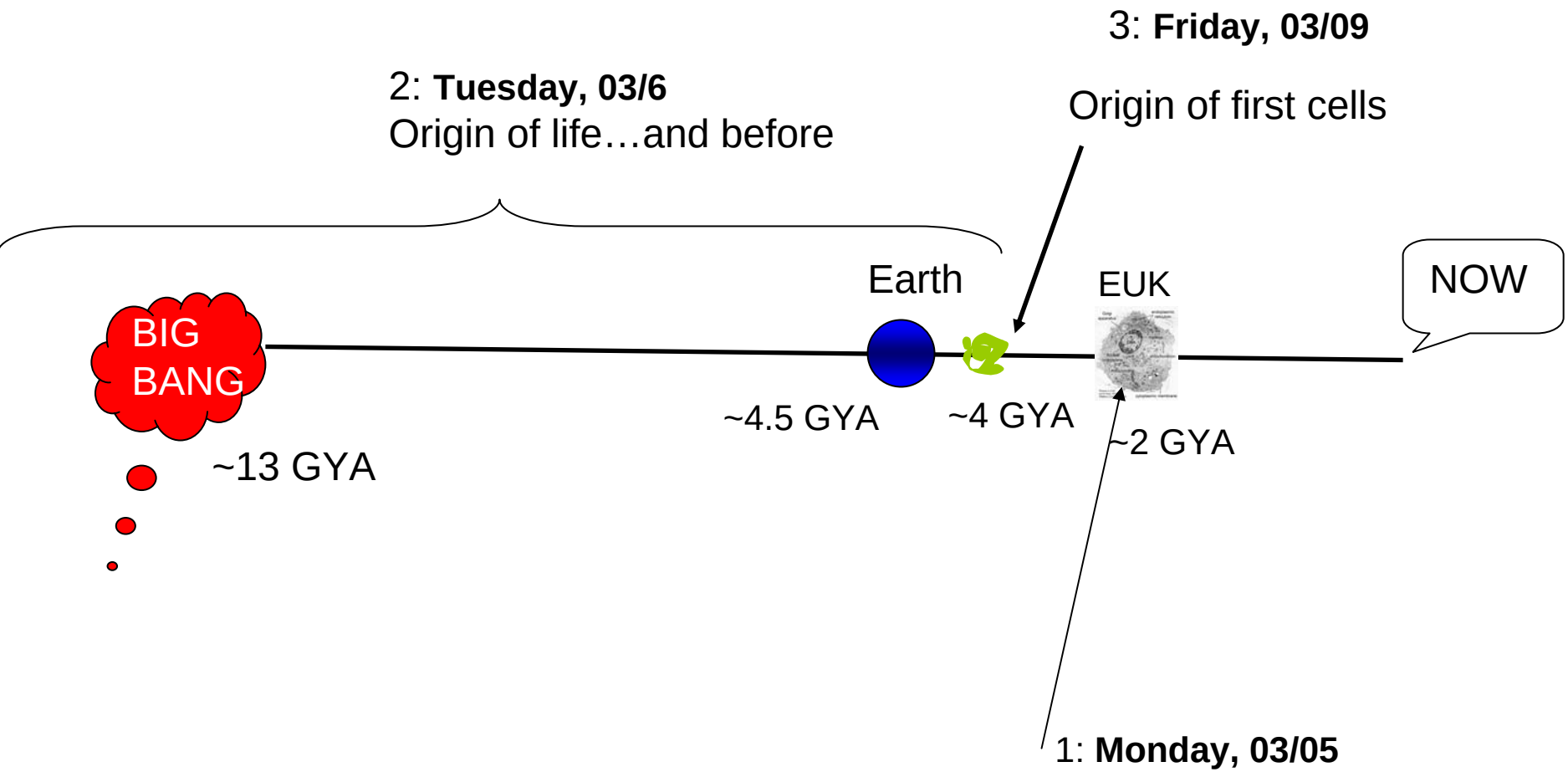
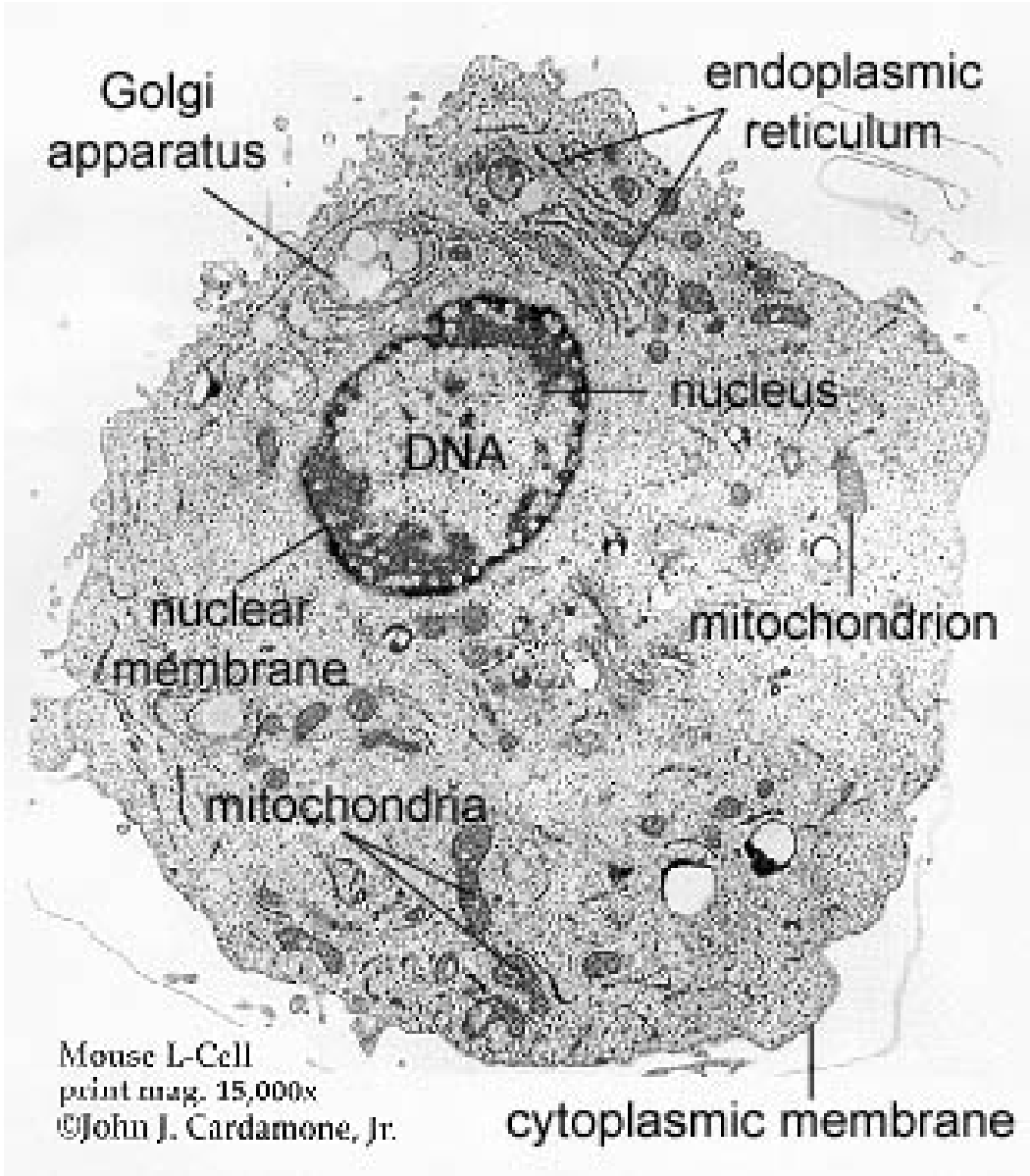
A three-part origin saga



