Evolution, diverse life styles and the extreme complexity of the mitochondrion of kinetoplastid flagellates

Pavel Flegontov, Jiří Týč, Tomáš Skalický, Eva Horáková, Eva Dobáková, Drahomíra Faktorová, Hassan Hashimi
(Institute of Parasitology, CAS and University of South Bohemia, České Budějovice [=Budweis], Czech Republic)

Jan Votýpka (Charles University, Prague, Czech Republic)
Vyacheslav A. Yurchenko (University of Ostrava, Czech Republic)
John Archibald (Dalhousie Univ., Halifax, Canada)
Laurie Read (SUNY, Buffalo, USA)

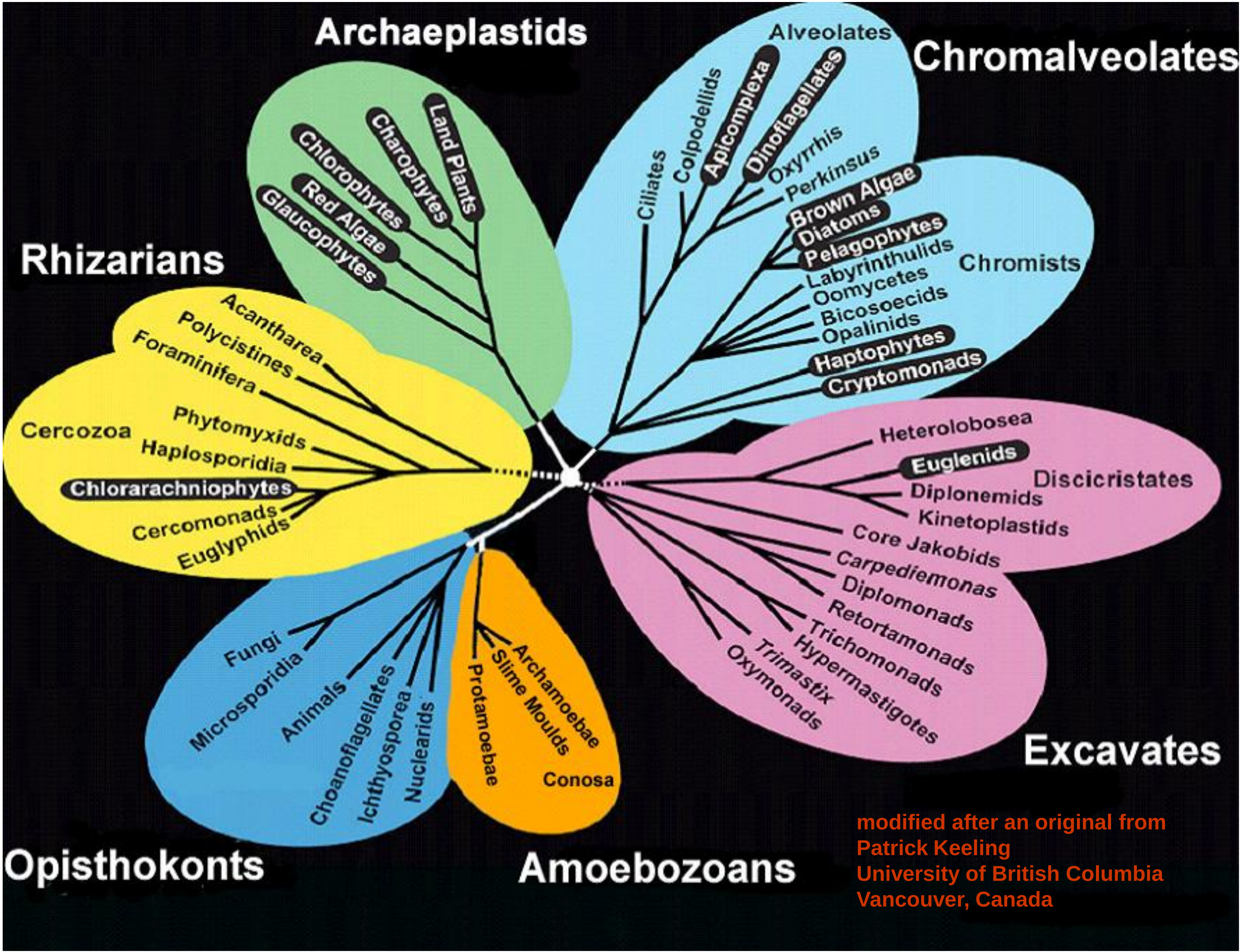
1. **Extreme complexity** of the kinetoplastid mitochondrion

2. **Diversity of life styles** – novel endosymbiotic relationship of a kinetoplastid with an amoeba

3. **Novel endosymbiotic bacterium** in kinetoplastids

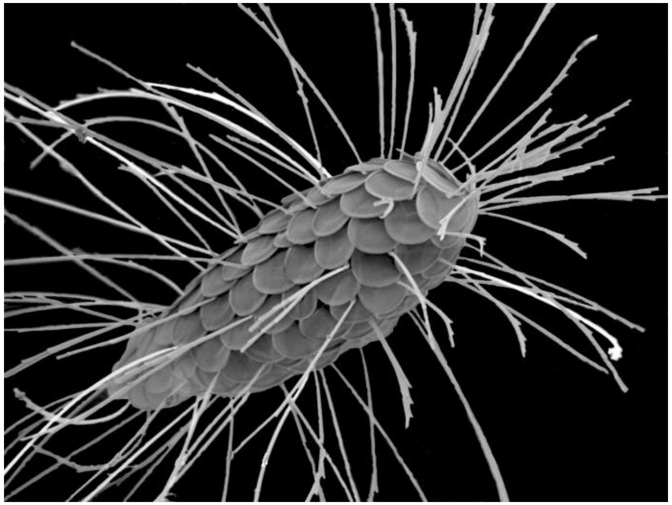
4. **An early-branching trypanosome** and solution to a 100 year-old debate

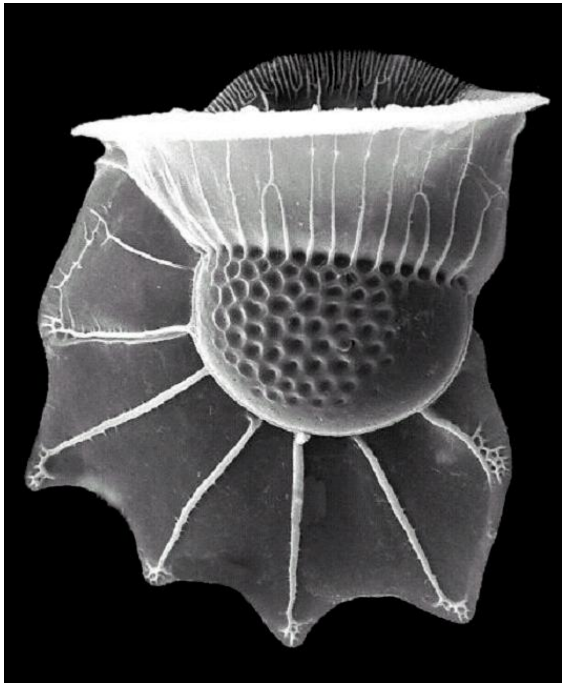
5. **Diversity of trypanosomatids** – what does it say?

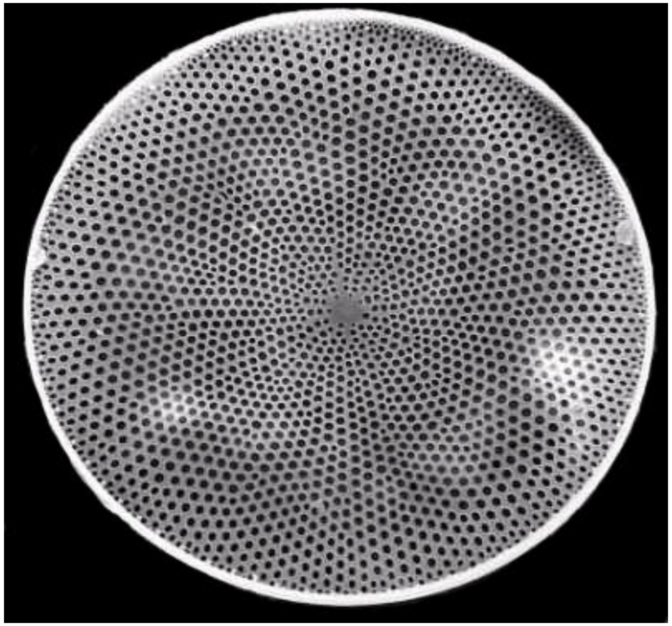


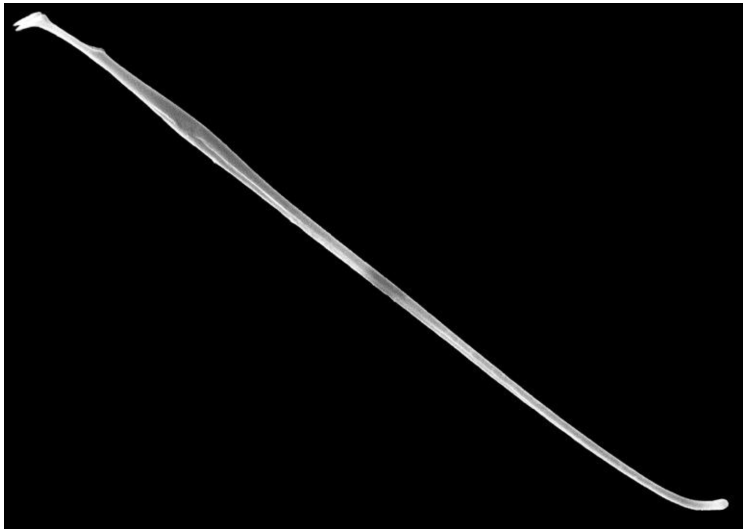
modified after an original from
Patrick Keeling
University of British Columbia
Vancouver, Canada

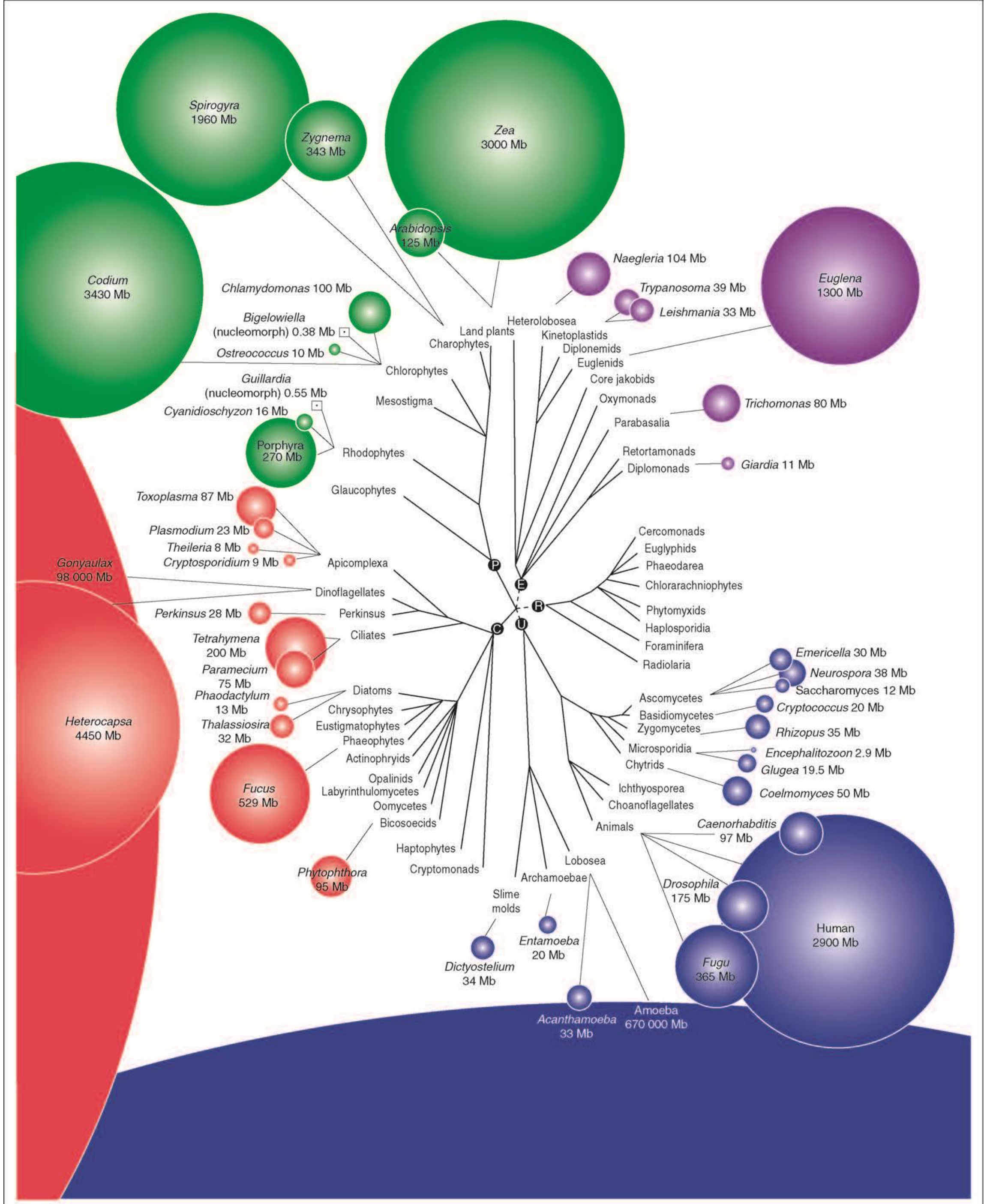
Enormous diversity of protists in morphology







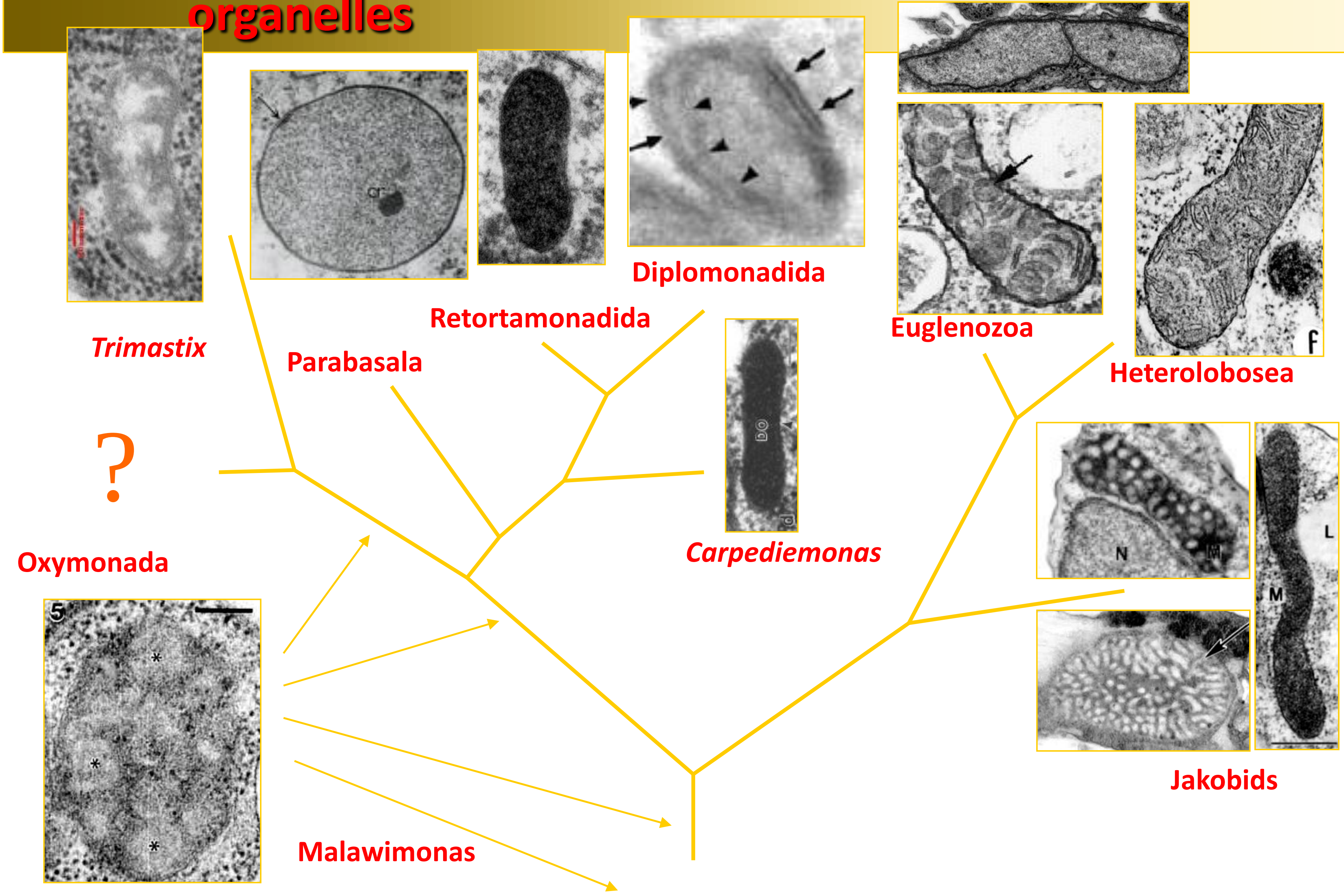




Current Opinion in Genetics & Development

Enormous diversity of genome size in protists

Enormous diversity of mitochondrion-derived organelles



Trimastix

Parabasala

Retortamonadida

Diplomonadida

Euglenozoa

Heterolobosea

?

