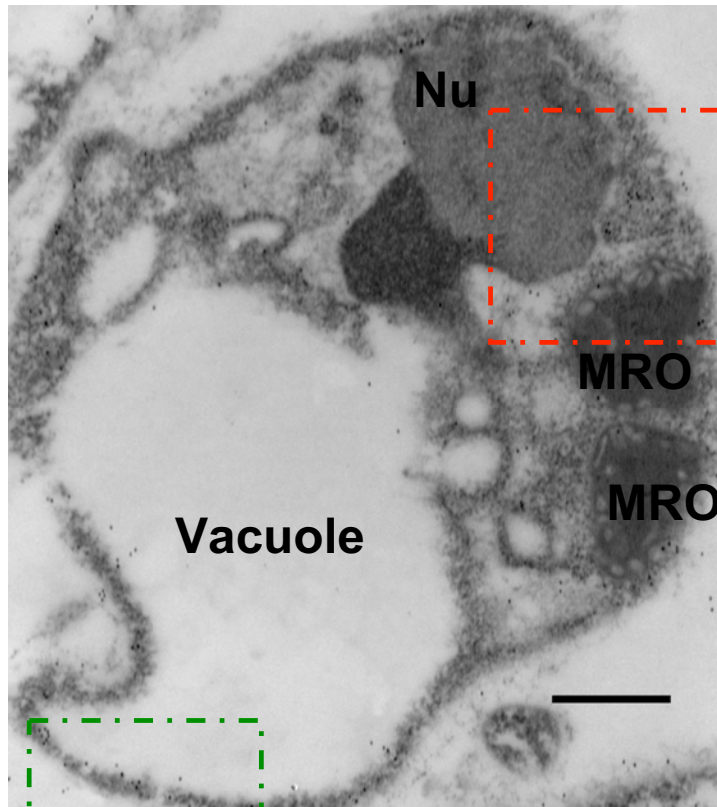
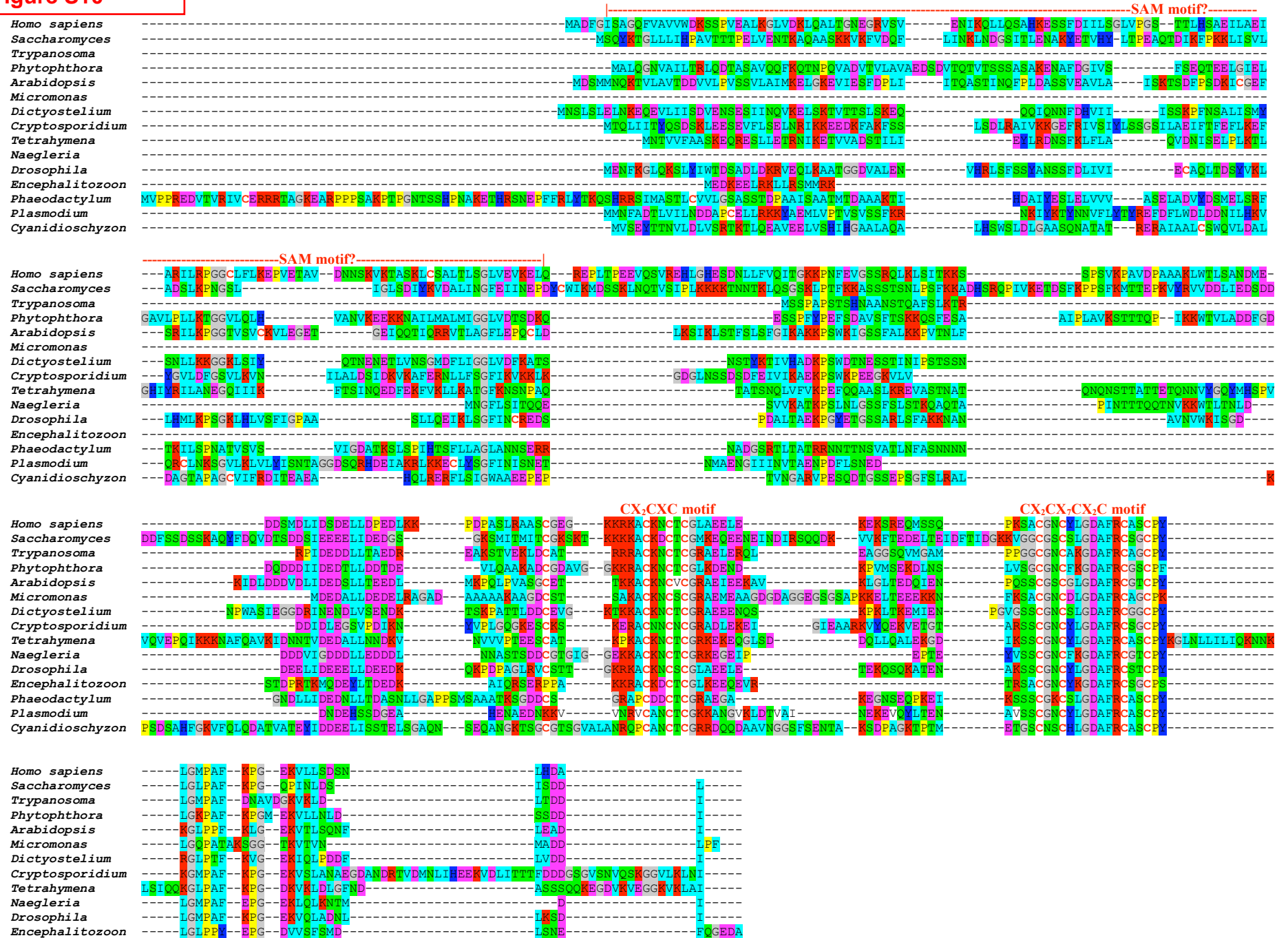