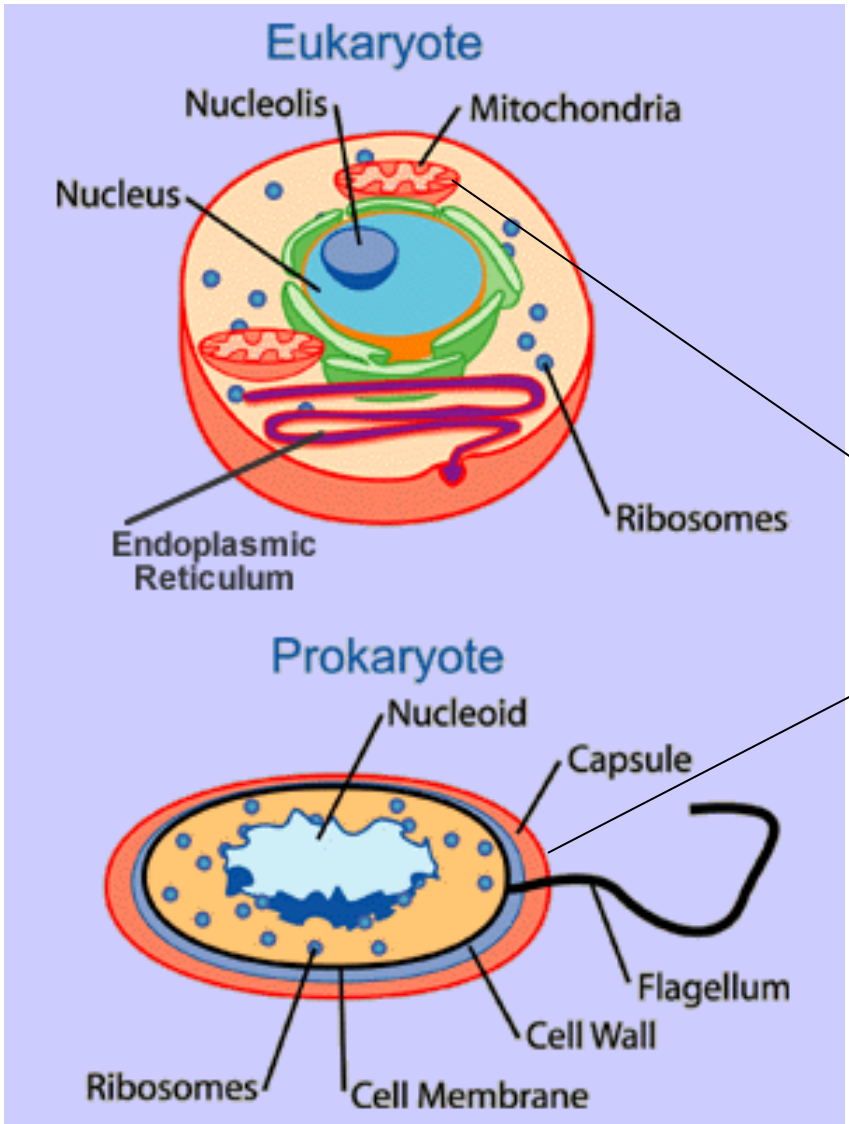
Origin of the eukaryotic cell: The endosymbiosis/intron hypothesis