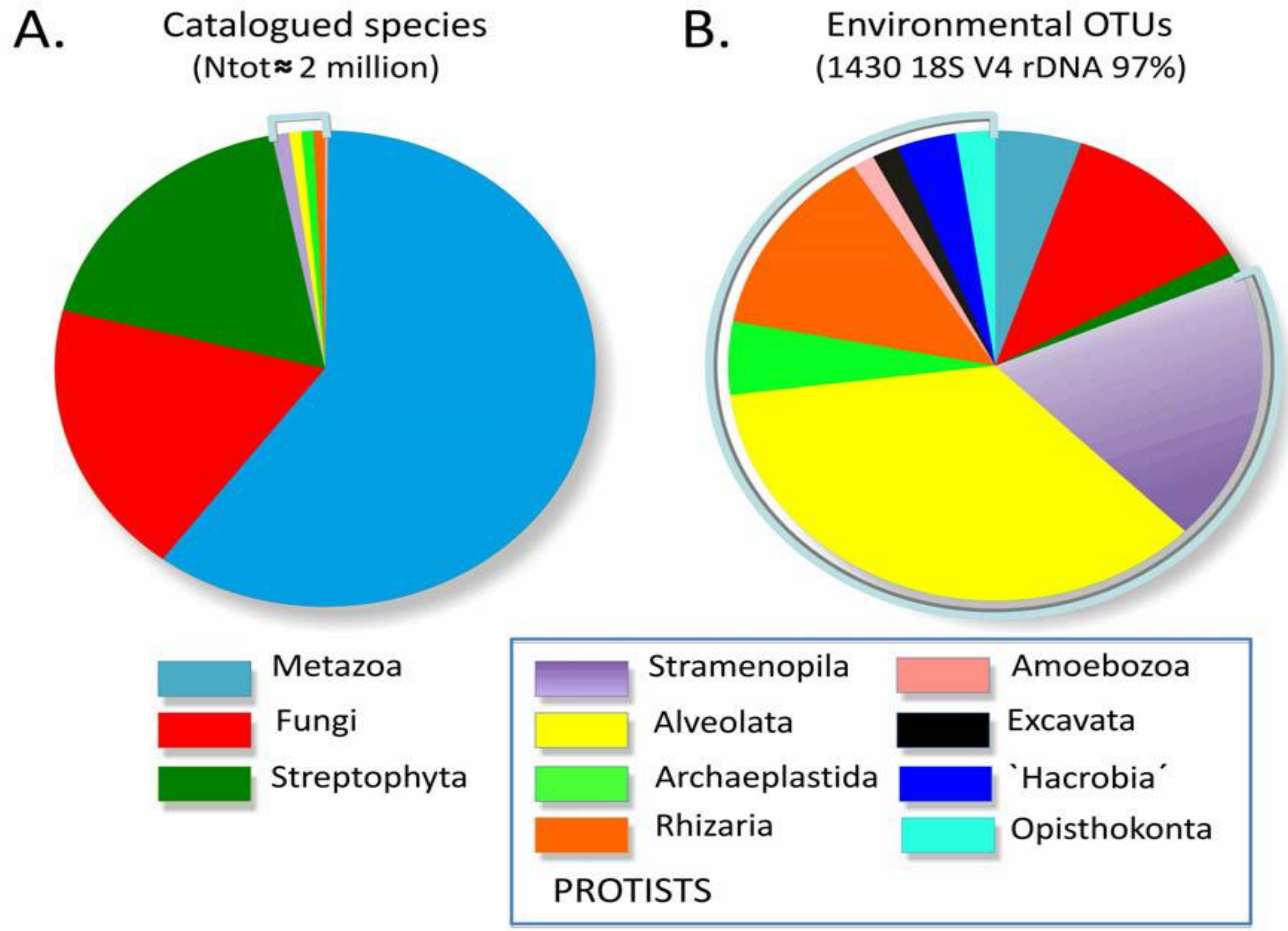
Oxymonada

Carpediemonas

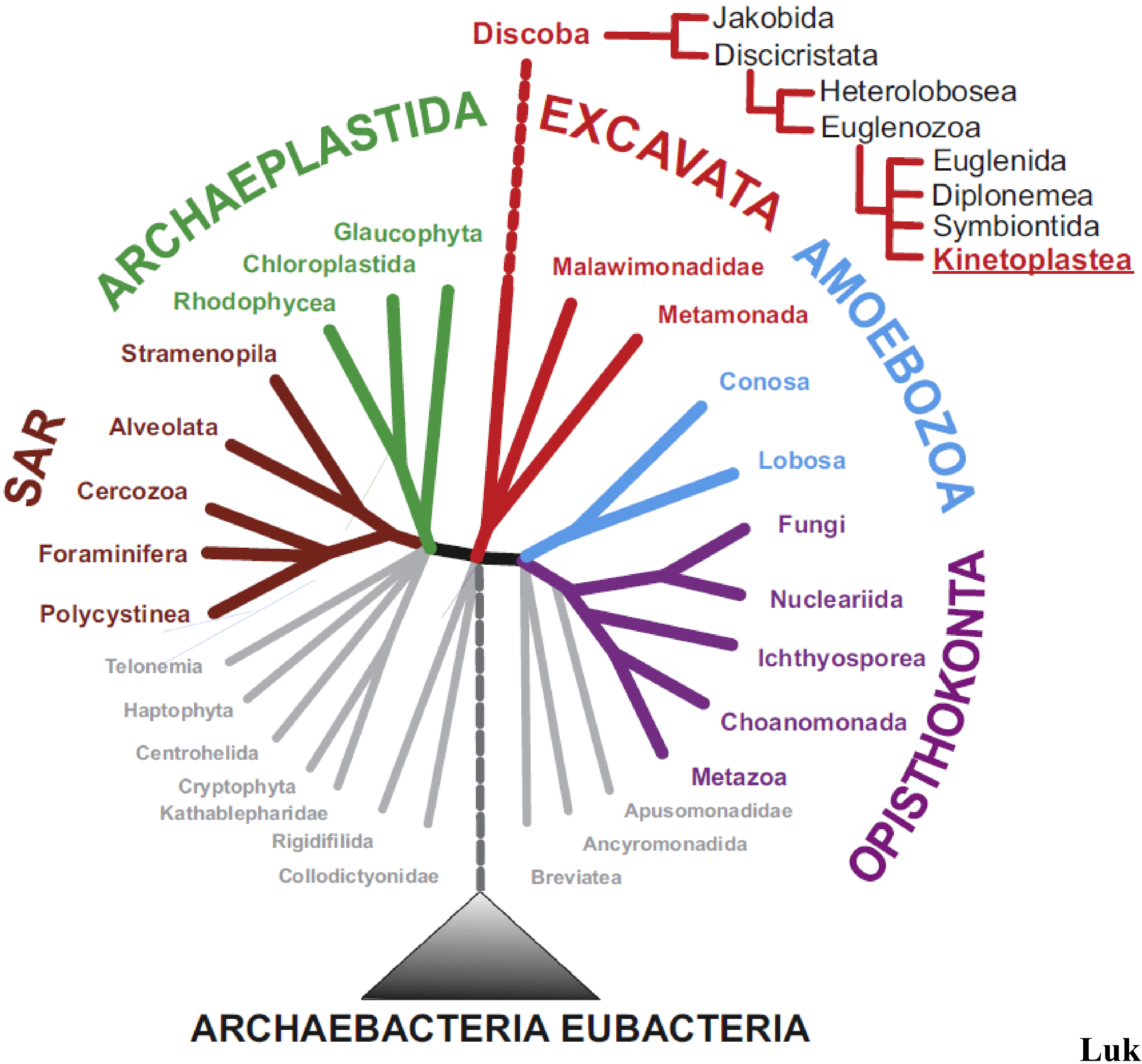
Jakobids

Malawimonas

Enormous diversity in number of species

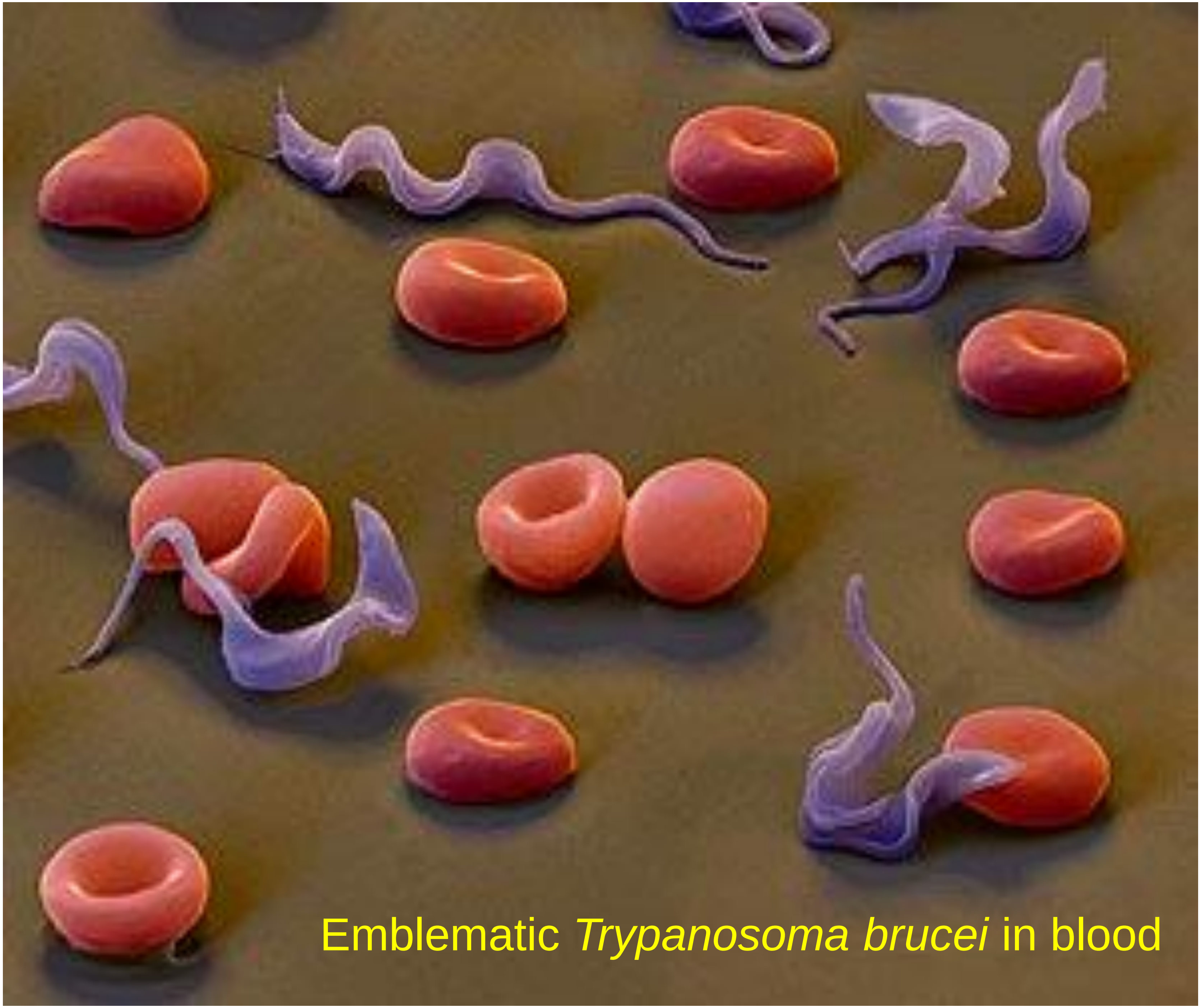


Morphological versus genetic views of total eukaryotic diversity. (A) Relative numbers of described species per eukaryotic supergroup. (B) Relative number of V4 18S rDNA Operational Taxonomic Units (97%) per eukaryotic supergroup, based on 59 rDNA clone library surveys of marine, fresh-water, and terrestrial total eukaryotic biodiversity.



Current view
of the
subdivision
of Discoba

Lukeš et al., MBP 2014

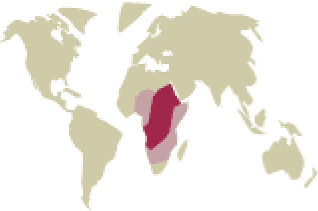


Emblematic *Trypanosoma brucei* in blood

TriTryps

Trypanosoma brucei	*Trypanosoma cruzi*	*Leishmania*
extracellular parasite 	**intracellular parasite** 	**intracellular parasite** 
Tse-tse fly 	**Reduviid bugs** 	**Sand fly** 
Sleeping sickness	**Chagas disease**	**Leishmaniases**

The scourge of trypanosomiases



Sleeping sickness

- 66 000 deaths per year
- 60 million people are at risk of contracting sleeping sickness

Leishmaniasis

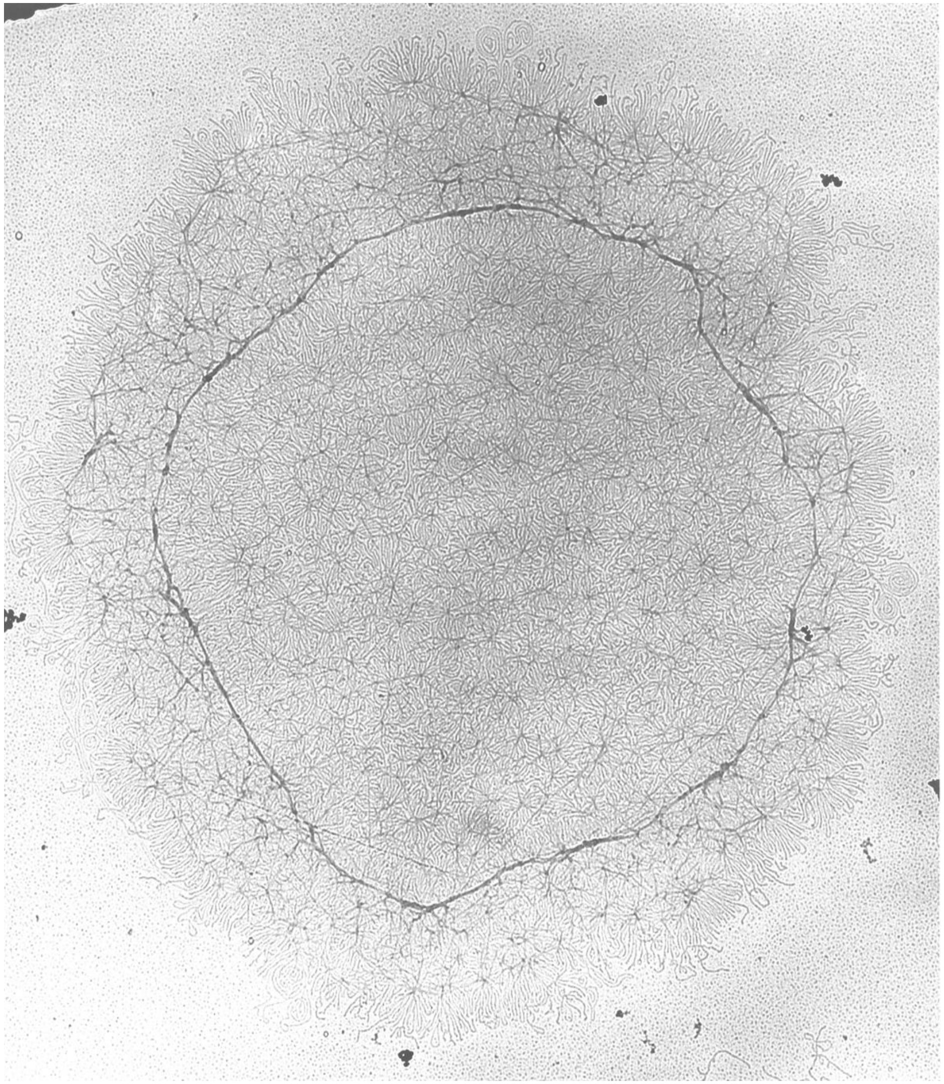
- 350 million people in 88 countries around the world are threatened
- 1.5 – 2 million new cases of leishmaniasis estimated to occur annually

Chagas’ disease

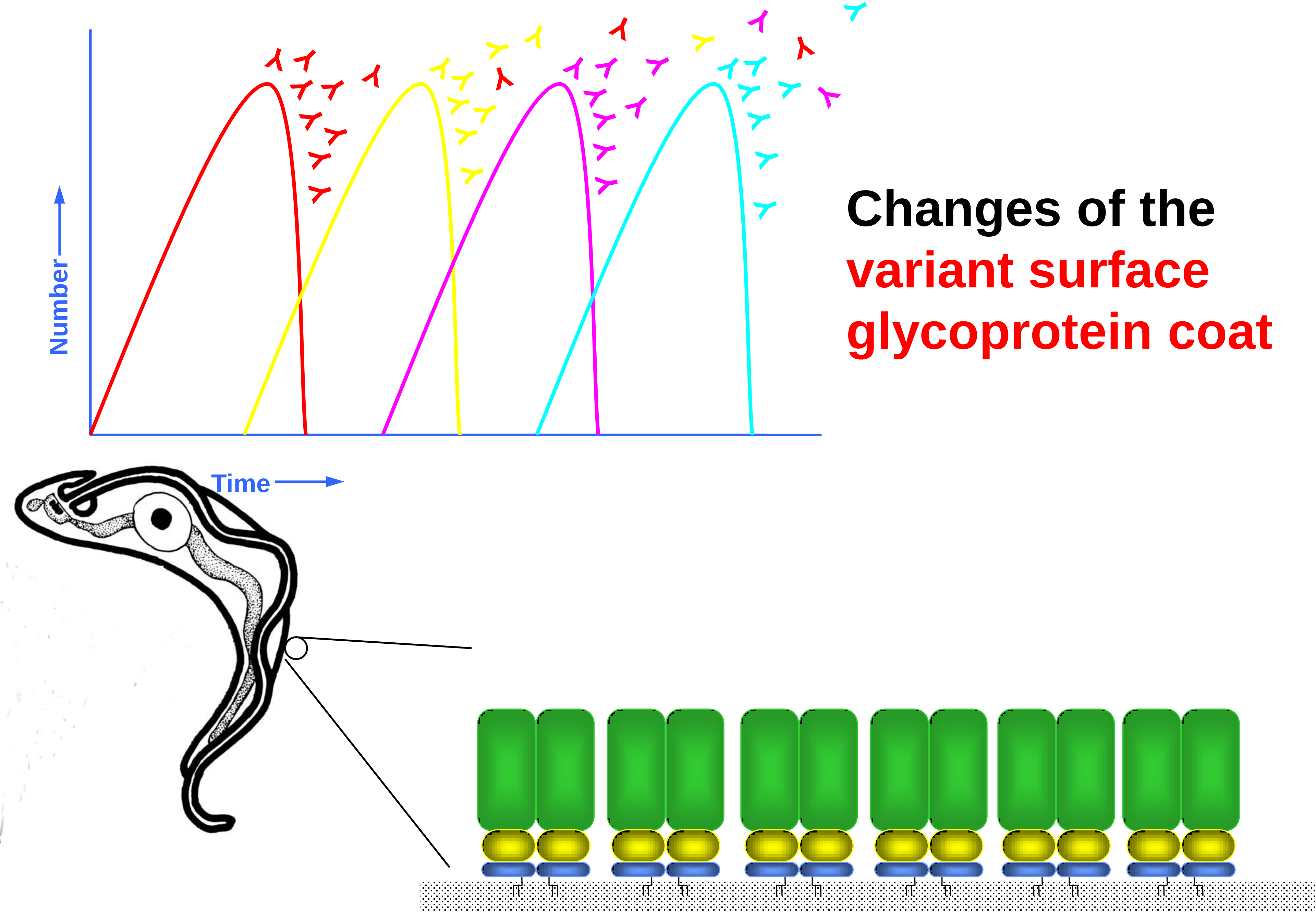
- 55 000 deaths per year
- 14-16 million people are infected each year

Trypanosomes have many unusual properties, but nothing beats what they do in their mitochondrion

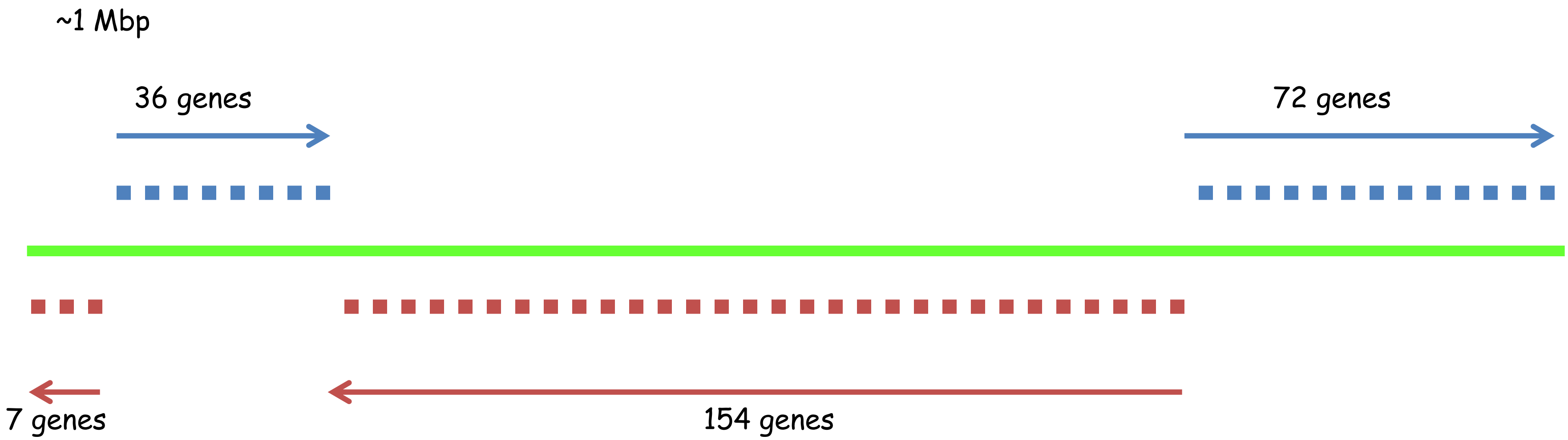
- antigenic variation
- glycosomes
- trans/splicing
- polycistronic transcription
- base J
- mt DNA network
- RNA editing in mt mRNAs