Figure S10



<i>Phaeodactylum</i>	---	L	K	P	A	F	---	K	P	G	---	E	E	L	V	L	D	---	L	O	D	---	F	---		
<i>Plasmodium</i>	---	K	G	L	P	A	F	---	C	P	G	---	E	N	V	R	L	N	L	D	---	N	E	F	N	---
<i>Cyanidioschyzon</i>	---	L	G	L	P	P	F	---	K	P	G	---	A	P	L	K	I	D	A	K	---	L	R	G	D	---

Figure S11

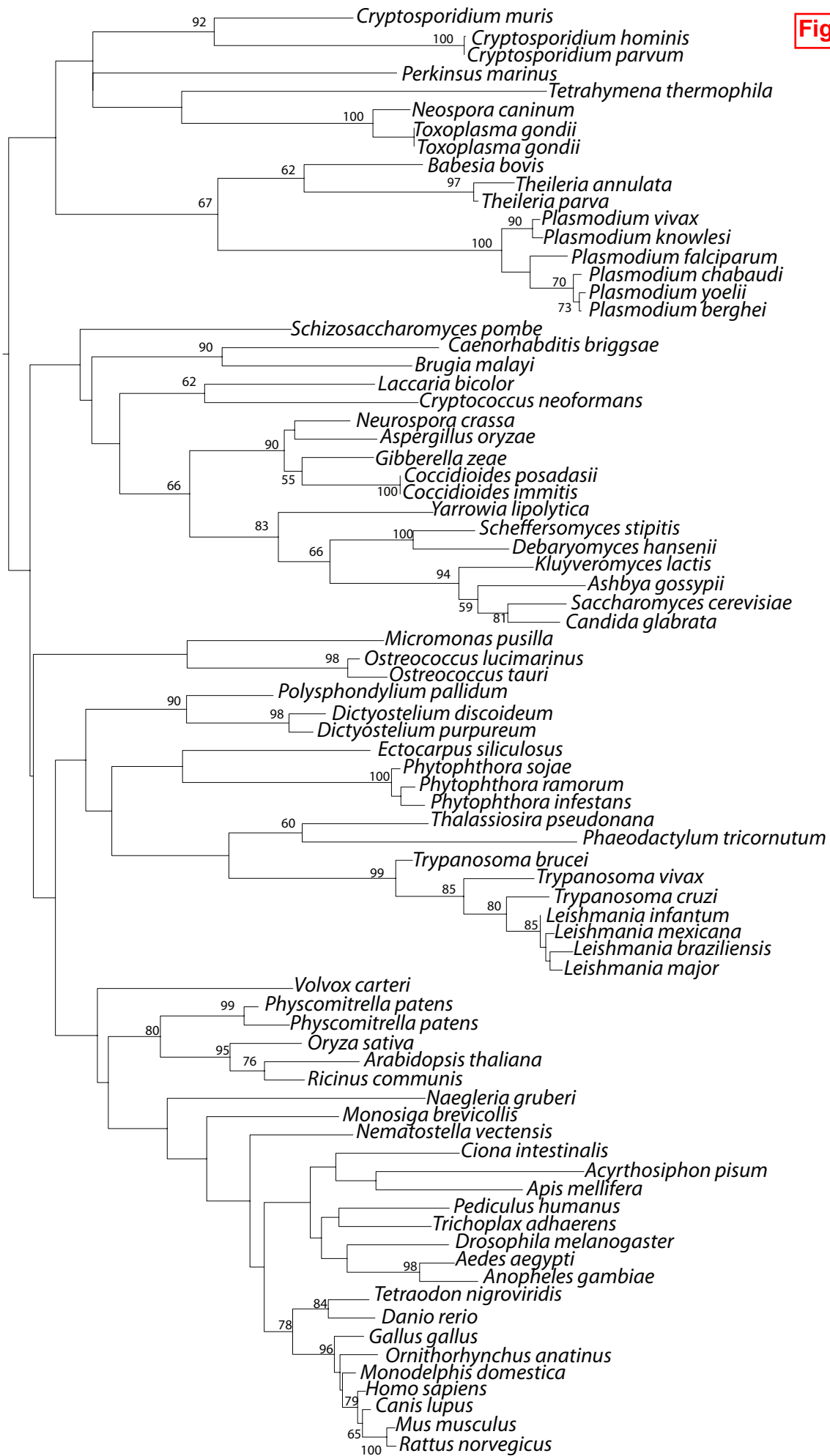
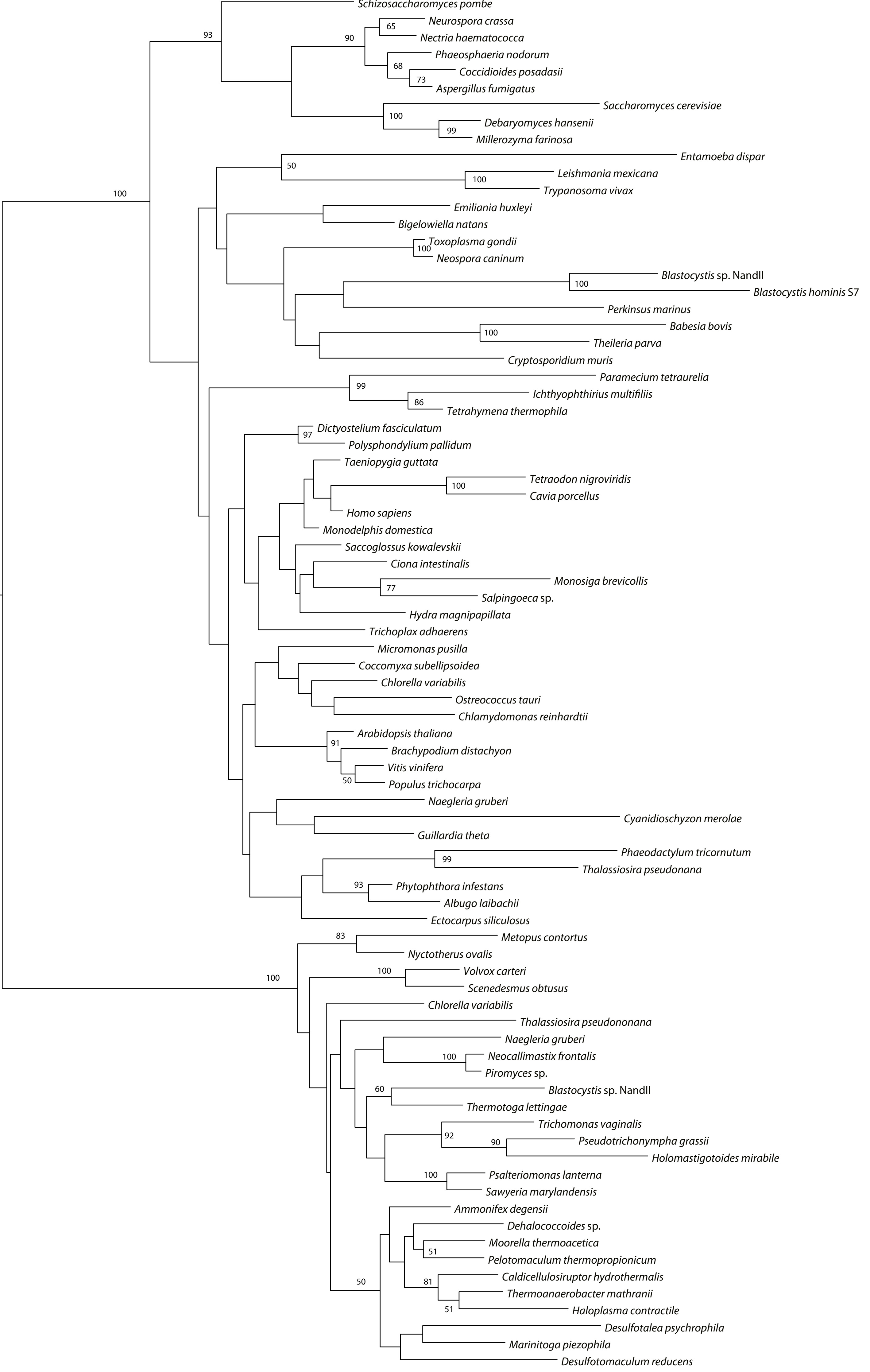