Eugene V. Koonin
NCBI, NIH, Bethesda

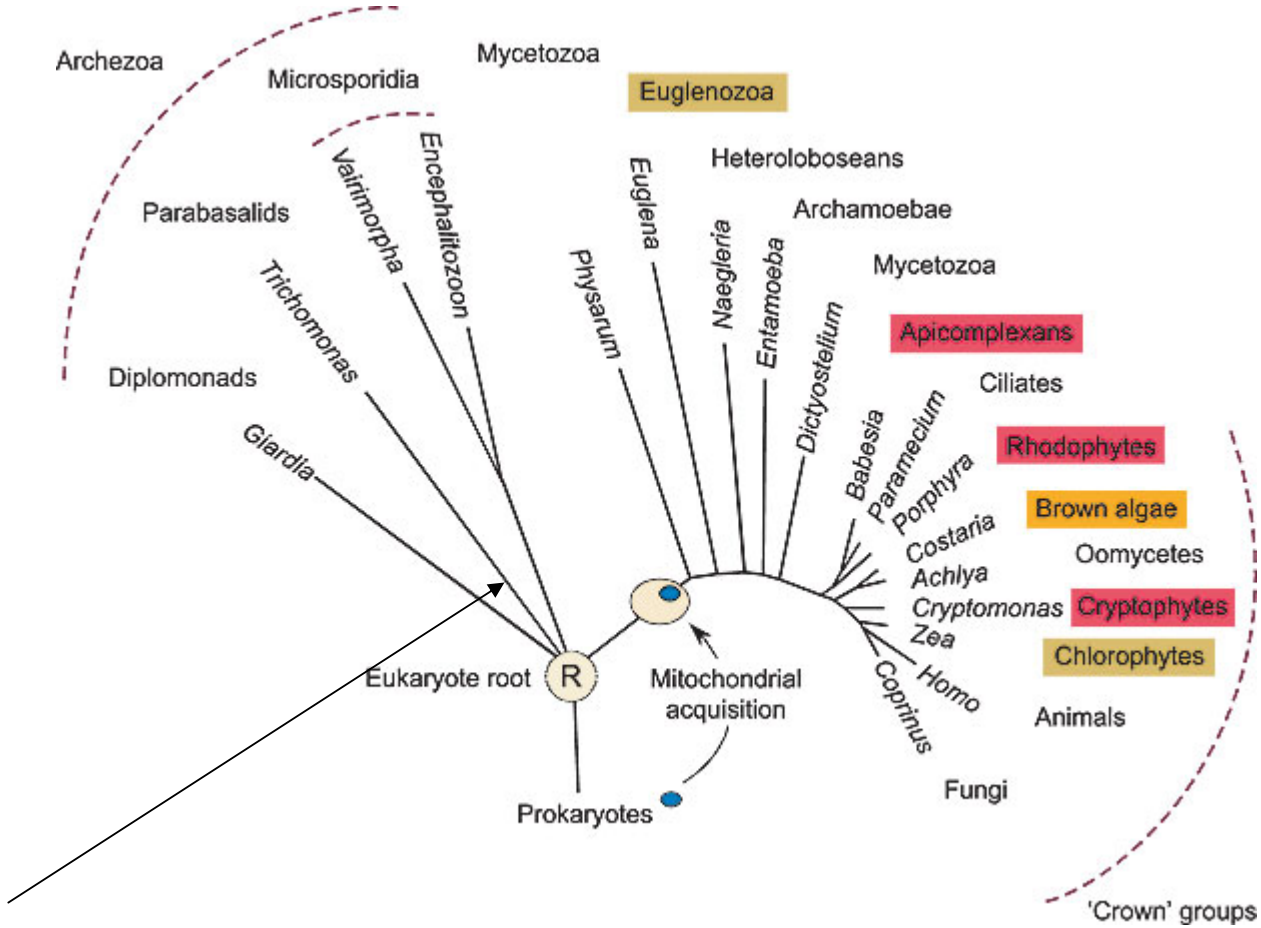
KITP, Santa Barbara, March 5, 2007


Golgi apparatus
endoplasmic reticulum
nucleus
DNA
nuclear membrane
mitochondrion
mitochondria
Mouse L-Cell
print mag. 15,000x
©John J. Cardamone, Jr.
cytoplasmic membrane

V_{euk}/V_{pro} = 1,000



Endosymbiosis



"Amitochondrial" eukaryotes

The archezoan hypothesis – the classic scenario of eukaryotic evolution

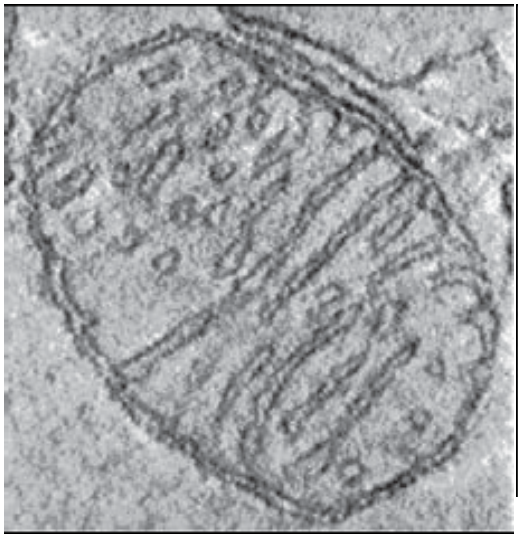
Embley TM, **Martin W**.