Changes of the **variant surface glycoprotein coat**



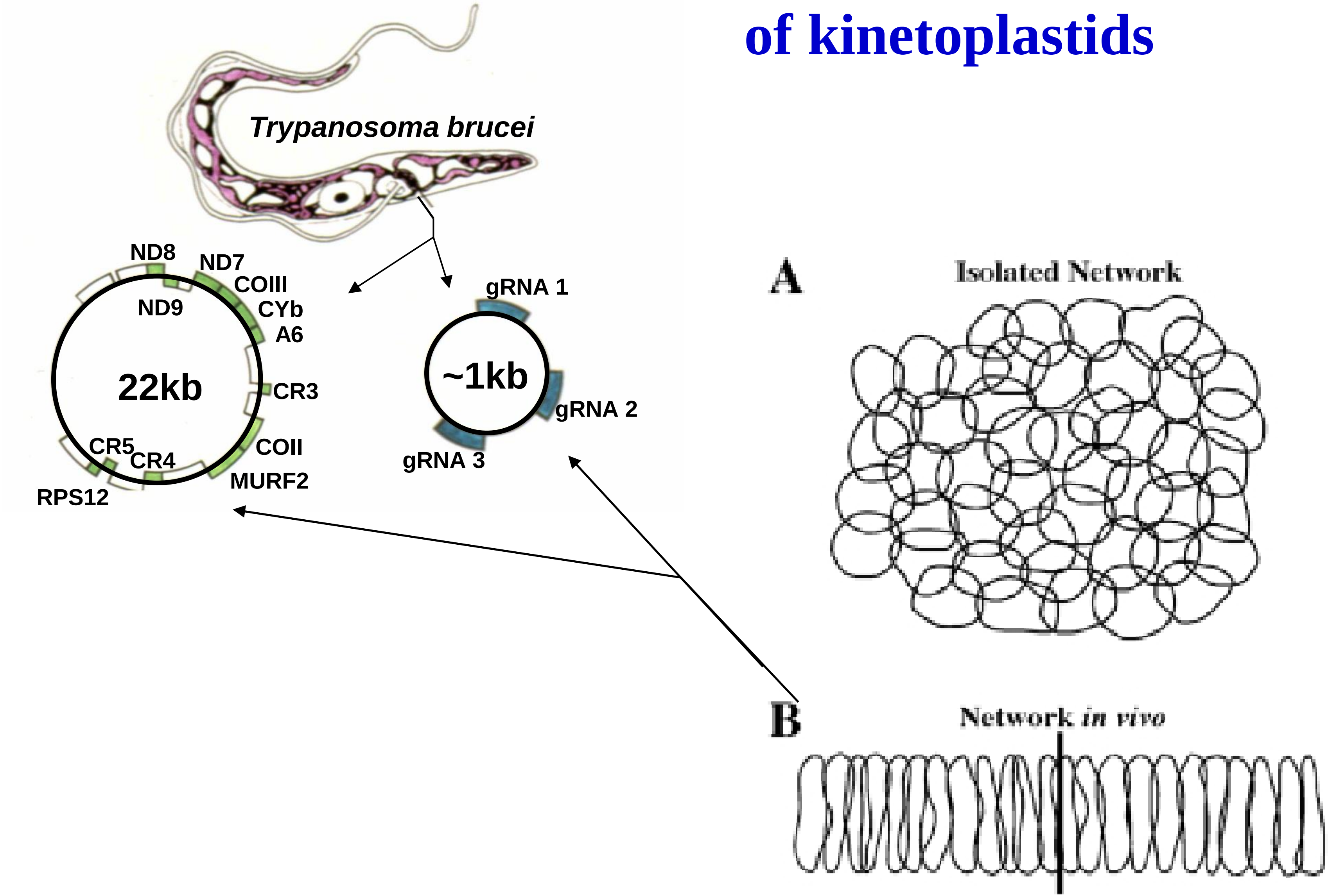
Nuclear genes are arranged in large, co-transcribed, tandem arrays



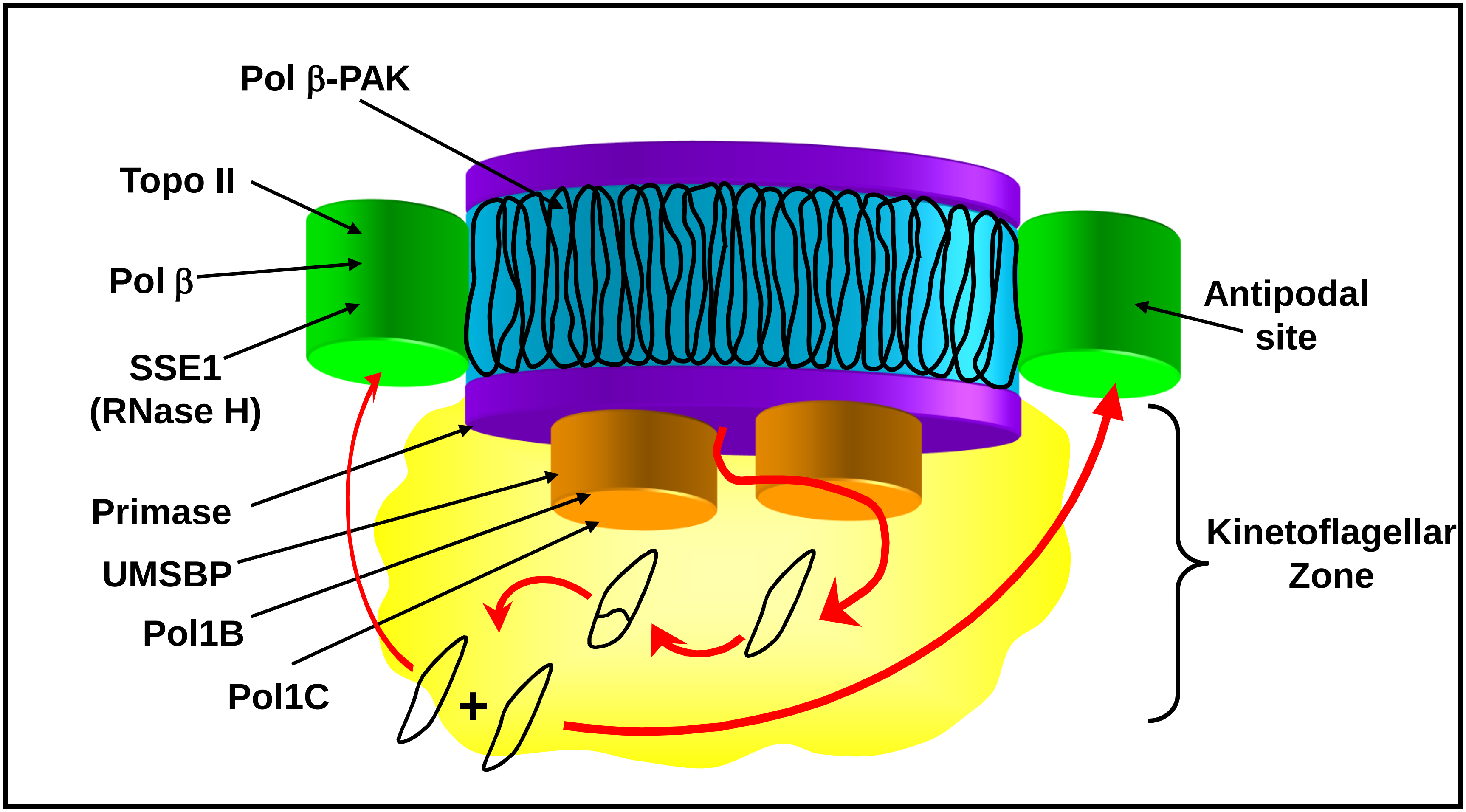
The production of a monocistronic mRNA is dependent on *trans* splicing of a capped mini-exon to the 5’ end of the mRNA in a process is linked to the cleavage and polyadenylation of the 3’ end of the upstream RNA

The selective advantage in such a strategy is unclear but it has several downstream consequences for genome structure and the regulation of gene expression

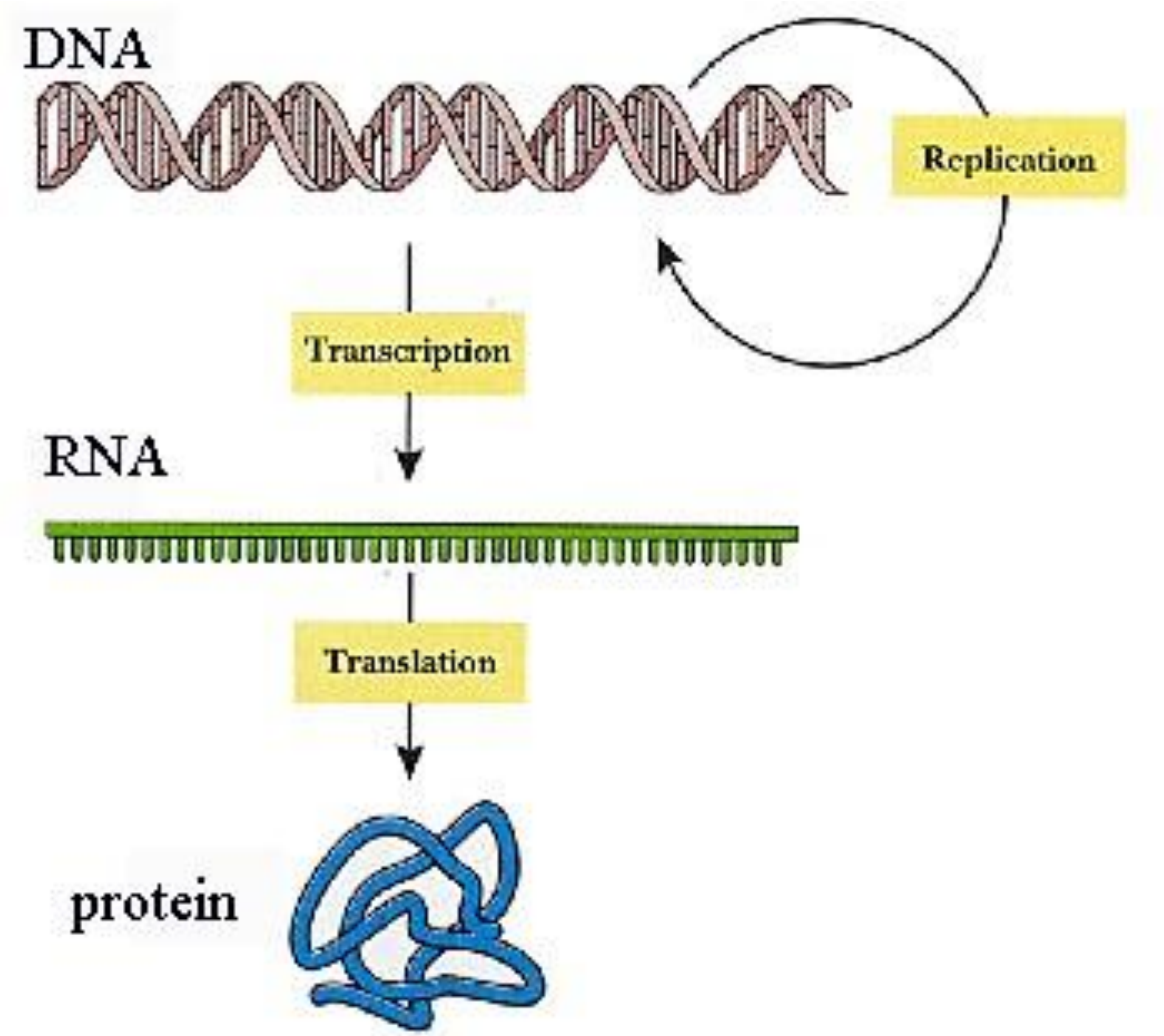
Mitochondrial genome (kinetoplast DNA) of kinetoplastids



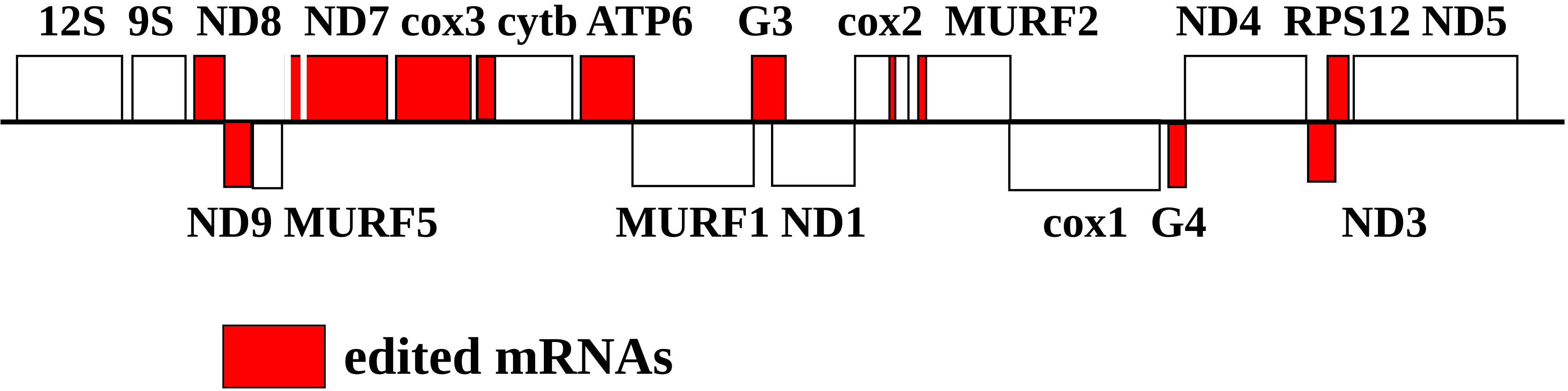
The kinetoplast DNA replication model



Central Dogma of Molecular Biology



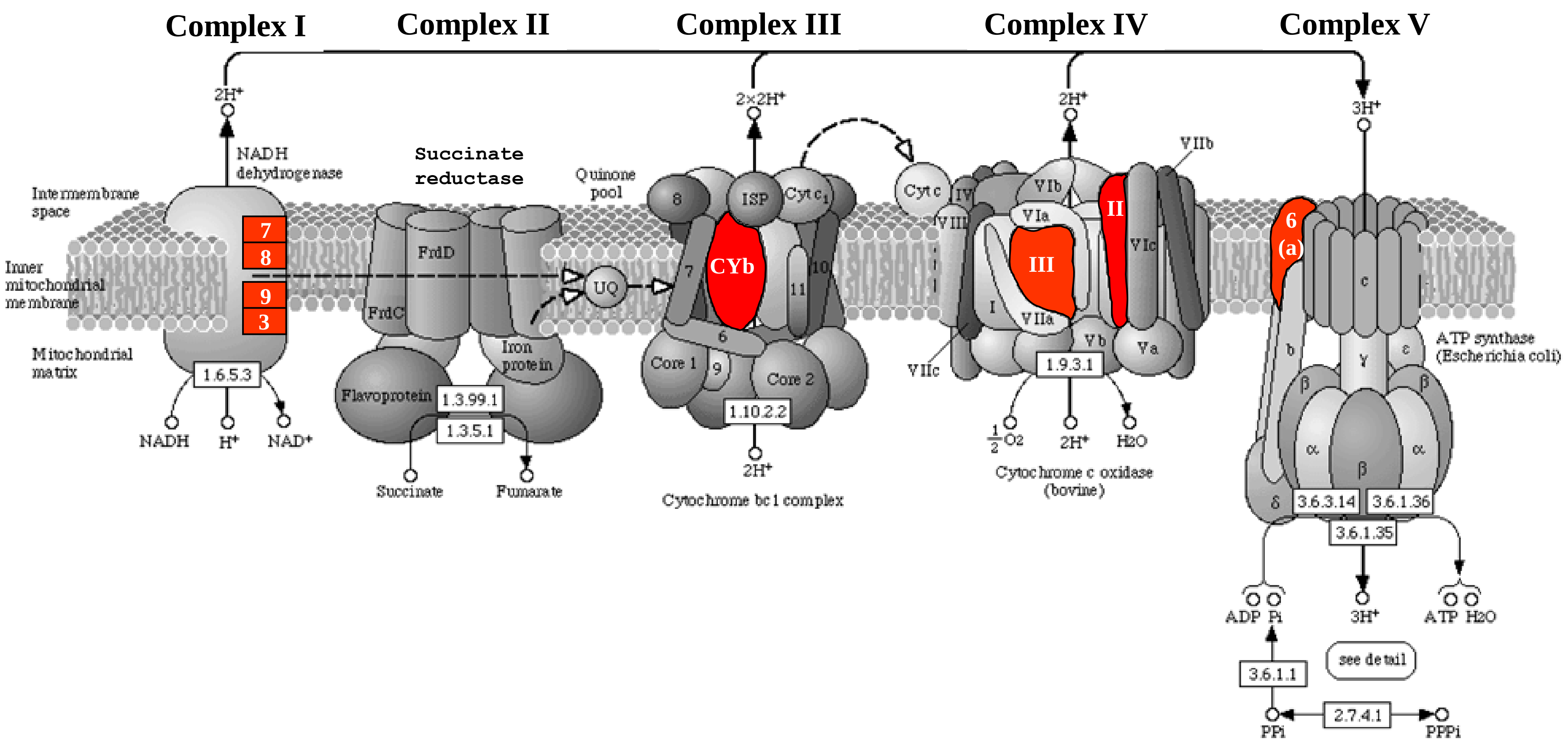
In *T. brucei*, extensive RNA editing is the rule: 9 of 18 mRNAs are extensively edited


12S 9S ND8 ND7 cox3 cytb ATP6 G3 cox2 MURF2 ND4 RPS12 ND5

ND9 MURF5 MURF1 ND1 cox1 G4 ND3

🟥 edited mRNAs

Most edited mRNAs encode subunits of the OxPhos system



Complex I

Complex II

Complex III

Complex IV

Complex V

Subunits encoded by mitochondrial mRNAs that require editing

Editing is complex but accurate

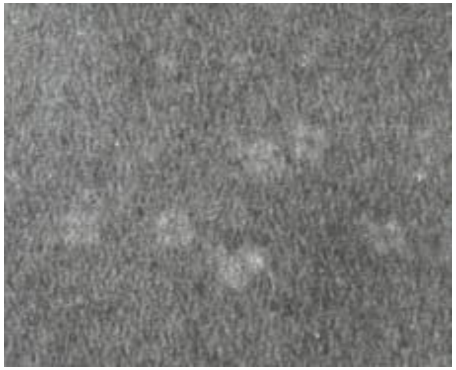
- Editing employs a series of coordinated enzyme steps
- Each enzyme contributes to the accuracy of editing
- gRNAs specify number and sites of U insertion/deletion
- Precisely edited and translatable mRNA are produced

The Editosome

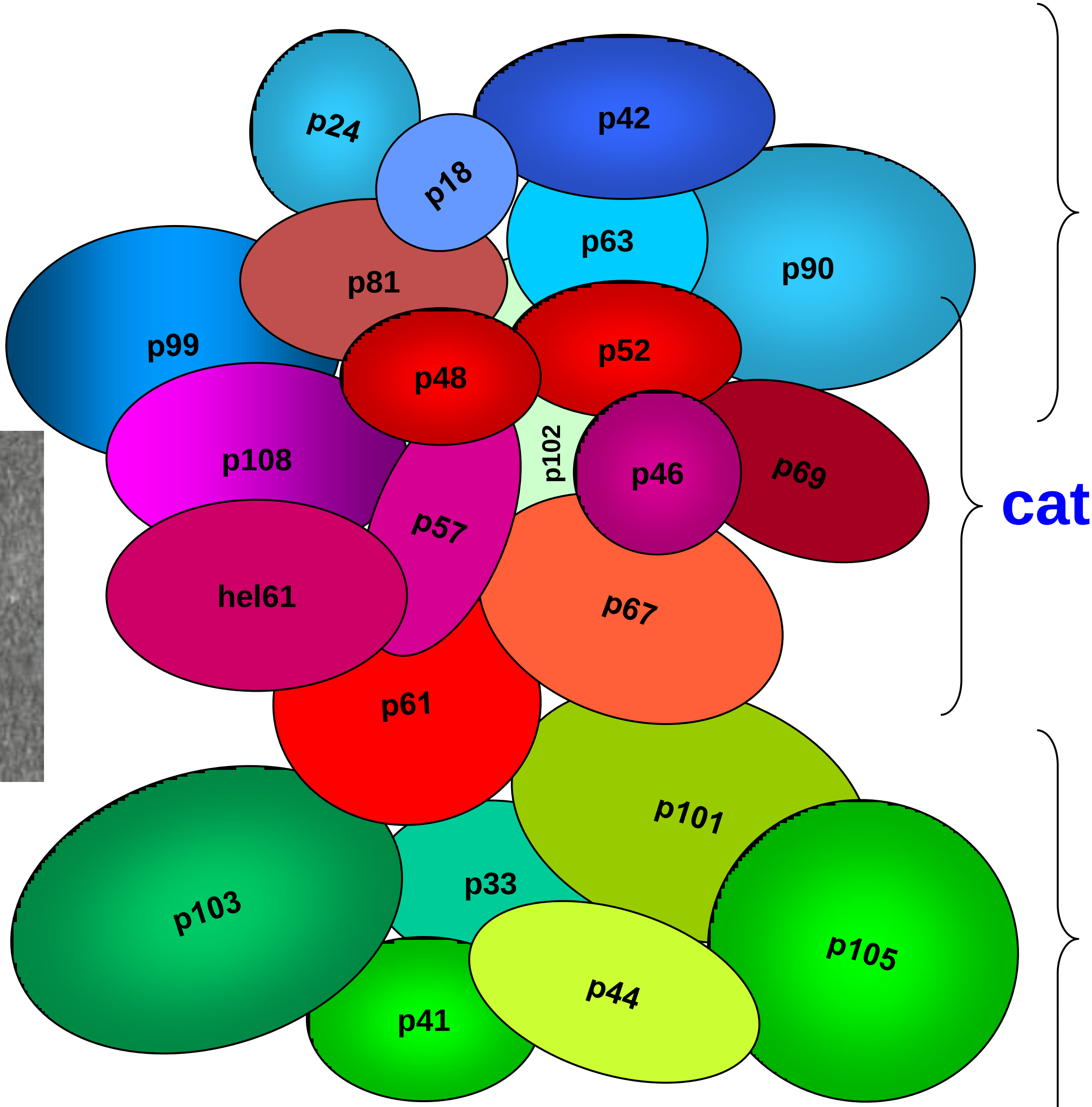
- ~1.6 MDa, 20S
- Contains 20 proteins, many exist in pairs
- Catalyzes insertion and deletion editing (endonuclease, exoUase, TUTase, ligase)
- Some catalysts are identical (TUTase, ligase, helicase)
- Likely accessory and regulatory proteins (RNA binding, RNA translocation, protein interactions, cofactor binding)