Figure S12



0.3

Figure S13

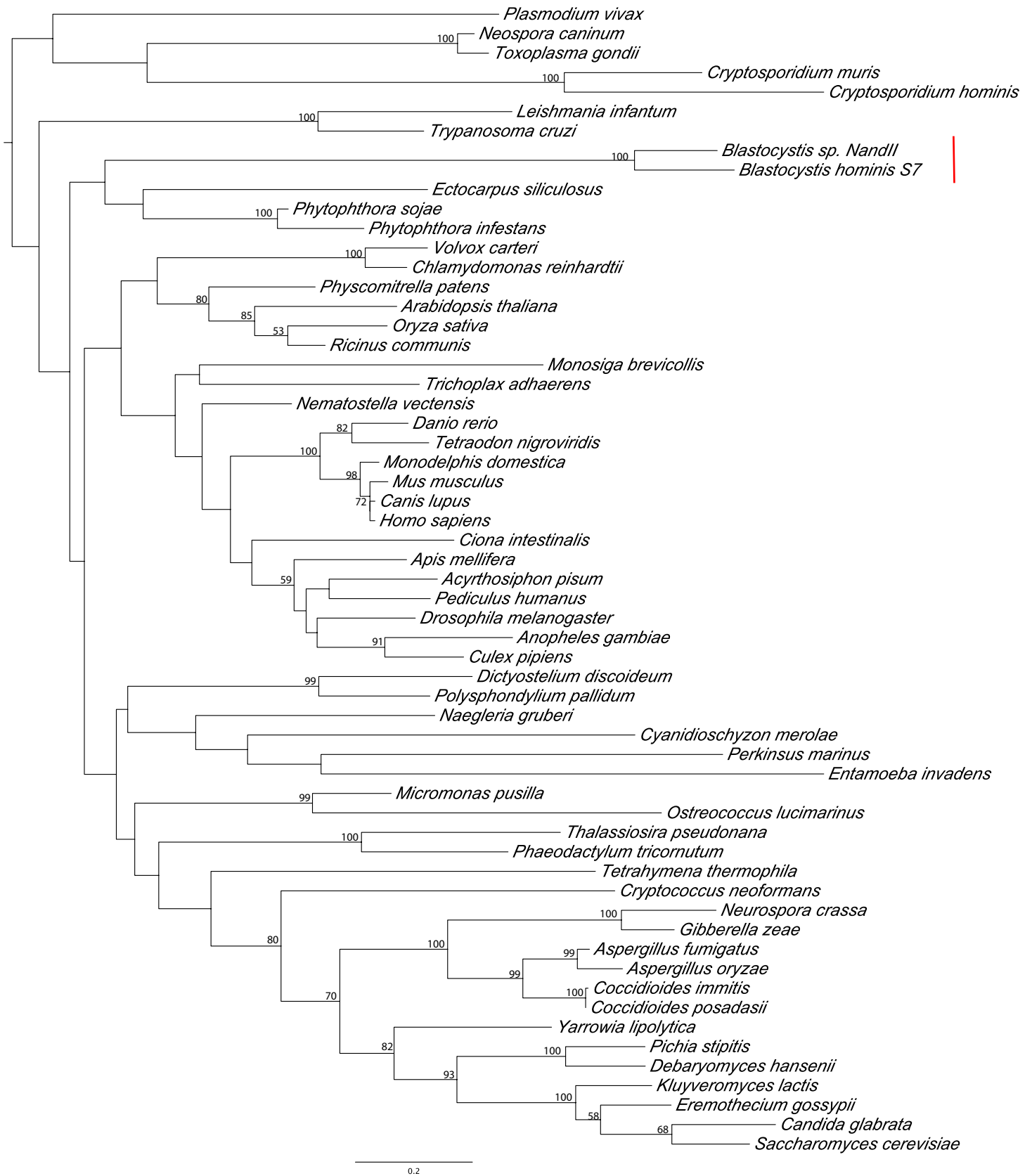