Eukaryotic evolution, changes and challenges

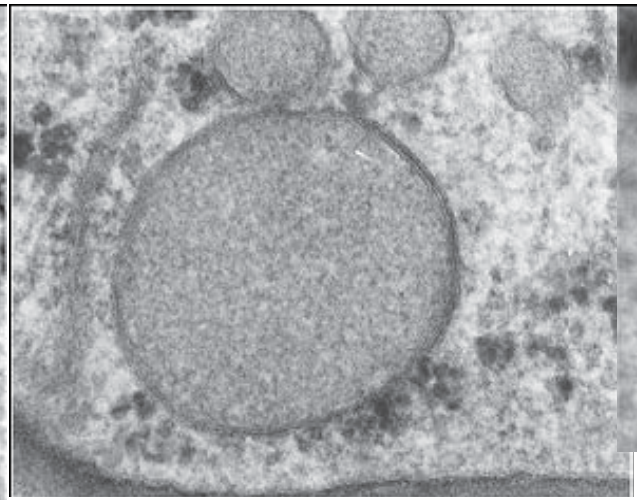
The idea that some eukaryotes primitively lacked mitochondria and were true intermediates in the prokaryote-to-eukaryote transition was an exciting prospect. It spawned major advances in understanding anaerobic and parasitic eukaryotes and those with previously overlooked mitochondria. But the evolutionary gap between prokaryotes and eukaryotes is now deeper, and the nature of the host that acquired the mitochondrion more obscure, than ever before.

Nature. 2006 Mar 30;440(7084):623-30

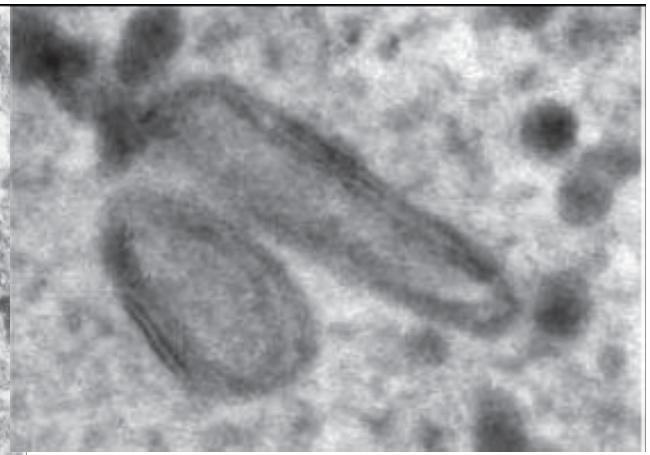
There are no (known) true amitochondrial eukaryotes!