Assembly and activity do not need mRNA or gRNA

Editosome Model



EM



upstream binding

catalytic

anchor binding

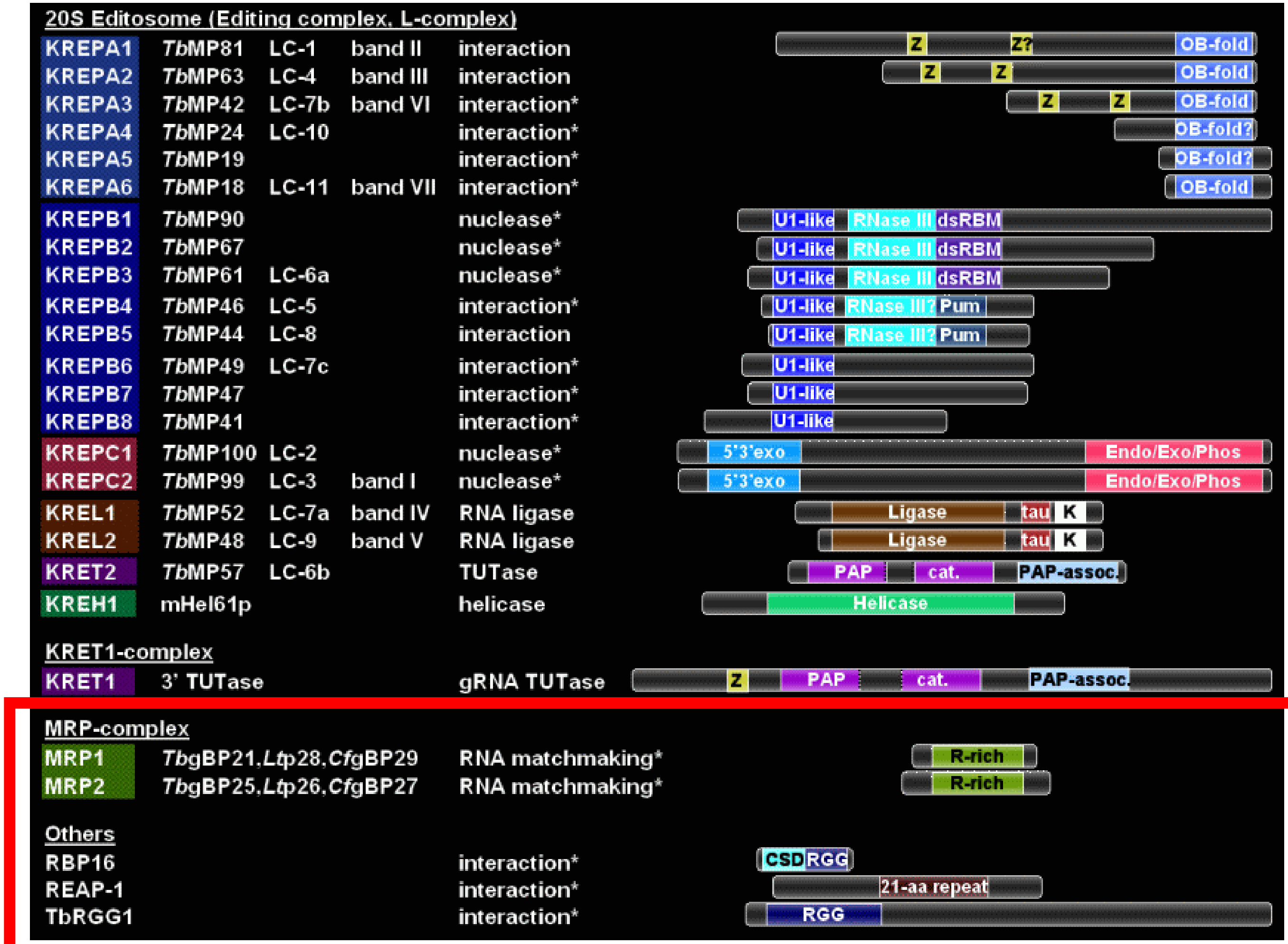
Insertion editing

Deletion editing

RNA editing

The core machinery: the 20 S editosome

and accessory factors



20S Editosome (Editing complex, L-complex)

KREPA1	*Tb*MP81	LC-1	band II	interaction	Z, Z?, OB-fold
KREPA2	*Tb*MP63	LC-4	band III	interaction	Z, Z, OB-fold
KREPA3	*Tb*MP42	LC-7b	band VI	interaction*	Z, Z, OB-fold
KREPA4	*Tb*MP24	LC-10		interaction*	OB-fold?
KREPA5	*Tb*MP19			interaction*	OB-fold?
KREPA6	*Tb*MP18	LC-11	band VII	interaction*	OB-fold
KREPB1	*Tb*MP90			nuclease*	U1-like, RNase III, dsRBM
KREPB2	*Tb*MP67			nuclease*	U1-like, RNase III, dsRBM
KREPB3	*Tb*MP61	LC-6a		nuclease*	U1-like, RNase III, dsRBM
KREPB4	*Tb*MP46	LC-5		interaction*	U1-like, RNase III?, Pum
KREPB5	*Tb*MP44	LC-8		interaction	U1-like, RNase III?, Pum
KREPB6	*Tb*MP49	LC-7c		interaction*	U1-like
KREPB7	*Tb*MP47			interaction*	U1-like
KREPB8	*Tb*MP41			interaction*	U1-like
KREPC1	*Tb*MP100	LC-2		nuclease*	5'3'exo, Endo/Exo/Phos
KREPC2	*Tb*MP99	LC-3	band I	nuclease*	5'3'exo, Endo/Exo/Phos
KREL1	*Tb*MP52	LC-7a	band IV	RNA ligase	Ligase, tau, K
KREL2	*Tb*MP48	LC-9	band V	RNA ligase	Ligase, tau, K
KRET2	*Tb*MP57	LC-6b		TUTase	PAP, cat., PAP-assoc.
KREH1	mHel61p			helicase	Helicase

KRET1-complex

KRET1	3' TUTase	gRNA TUTase	Z, PAP, cat., PAP-assoc.

MRP-complex

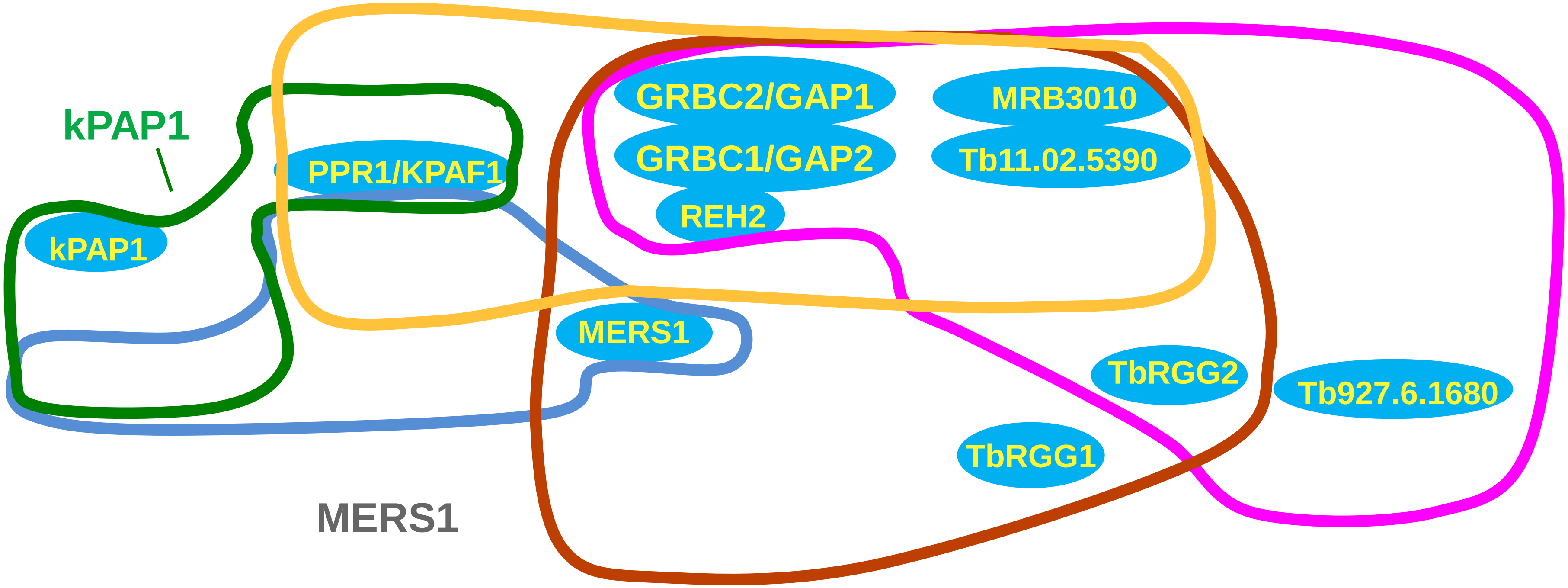
MRP1	*Tbg*BP21, *Lt*p28, *Cfg*BP29	RNA matchmaking*	R-rich
MRP2	*Tbg*BP25, *Lt*p26, *Cfg*BP27	RNA matchmaking*	R-rich

Others

RBP16	interaction*	CSD, RGG
REAP-1	interaction*	21-aa repeat
TbRGG1	interaction*	RGG

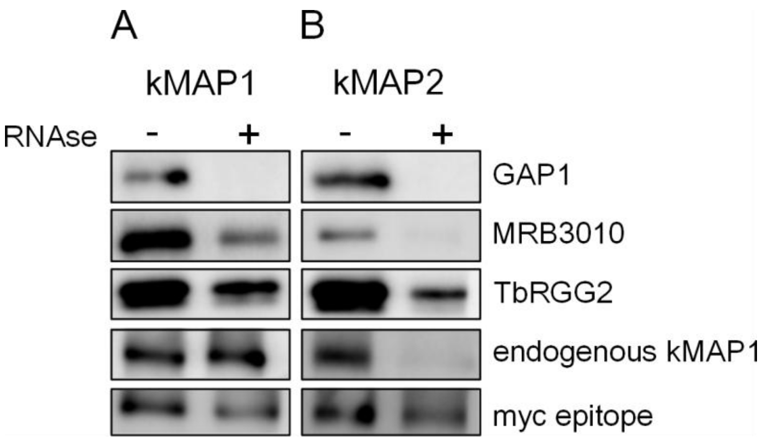
Stuart *et al.*, Trends Genet. 2005

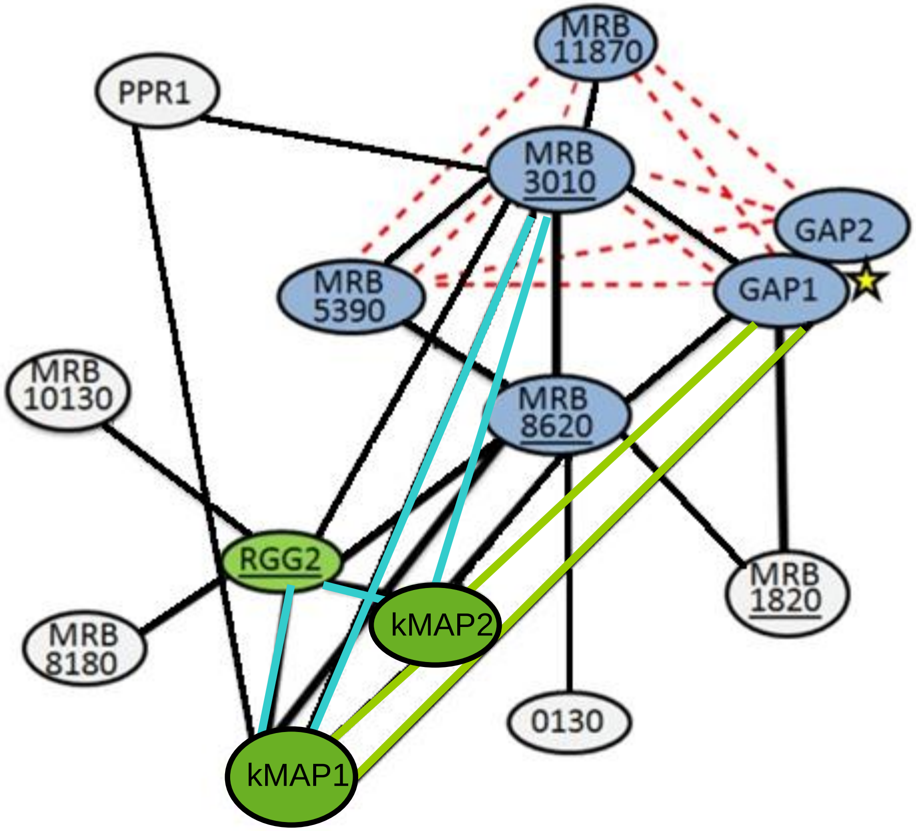
GAPs are members of novel Mitochondrial RNA binding (MRB1) complex involved in RNA biogenesis, identified in parallel by three labs



MRB1 (Hashimi, *et al.* 2008)
MRB1 (Panigrahi, *et al.* 2009)
GRBC (Weng, *et al.* 2008)
MERS1 (Weng, *et al.* 2008)
kPAP1 (Etheridge, *et al.* 2008;
Weng, *et al.* 2008)

Both kMAPs are incorporated in equivalent ribonuclear particles



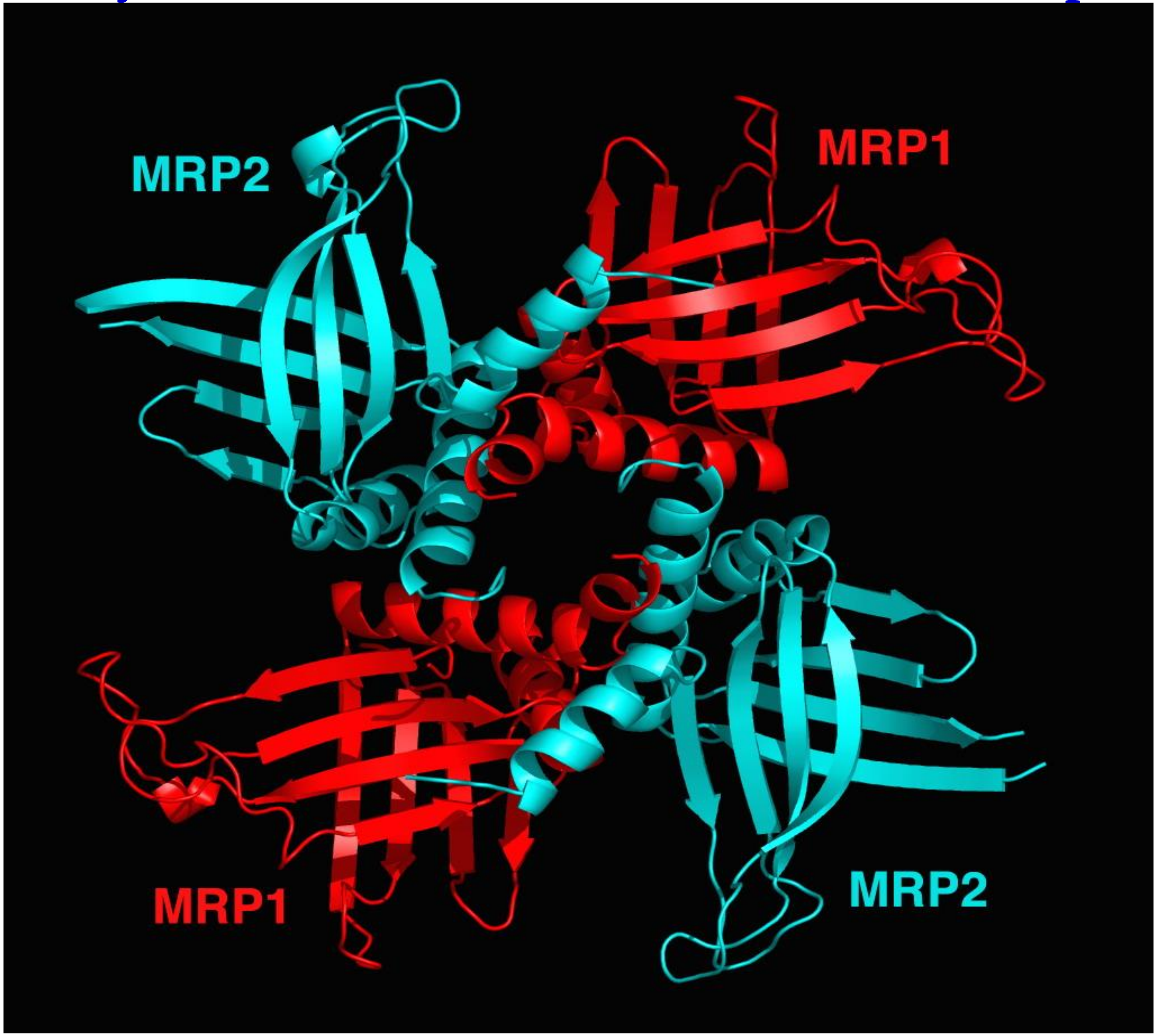


● RNA-dependent interaction

● RNA-enhanced interaction

Ammerman *et al.*, NAR 2012

Crystal structure of the MRP1/MRP2 complex



Ribbon diagram with MRP1 in red (subunits A and C) and MRP2 in cyan (subunits B and D)

Schumacher *et al*., Cell 2006

Proposed „advantages” of editing

1. **Extra level of regulation** – why not if it is already there, but as a reason for its invention not very convincing

2. **Counteracting mutations** that accumulated during periods when the genome was inactive – no evidence

3. **Efficient defense against viruses and transposons** – their RNA would be made harmless by editing; pure speculation - no evidence