Figure S14

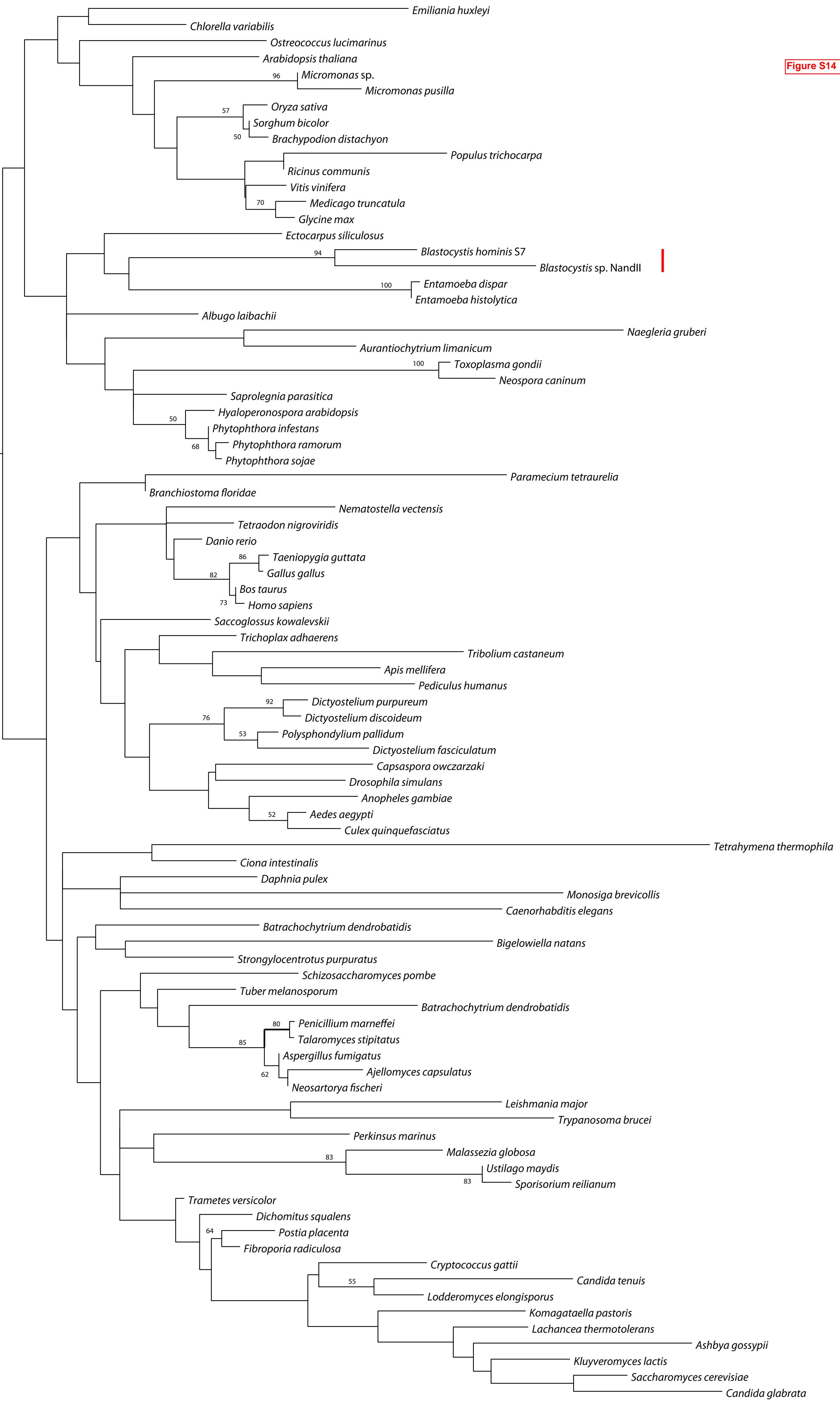


Figure S15