Animal mitochondrion



Hydrogenosome from an anaerobic fungus

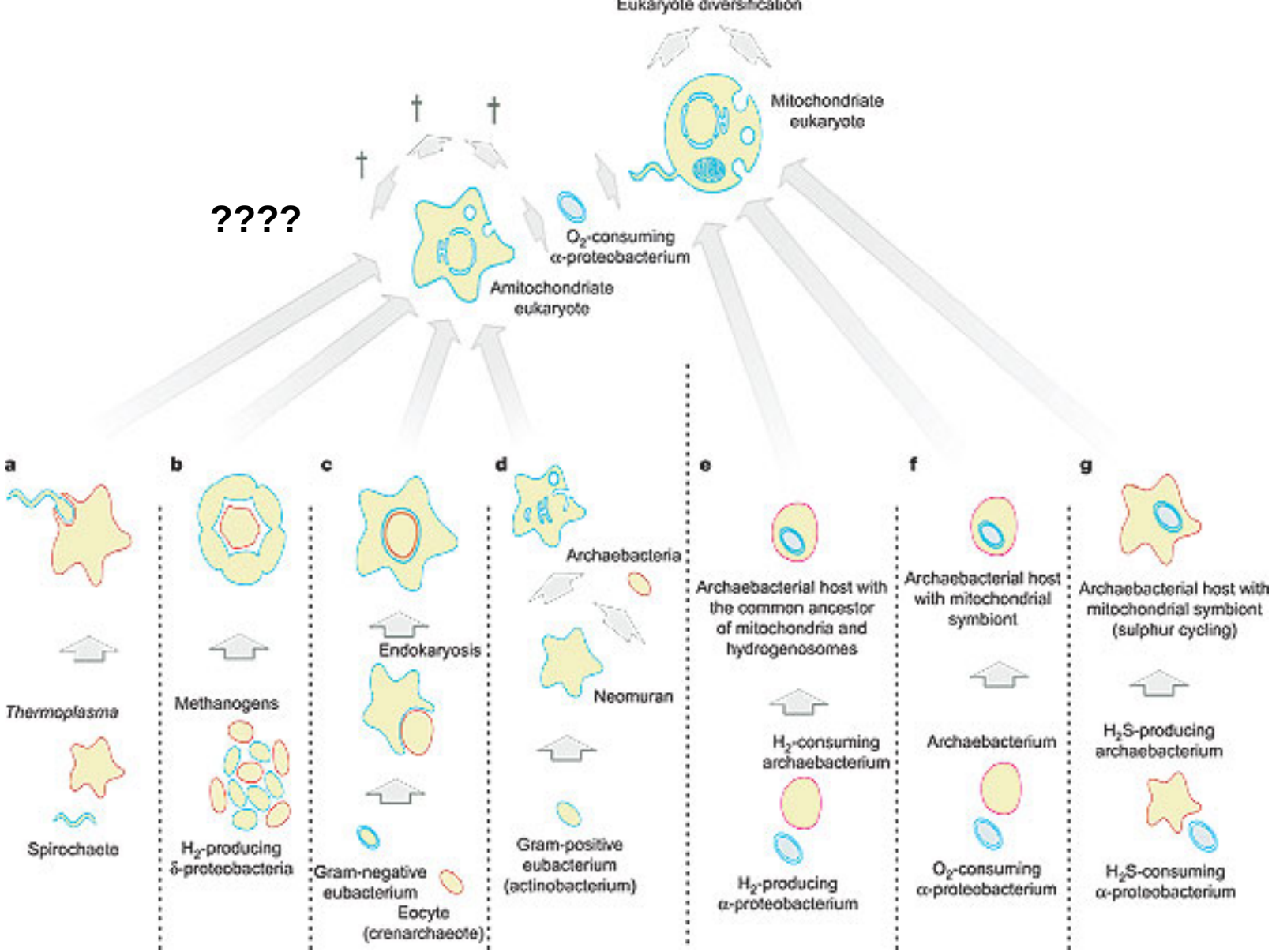


Mitosomes from *Giardia*

van der Giezen M, Tovar J. **Degenerate mitochondria.**

EMBO Rep. 2005 Jun;6(6):525-30.

All “archezoa” possess:
-mitochondrial genes in nuclear genomes
-degenerate derivatives of mitochondria
They are not archezoa at all!