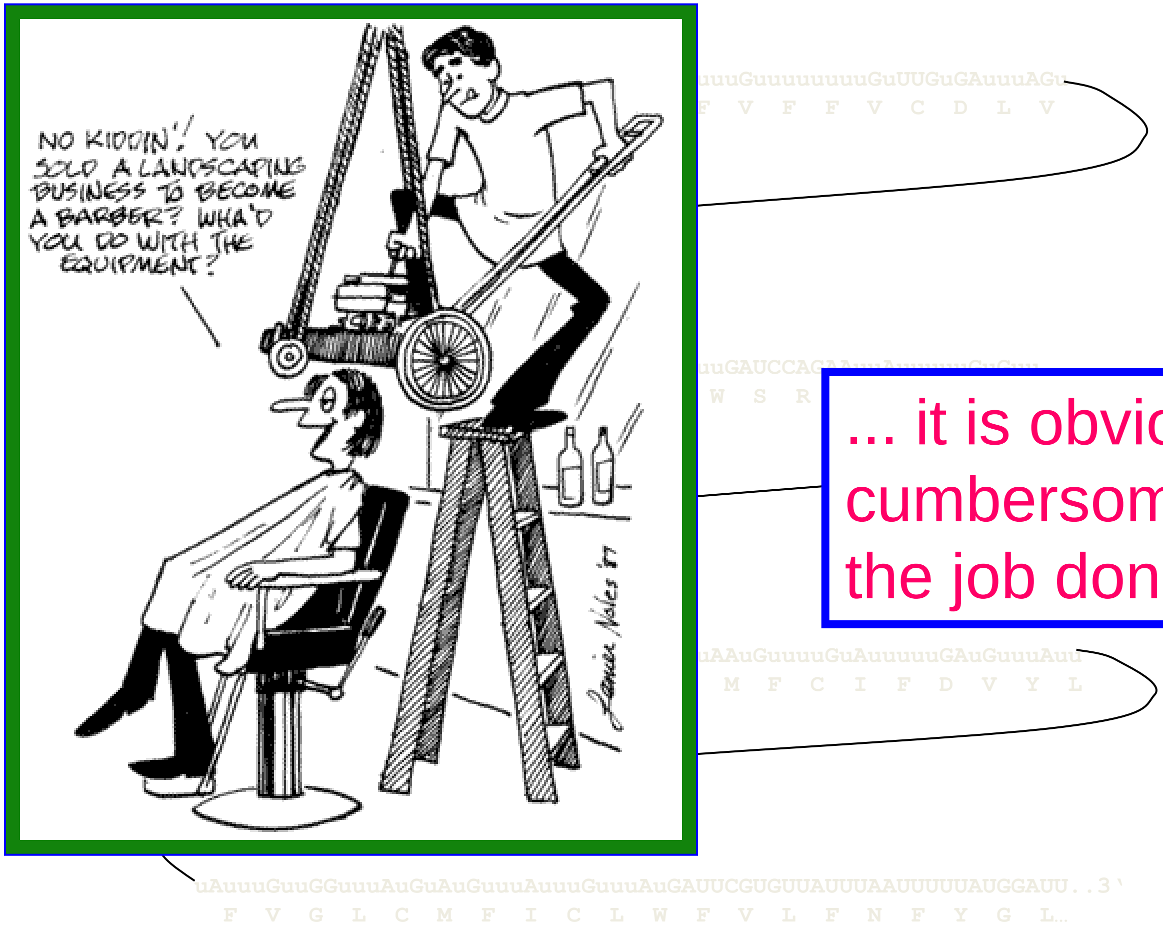
4. **Accelerating evolution by increasing the rate of genetic variation** – edited sequences may indeed evolve faster; rather a consequence than the reason for invention

----

5. **Multiple proteins encoded by one gene** – viruses and trypanosomes

6. **Advantage of gene fragmentation** - protecting them from during life stages of fierce competition, when the gene products are not needed and would be lost

7. **Acquired resistance to certain type of mutations that accummulated throught time**; complexity for complexity



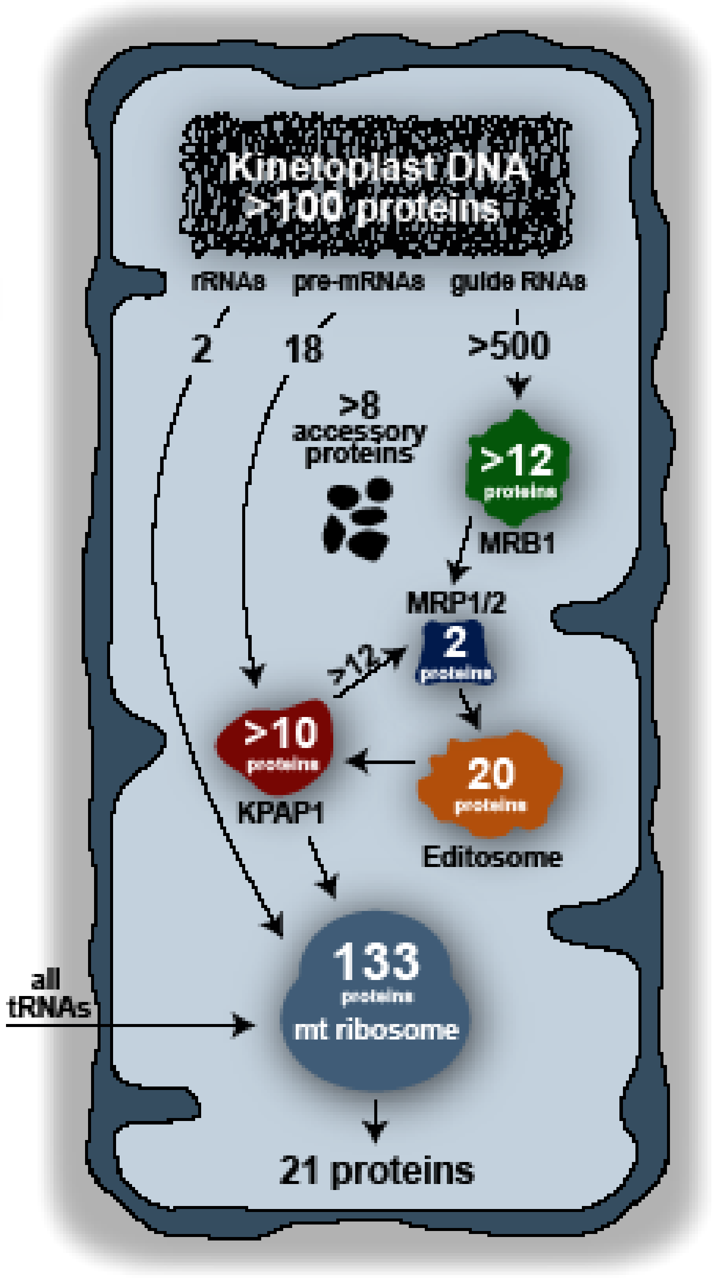
... it is obviously a rather cumbersome way to have the job done

Uncertain origin and maintenance:
Our favorite explanation is "Irremediable complexity"

Gray et al., *Science*, 2010

Current view at mitochondrial transcription, editing and processing in *Trypanosoma brucei*

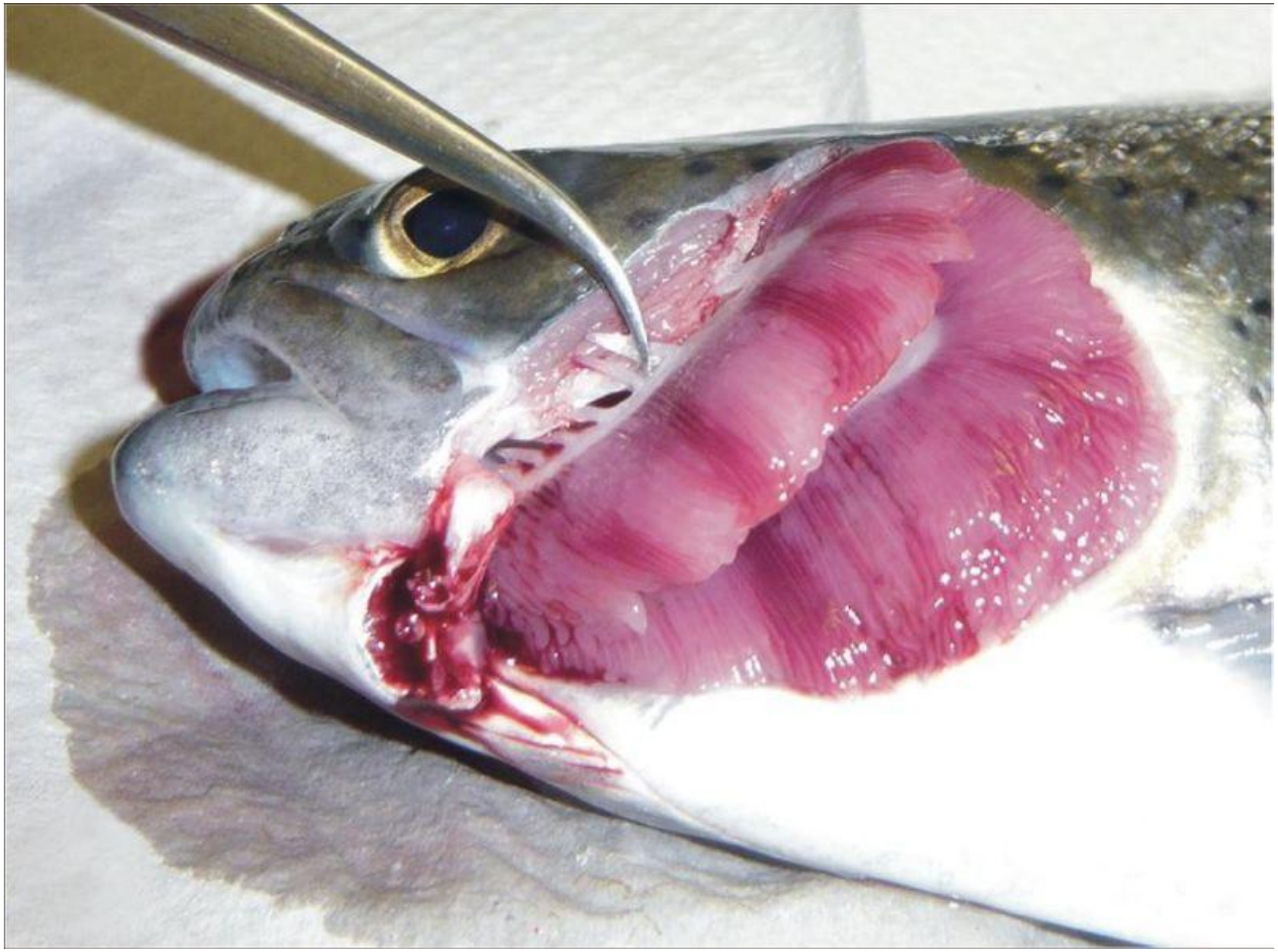
At least ~ 300 proteins are needed for the production of just 21 proteins in the procyclic stage !
And just 1 protein in the bloodstream stage!!!



Lukeš et al., *IUBMB Life*, 2011

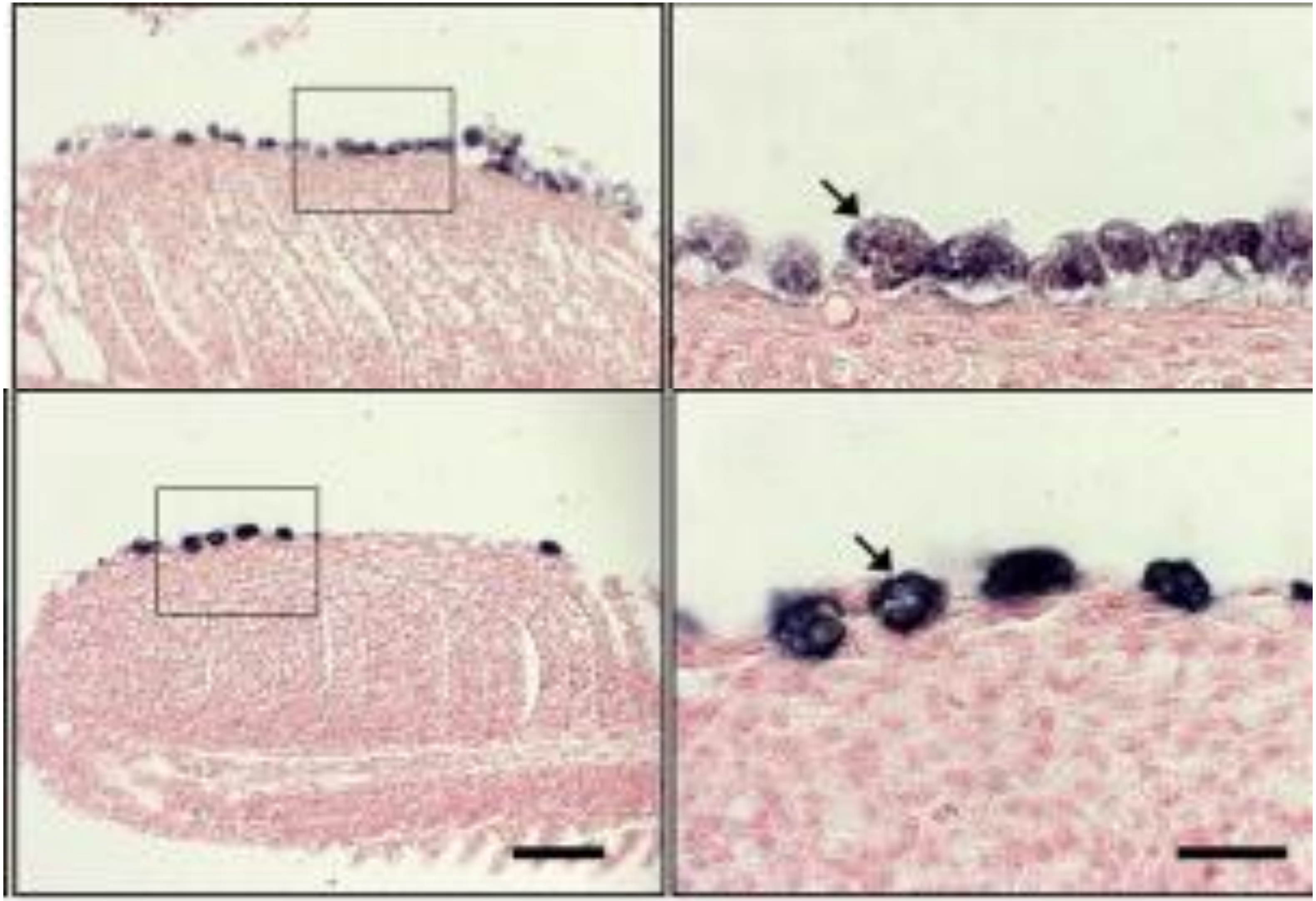
The strange life of the kinetoplastid *Perkinsela*

Infected fish with *Neoparamoeba* sp. - amoebic gill disease



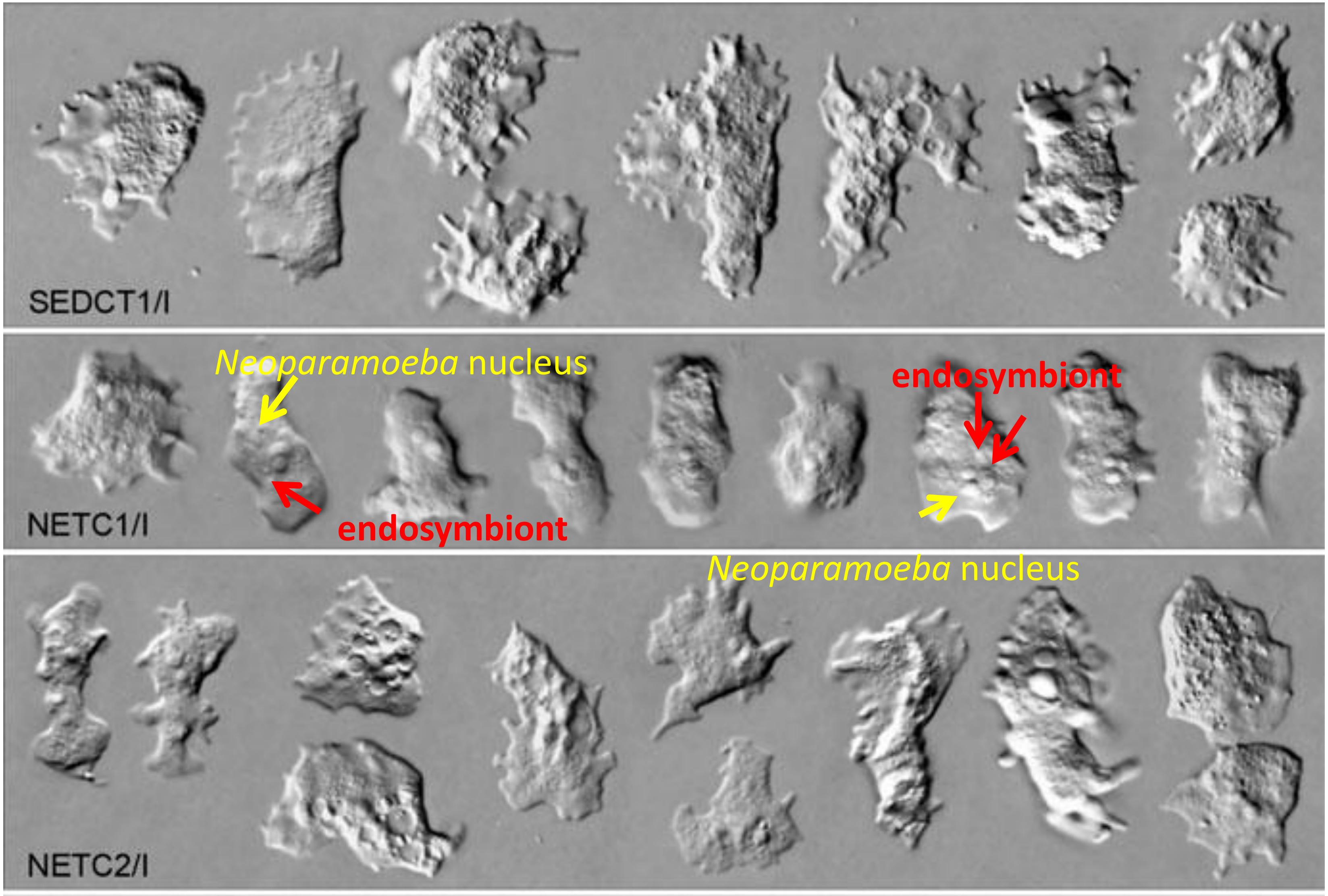
Neoparamoeba perurans

an agent of amoebic gill disease of Atlantic salmon (*Salmo salar*)