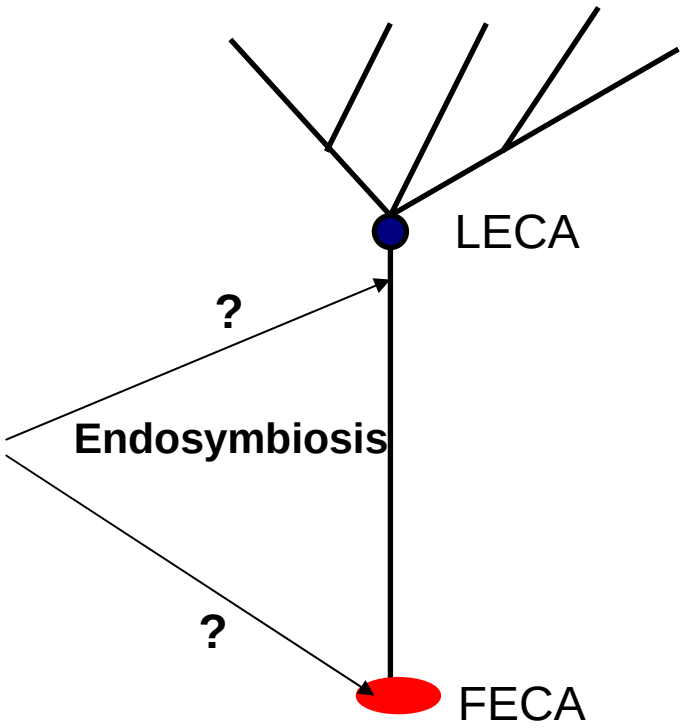


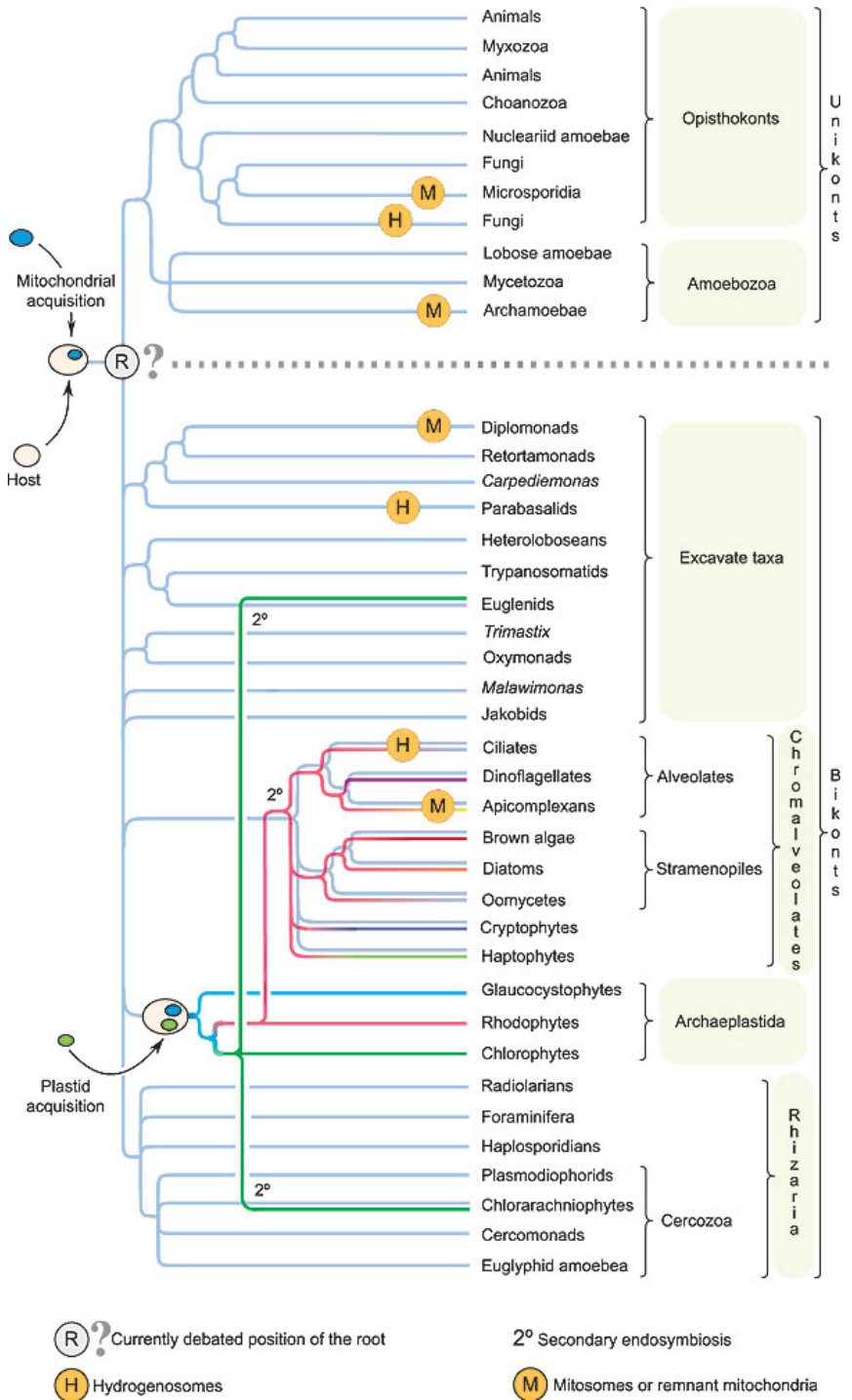
Hypotheses on the origin of eukaryotes

Embley, Martin, Nature 2006

The symbiosis hypotheses:
All modern eukaryotes diverged after (and as a result of??) the mitochondrial endosymbiosis…

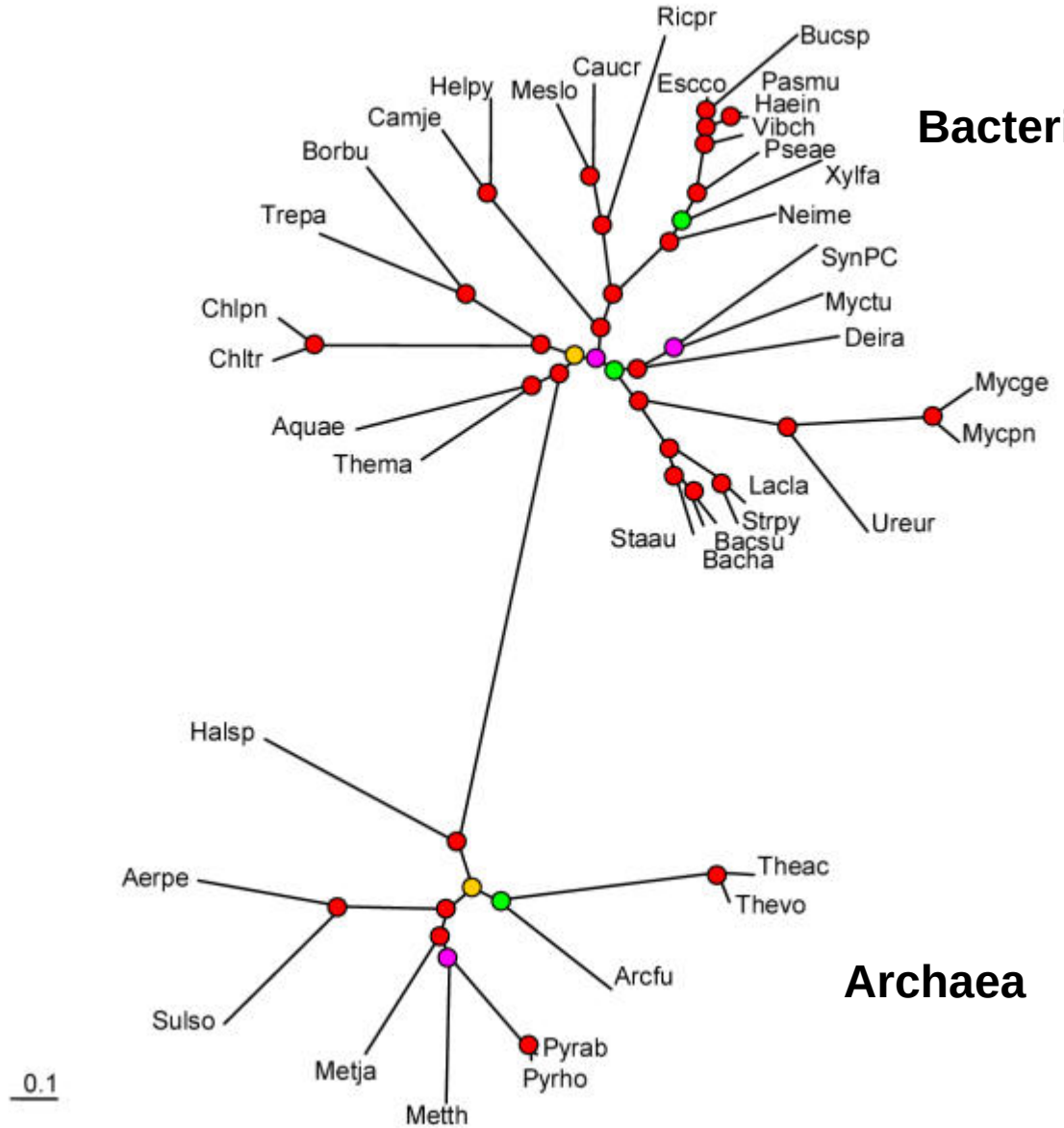
But what was the host??





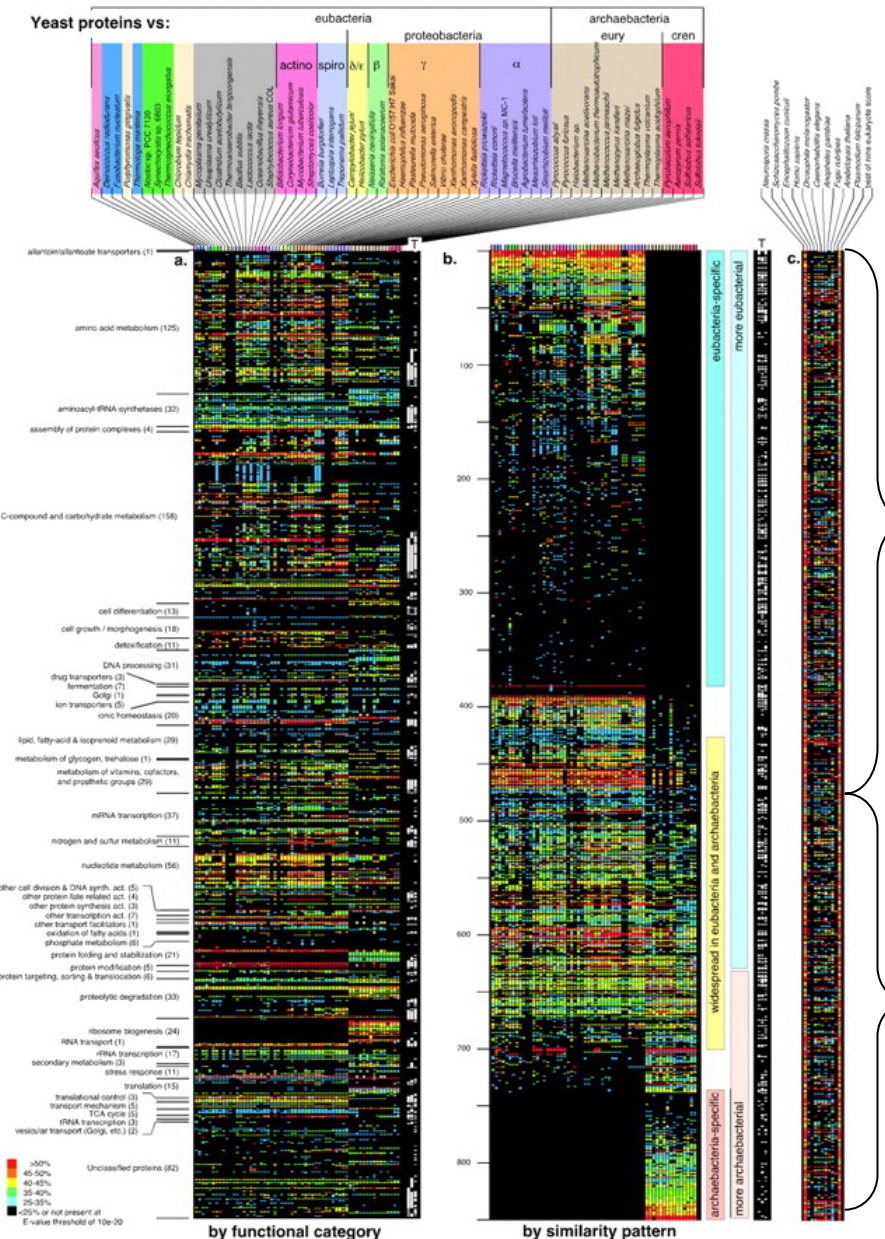
Archaea and Bacteria are the two domains of prokaryotes (Woese 1977)

Archaea and Bacteria are well separated in trees and differ qualitatively:
Non-homologous systems for
-DNA replication
-membrane biogenesis



Wolf et al. BMC Evol Biol. 2001 Oct 20;1:8

The eukaryotic gene set is a mixture of archaeal genes, bacterial genes, ancestral genes of uncertain origin and unique genes



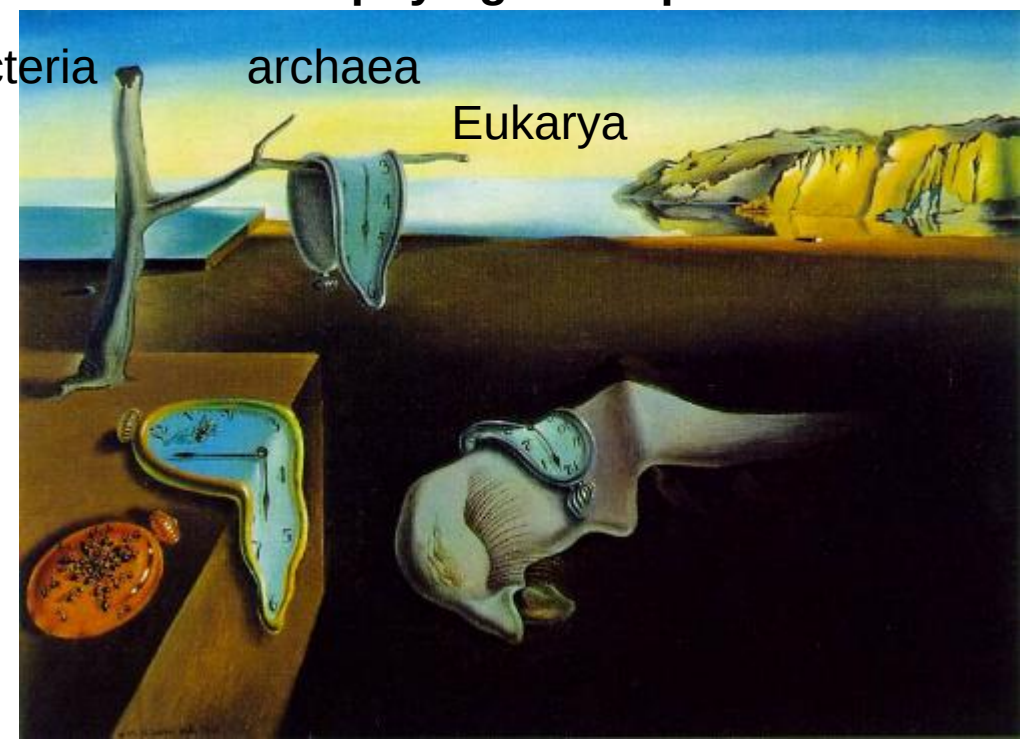
Bacterial – mostly, operations (metabolism etc)