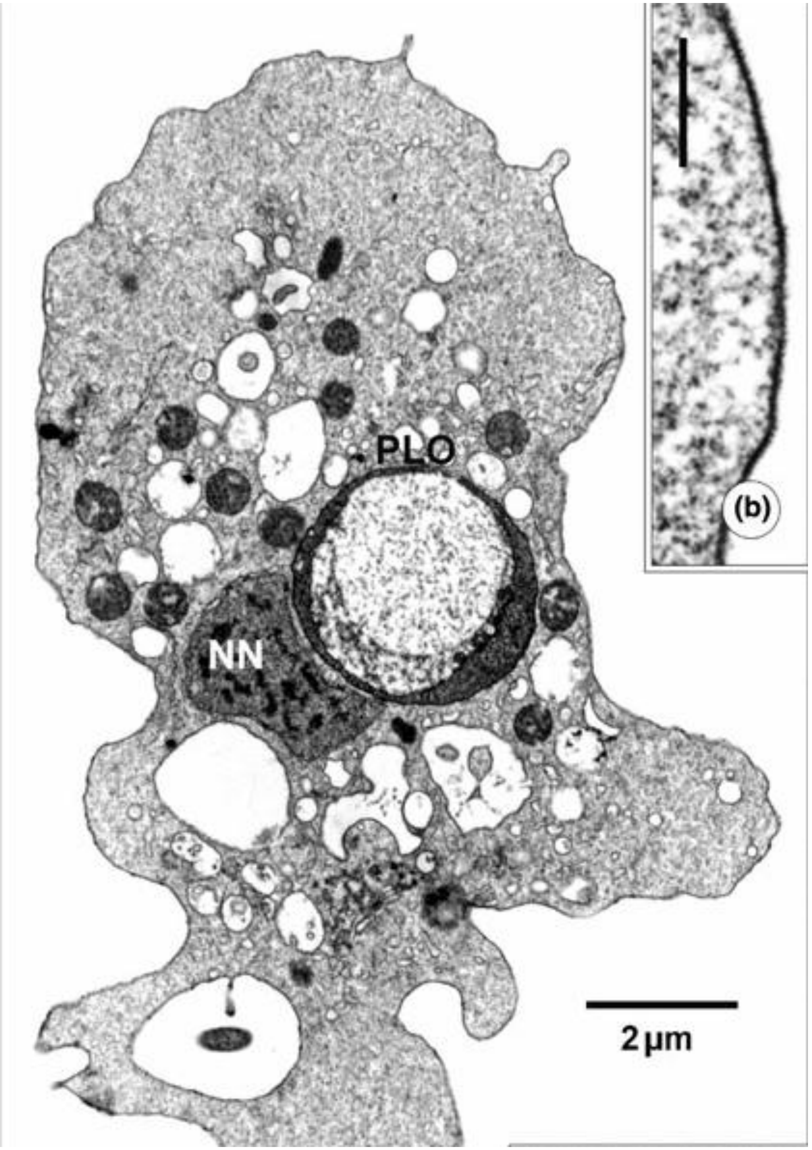
Young et al. IJP 2007

Trophozoites representing clonal cultures of *Neoparamoeba* strains

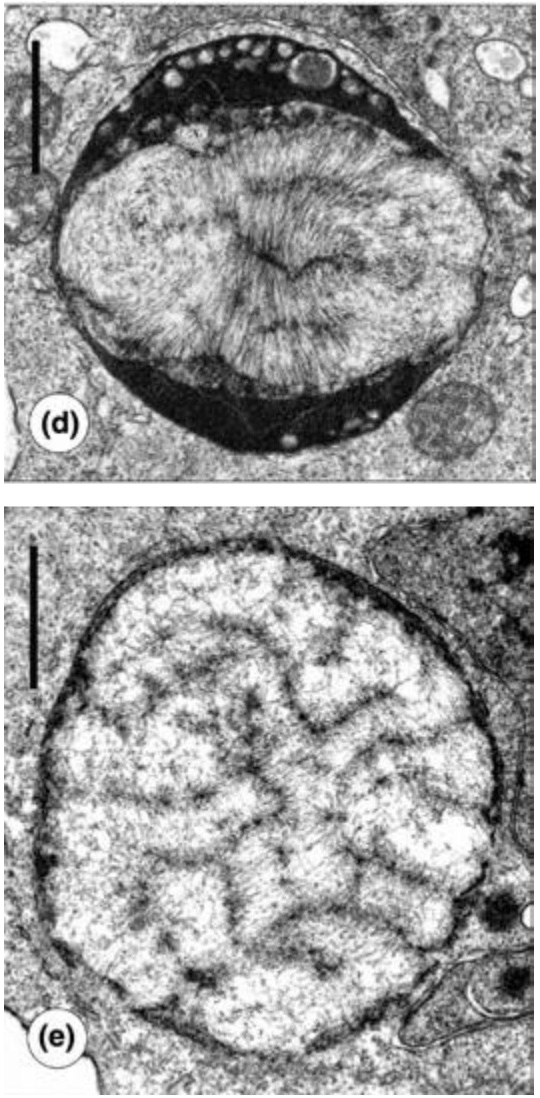


Dyková et al. J.Fish Dis. 2005

Ultrastructure of *Neoparamoeba* with “its” *Perkinsela*



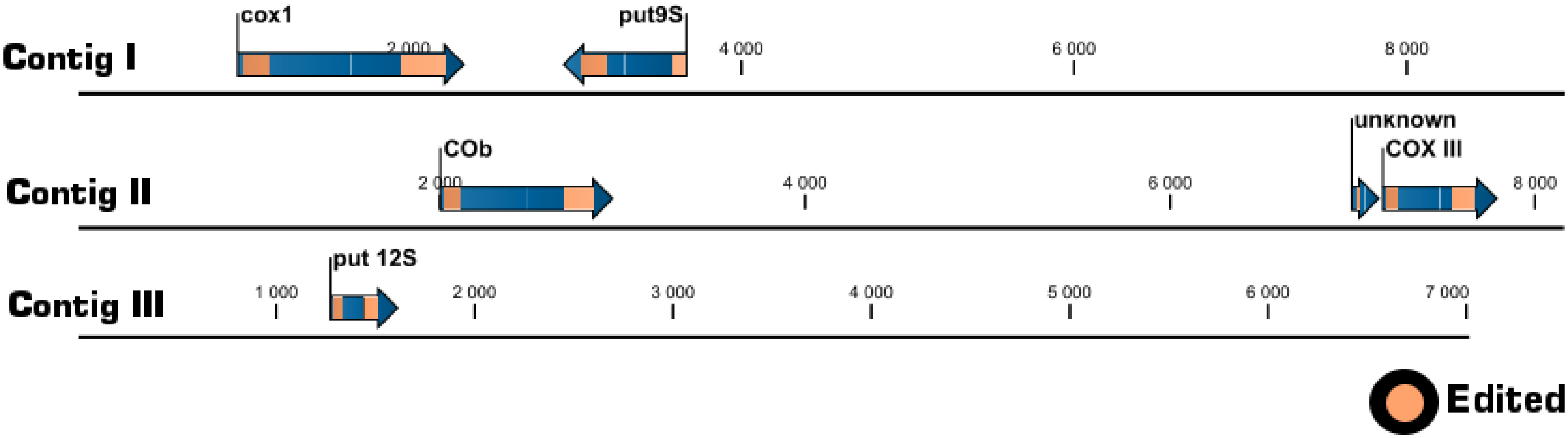
Trophozoite, NN – amoeba nucleus, PLO – *Perkinsela amoebae*-like endosymbiont

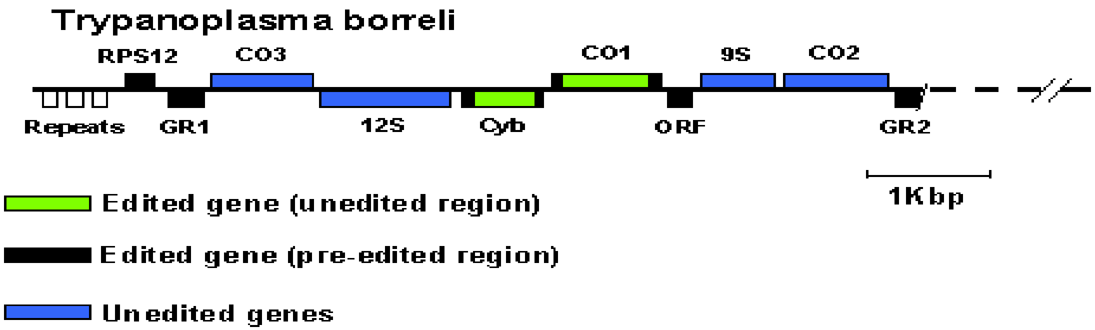


(d, e) Endosymbionts with a huge kinetoplast (bar = 1 µm)

Dyková et al. Europ. Protistol. 2003; J. Fish Dis. 2005

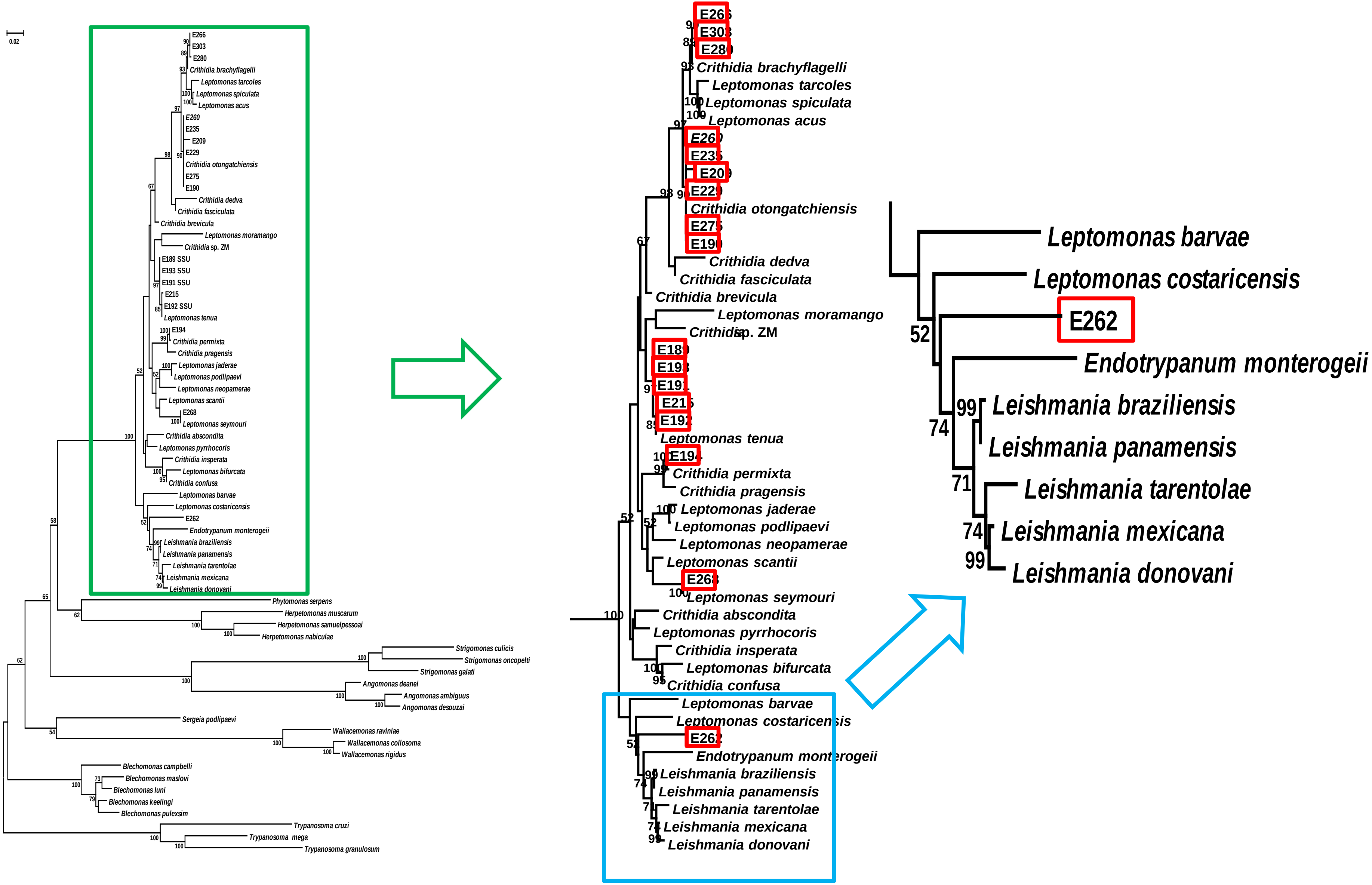
We have sequenced the genomes of *Neoparamoeba* and *Perkinsela*, and assembled its kDNA into contigs



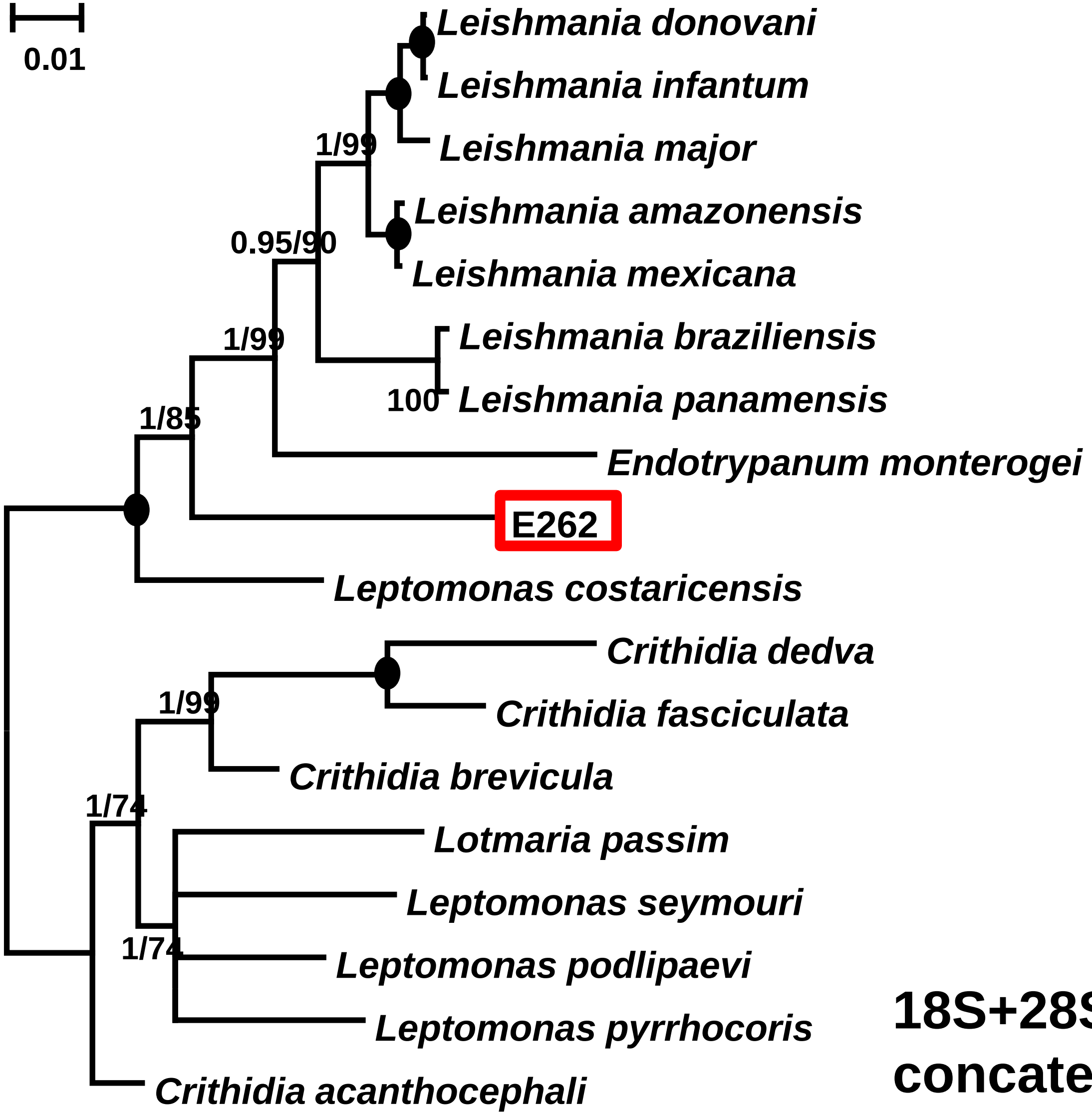


http://dna.kdna.ucla.edu/trypanosome/images/tbo1.GIF

Exploration of diversity in Ecuador resulted in finding *Novymonas*, a new genus of endosymbiont-containing trypanosomatids closely related to *Leishmania*

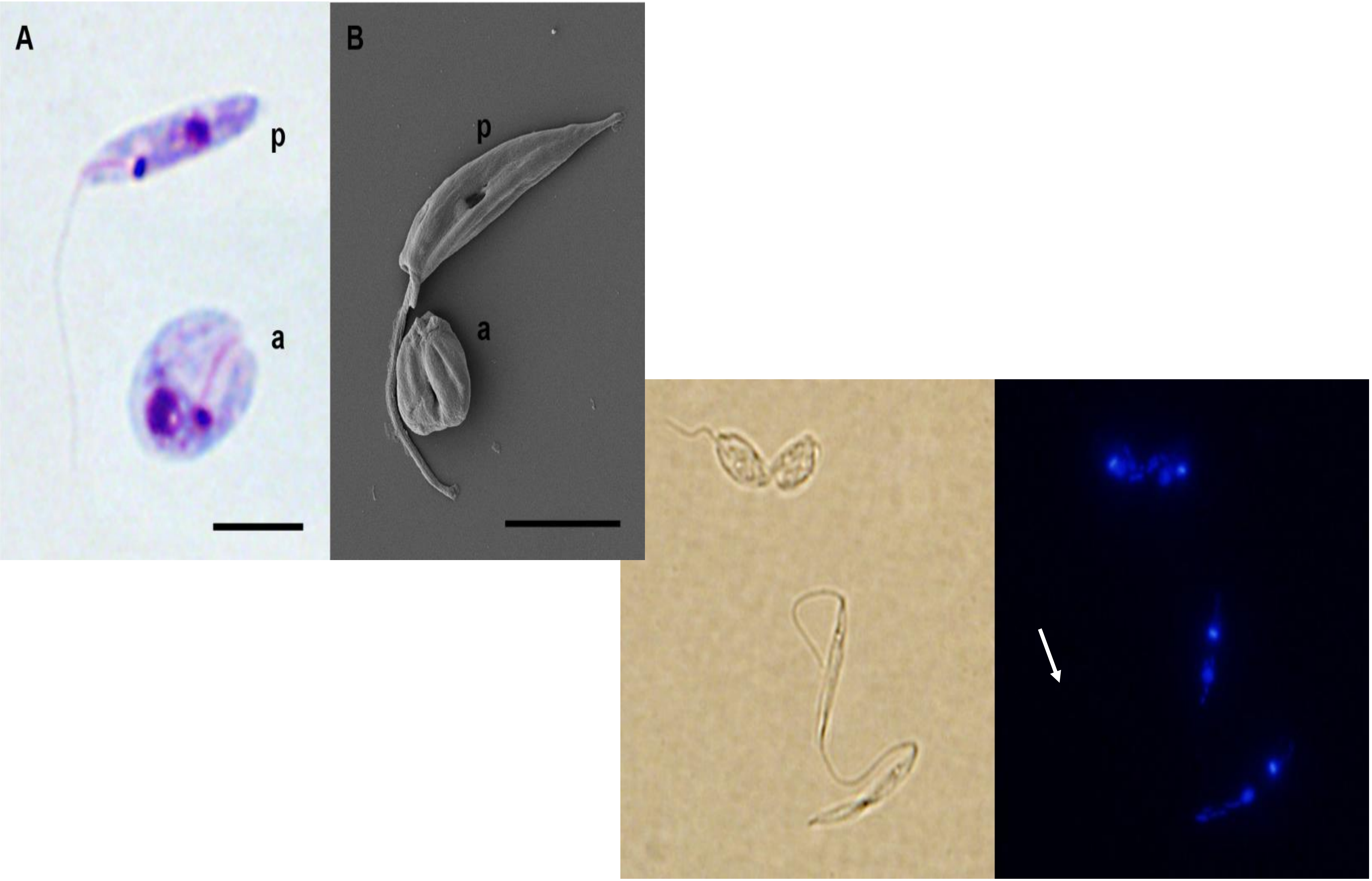


Phylogenetic position of the E262 isolate

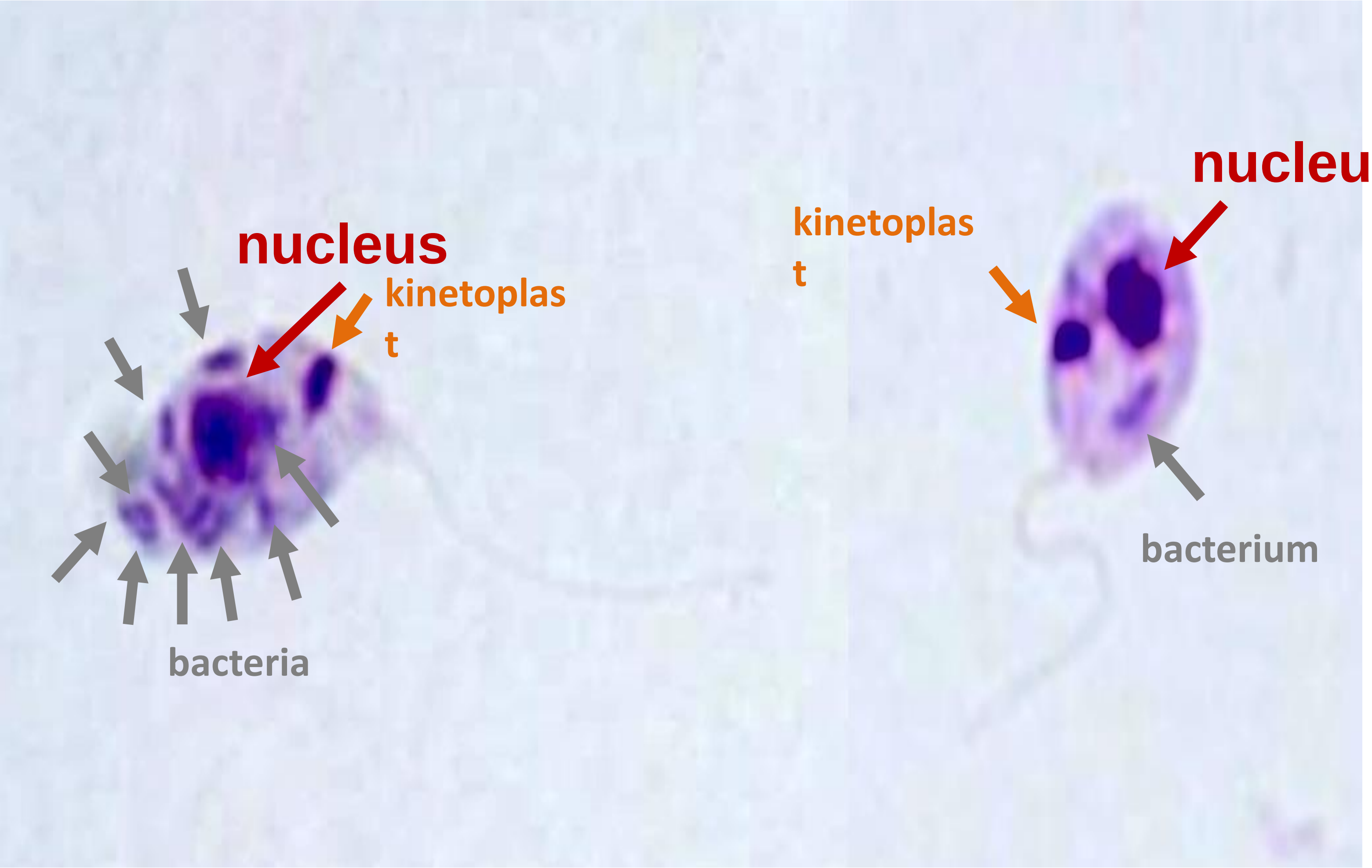


18S+28S+HSP83 concatenated tree

E262 has amastigotes but even more importantly, an endosymbionts

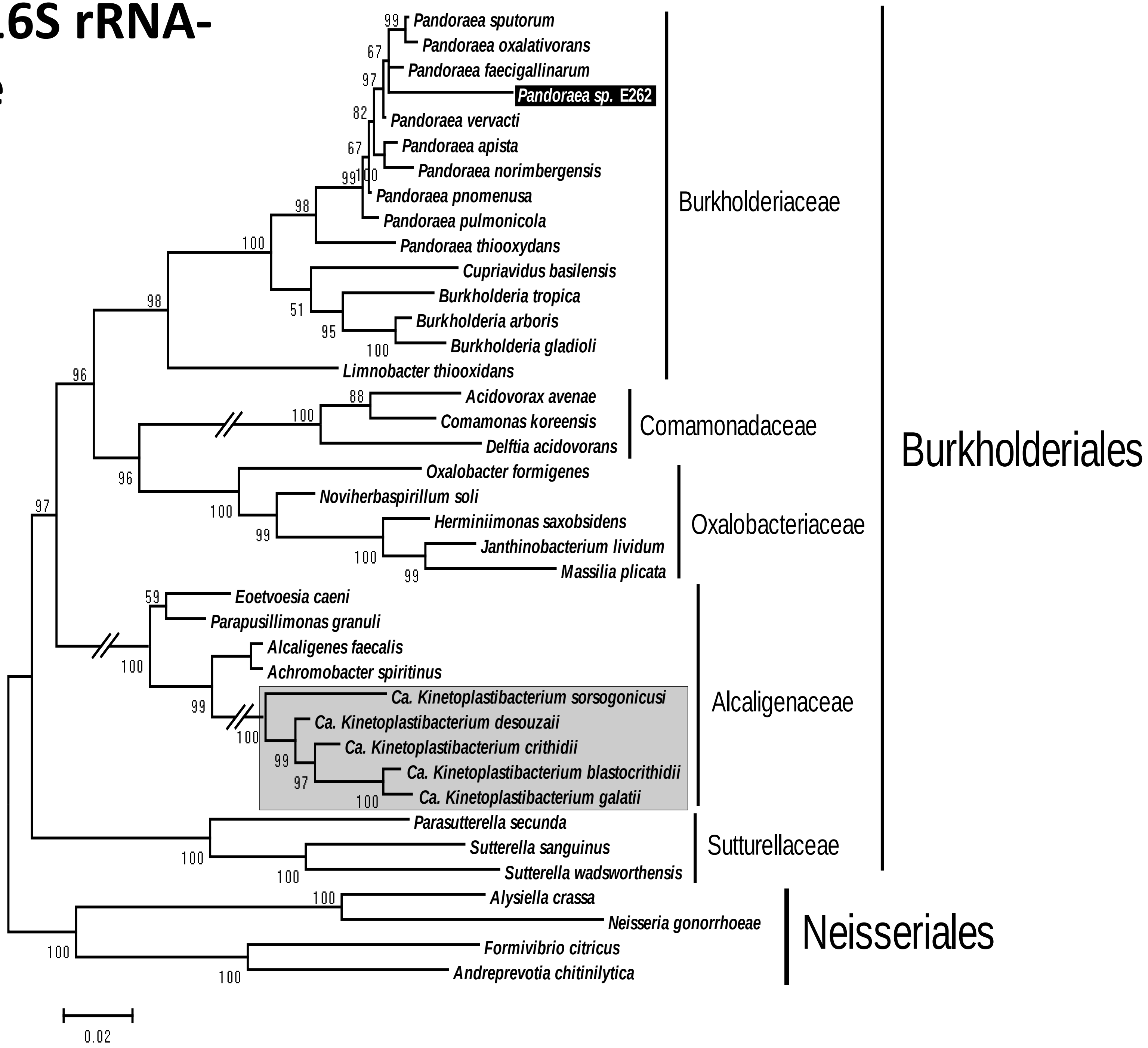


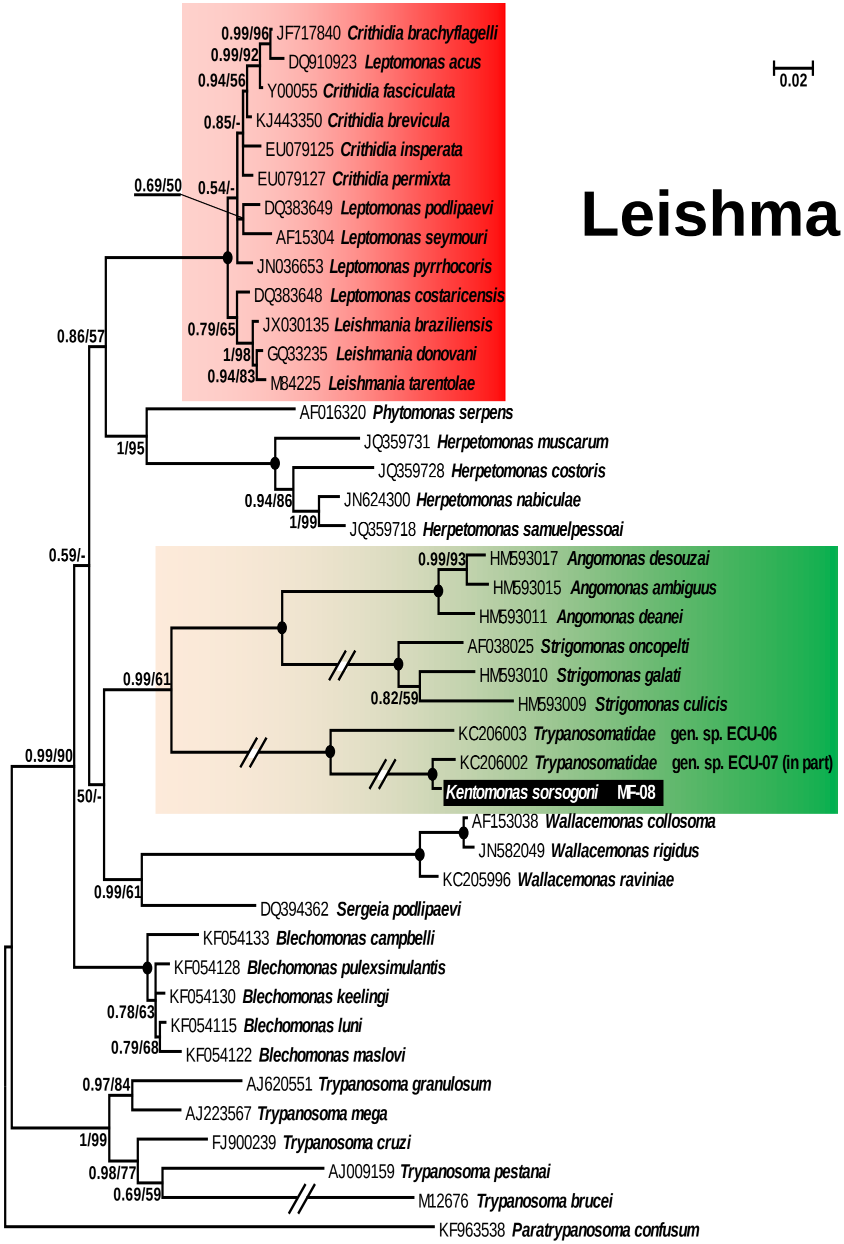
Endosymbionts: Giemsa staining



The number of bacteria in the trypanosomatid cell is inconstant!

Bayesian 16S rRNA-based tree





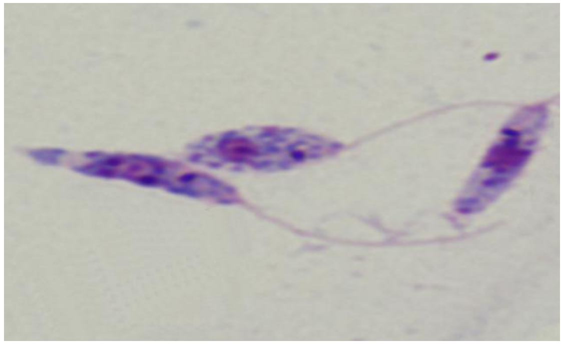
Leishmaniinae

Strigomonadinae

Votýpka et al., Protist 2014

The case of *Paratrypanosoma* *"confusa"*

- isolated from *Culex pipiens* in Prague (of all exotic places…?!?)
- successfully cultured
- unique phylogenetic position!



Paratrypanosoma - phylogenetic analyses

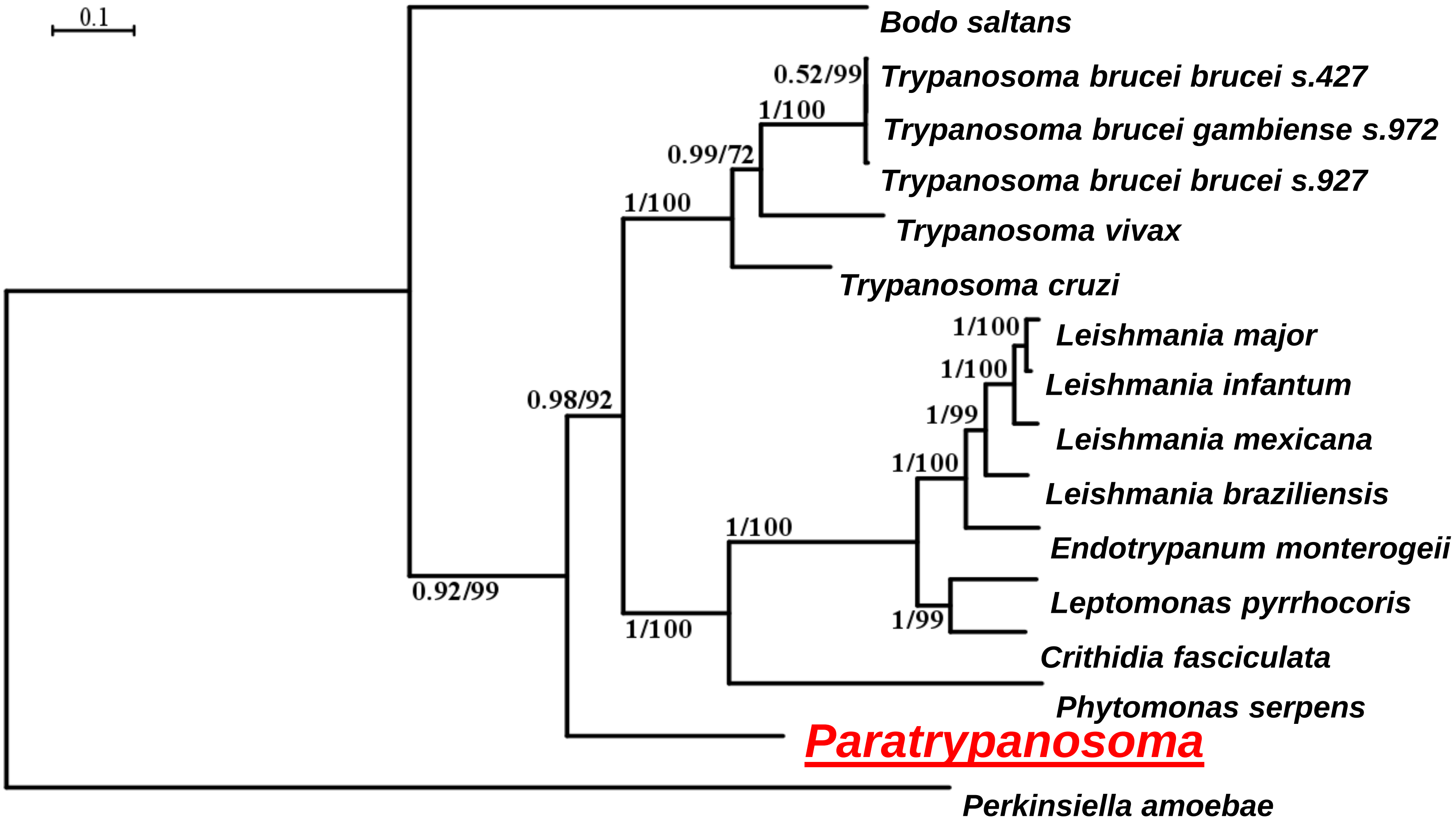
- Limits of out-groups
  - *Bodo saltans* – not enough genes identified
  - *Diplonema* – most identified genes are different from those available for *Paratrypanosoma* and *Bodo*
- *Perkinsela amoebae* (the most early-branching bodonid living as an endosymbiont in the amoebae on fish gills!) is the most suitable out-group – we are sequencing its whole genome

Analyzing datasets for a growing number of genes available for enough species (currently 32)

used RaxML program - maximum likelihood, LG substitution model (WAG and JTT also tried) and rapid bootstrapping

Paratrypanosoma -concatenated tree

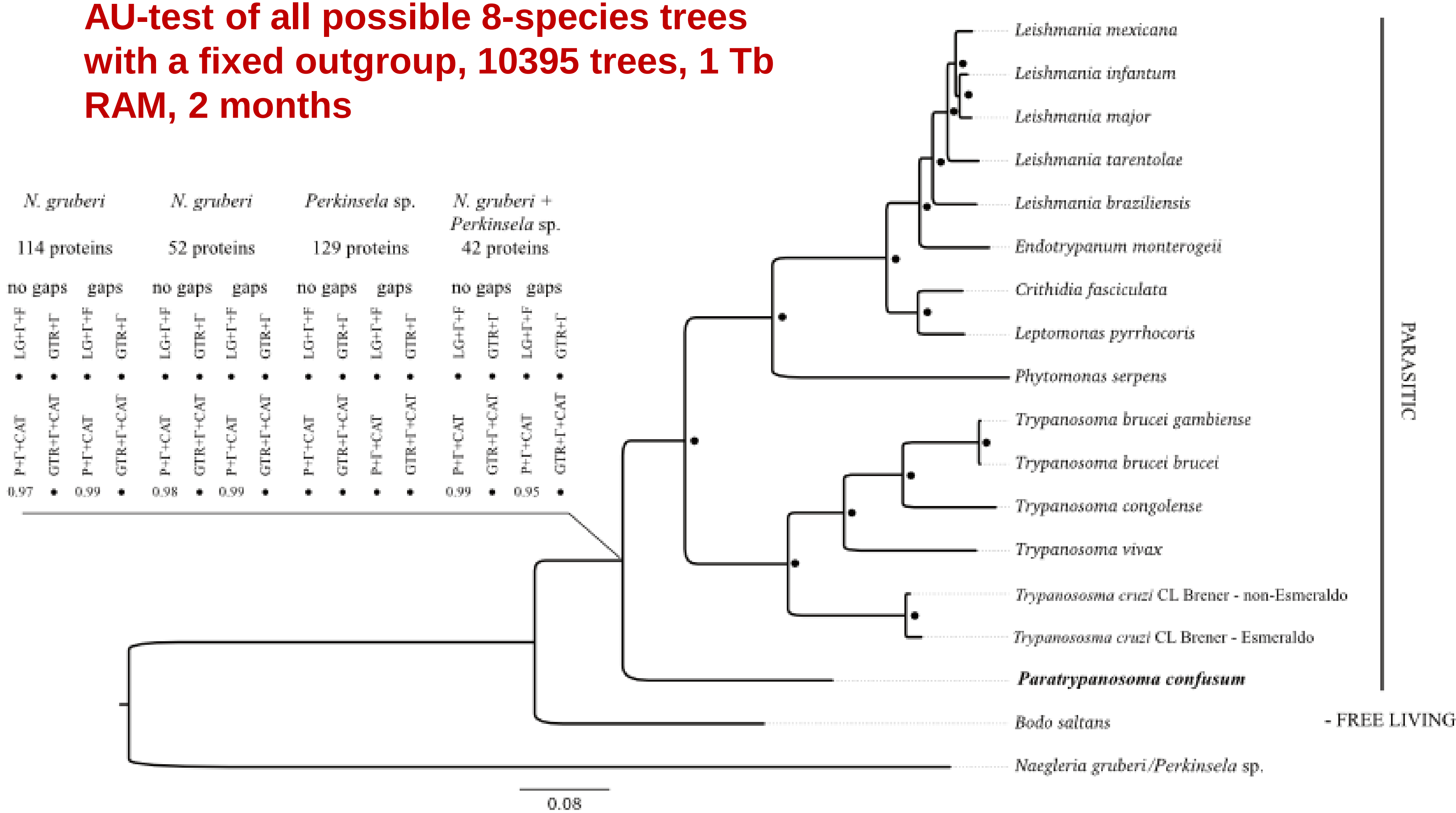
- **Composed of 16 genes and 16 species**
- used Bayesian analysis (Phylobayes) and RAxML and LG model
- maximum parsimony (PAUP) analysis also supports this branching



Flegontov et al., Curr. Biol. 2013

Basal position of *Paratrypanosoma* is unambiguous

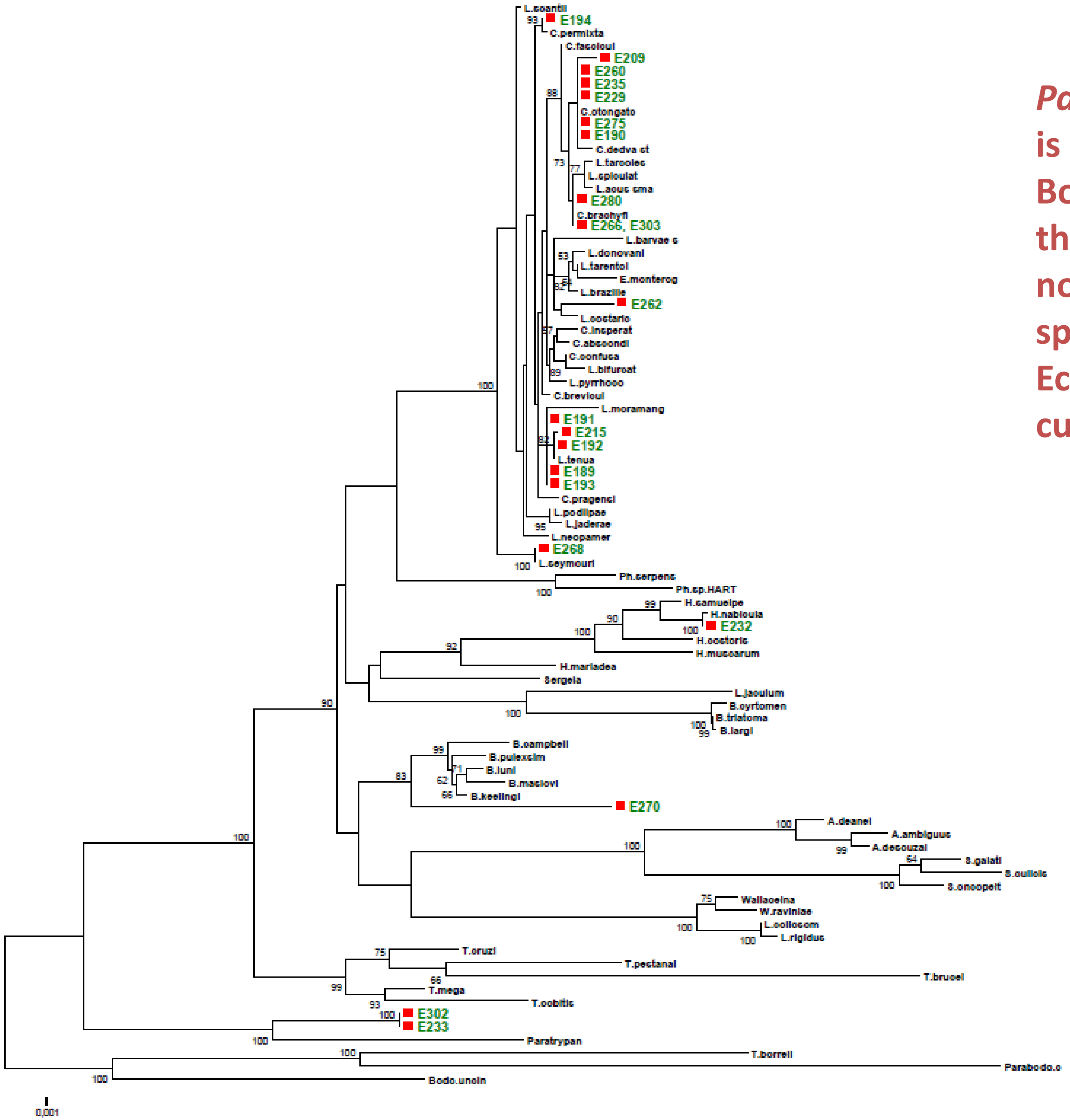
AU-test of all possible 8-species trees with a fixed outgroup, 10395 trees, 1 Tb RAM, 2 months



N. gruberi	114 proteins	no gaps	LG+Γ+F	•	P+Γ+CAT	0.97
			GTR+Γ	•	GTR+Γ+CAT	•
		gaps	LG+Γ+F	•	P+Γ+CAT	0.99
			GTR+Γ	•	GTR+Γ+CAT	•
N. gruberi	52 proteins	no gaps	LG+Γ+F	•	P+Γ+CAT	0.98
			GTR+Γ	•	GTR+Γ+CAT	•
		gaps	LG+Γ+F	•	P+Γ+CAT	0.99
			GTR+Γ	•	GTR+Γ+CAT	•
Perkinsela sp.	129 proteins	no gaps	LG+Γ+F	•	P+Γ+CAT	•
			GTR+Γ	•	GTR+Γ+CAT	•
		gaps	LG+Γ+F	•	P+Γ+CAT	•
			GTR+Γ	•	GTR+Γ+CAT	•
N. gruberi + *Perkinsela* sp.	42 proteins	no gaps	LG+Γ+F	•	P+Γ+CAT	0.99
			GTR+Γ	•	GTR+Γ+CAT	•
		gaps	LG+Γ+F	•	P+Γ+CAT	0.95
			GTR+Γ	•	GTR+Γ+CAT	•

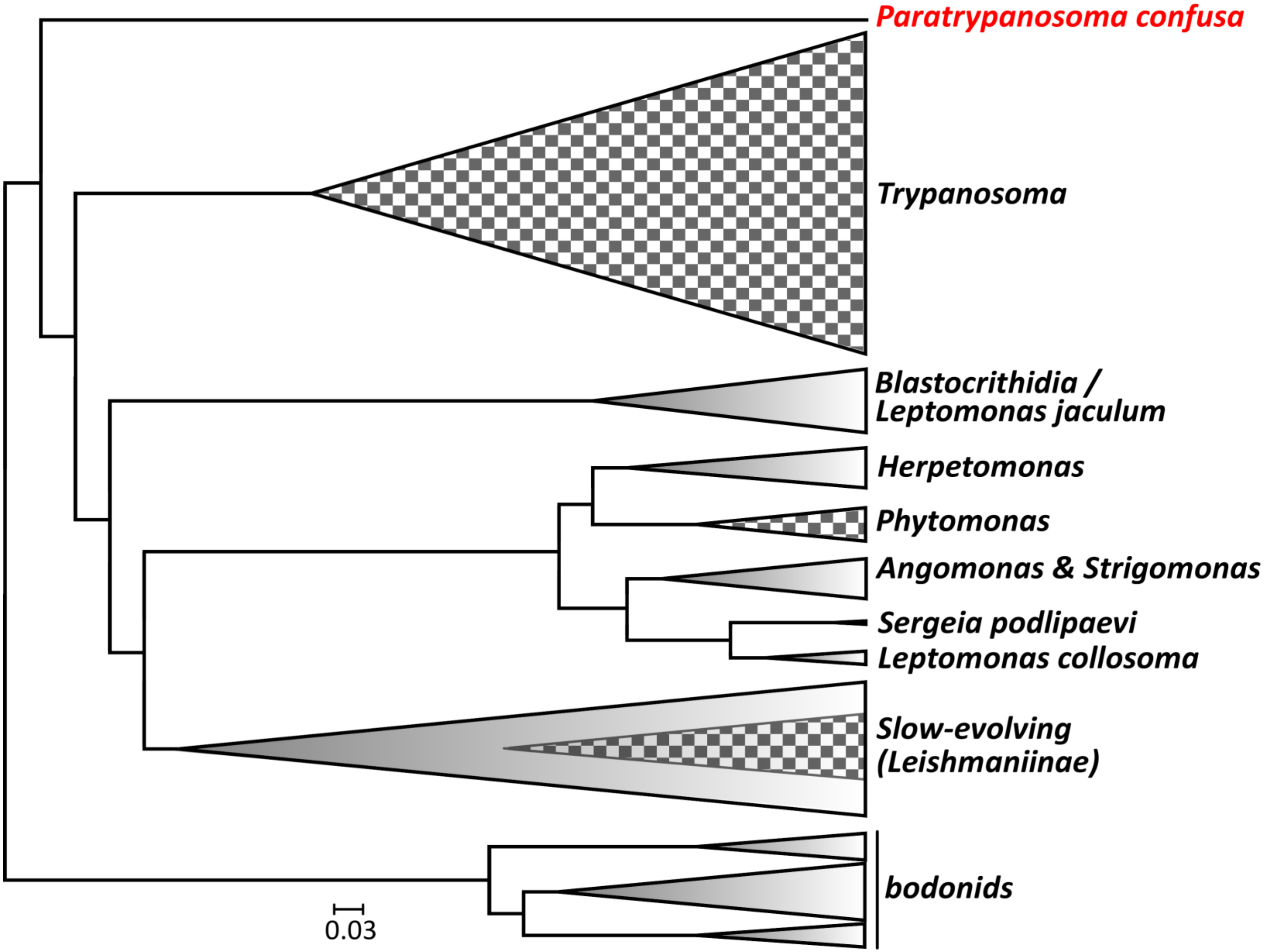
ML concatenated tree

Flegontov et al., Curr. Biol. 2013

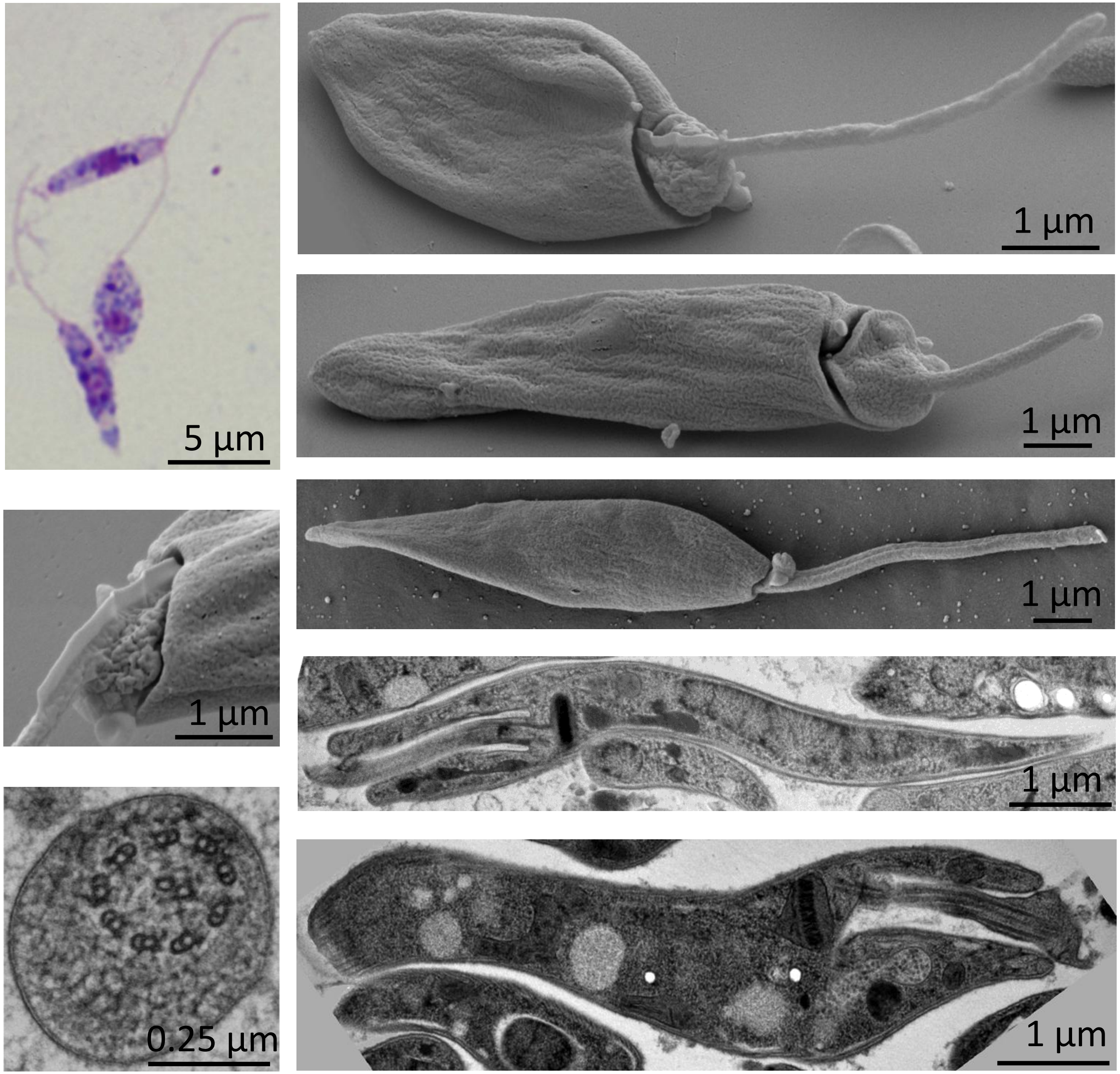


Paratrypanosoma is not only in Bohemia and the US, but we now have new species from Ecuador in culture

From now on every dixenous trypanosomatid has a monoxenous basal clade

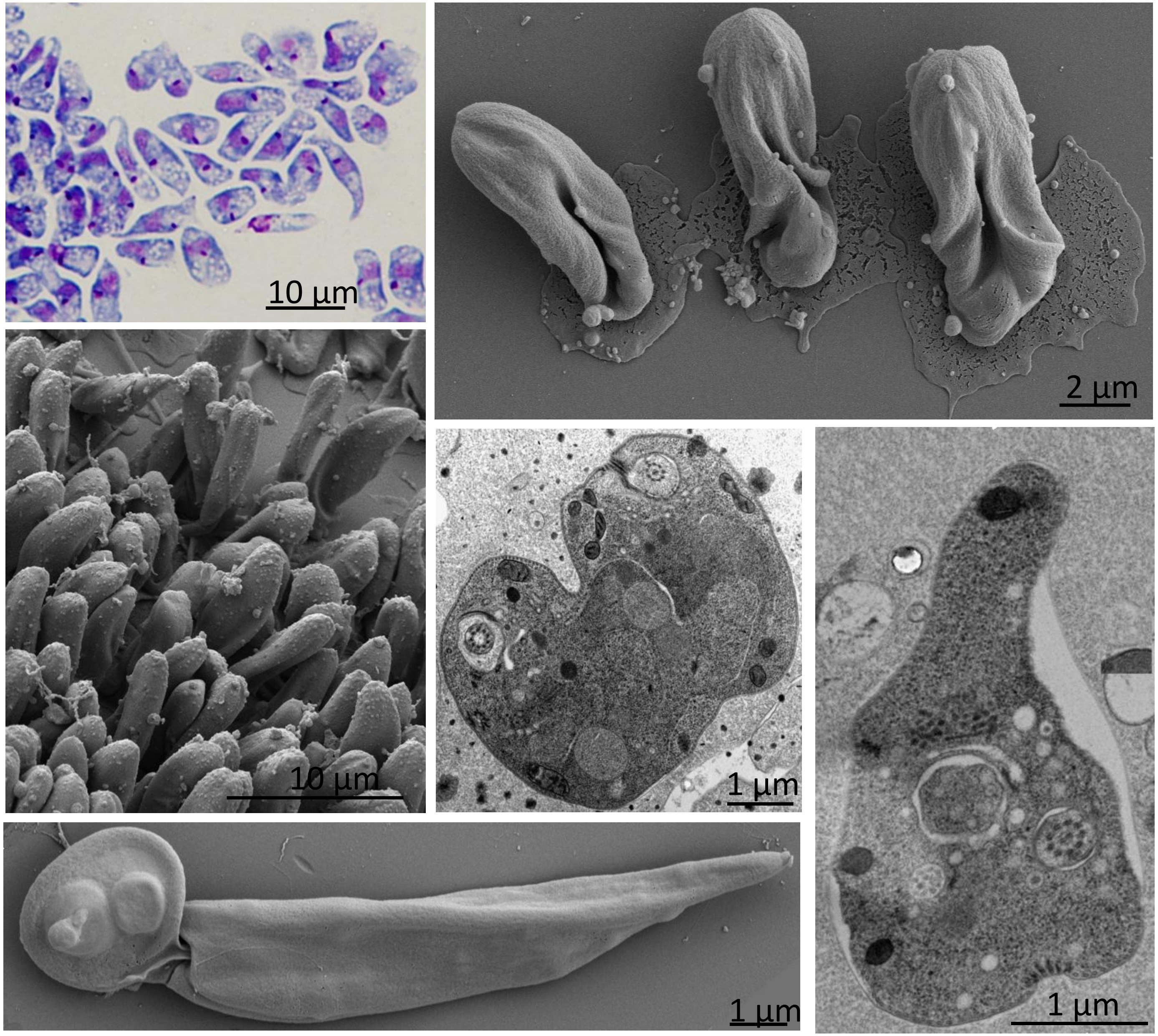


Swimmer stage of *Paratrypanosoma*

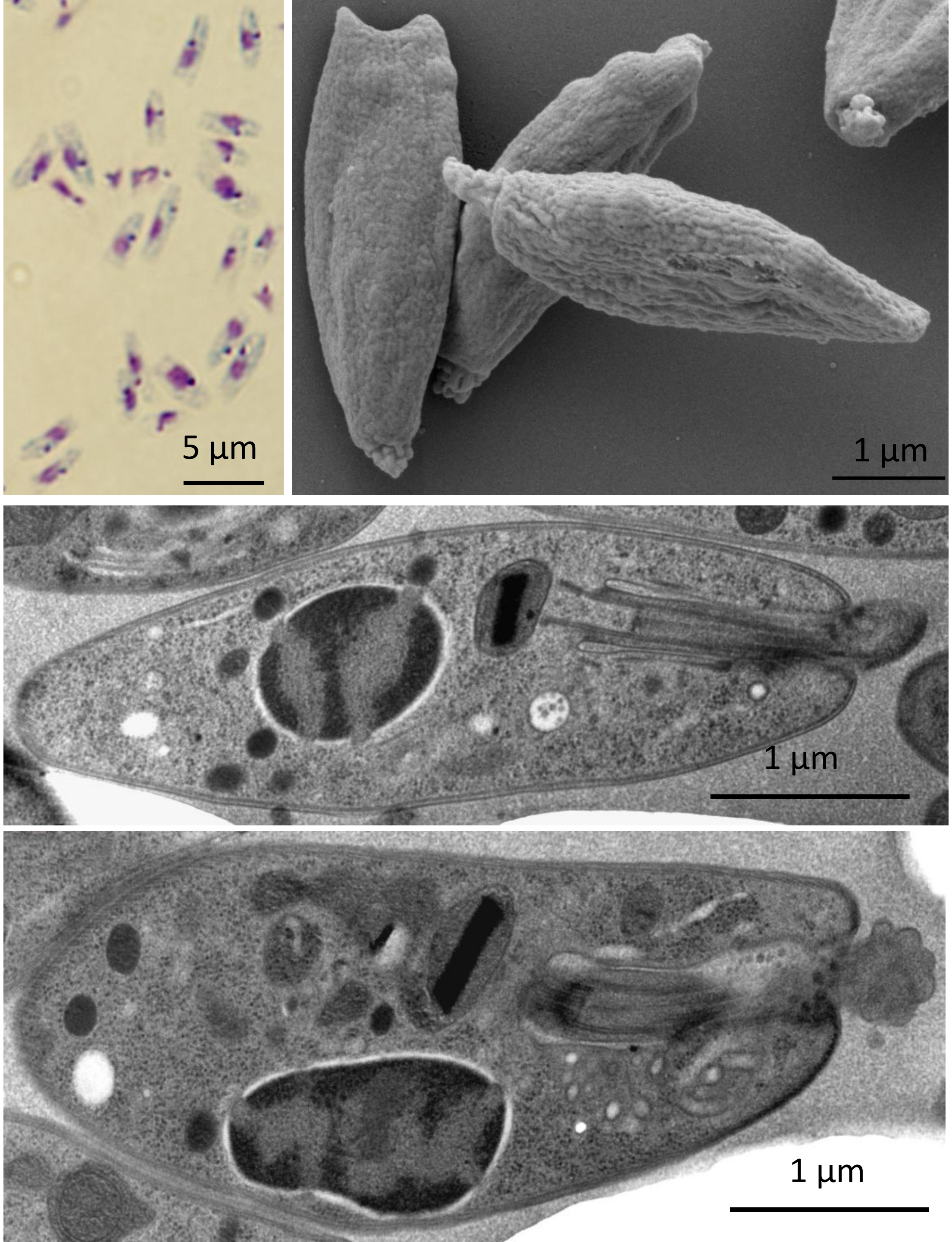


Classical-looking promastigotes, yet with an extended basal part of the flagellum and a tiny undulating membrane

Sessile stage of *Paratrypanosoma*



Stages that internalize their flagellum
and become firmly attached to the surface



Another sessile stage of *Paratrypanosoma* on agar plates

Stages that become attached to a soft surface look yet different – their end is tapered and they retain a very short flagellum

Upper: Alexander Haindrich, Jiří Týč, Hassan Hashimi, Julius Lukeš, Somsuvro Basu, Pavel Poliak, Milan Jirků, Zhenqiu Huang, Pavel Flegontov, Tomáš Skalický

Lower: Sabine Kaltenbrunner, Eva Horáková, Eva Stříbrná, Gabriela Ridvanová, Julie Kovářová, Olga Flegontova, Piya Changmai

Acknowledgements

our lab
in České Budějovice
(Budweis)
South Bohemia

Grant support

Czech Grant Agency
Czech Academy of Sciences
Czech Ministry of Education
Canadian Institute for Advanced Research