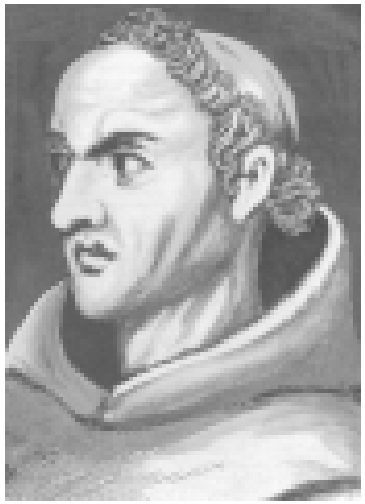
Archaeal – mostly informational (replication-transcription-translation)

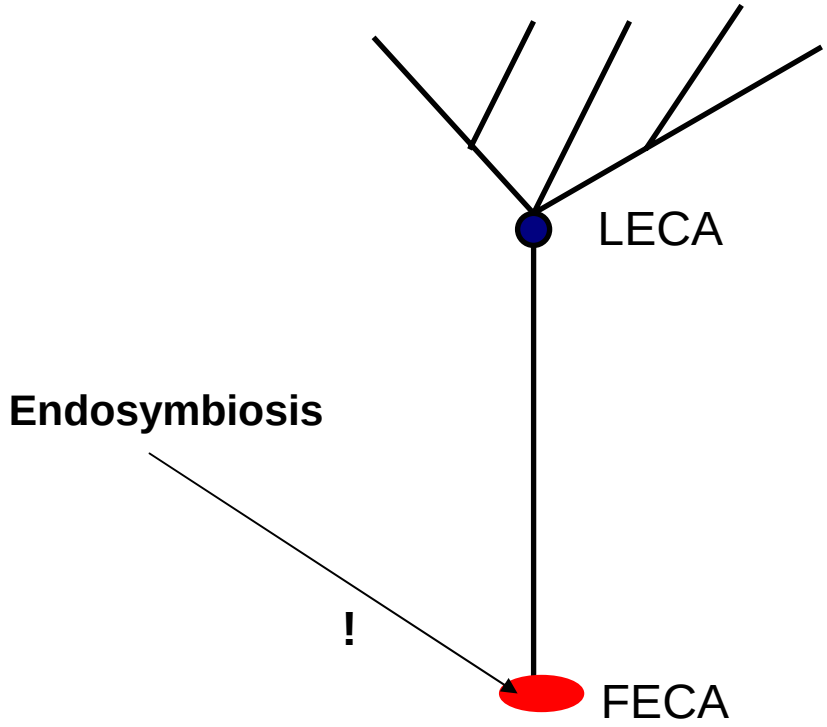
Esser et al. Mol Biol Evol. 2004;21:1643-60

Harris JK, Kelley ST, Spiegelman GB, Pace NR.
The genetic core of the universal ancestor.
Genome Res. 2003 Mar;13(3):407-12
We used the Clusters of Orthologous Groups database and information from published genomes to search for other universally conserved genes that have the same phylogenetic pattern as ribosomal RNA, and therefore constitute the ancestral genetic core of cells. Our analyses identified a small set of genes that can be traced back to the universal ancestor and have coevolved since that time.

- **Most of the core genes belong to the translation and transcription systems and show the "canonical" phylogenetic pattern**







Occam's razor suggests that mitochondrial endosymbiosis triggered eukaryogenesis, i.e., the host for the symbiont was a "garden variety" archaeon

If so, aftermath of the symbiosis involved a truly momentous transformation of the system

Has endosymbiosis triggered eukaryogenesis?

If so, HOW?

The eukaryotic cell is an epitome of (seemingly) irreducible organizational complexity

ALL eukaryotes have:

- Nucleus
- Endomembrane systems (ER)
- Cytoskeleton
- Mitochondria
- Ubiquitin signaling system
- Introns and spliceosomes

No clear intermediates for any of this….

Hypothesis:

- ***Endosymbiosis triggered eukaryogenesis***
- ***The nucleus and, possibly, other major features of the eukaryotic cell evolved as defense against massive invasion of introns from the endosymbiont into protein-coding genes***

Martin, Koonin. 2006. Nature; Koonin, 2006, Biology Direct

Spliceosomal introns: the hallmark of eukaryotic genes

There are no known eukaryotes without at least a few introns and a **(nearly) fully-fledged spliceosome**

Simpson AG, MacQuarrie EK, Roger AJ.
Eukaryotic evolution: early origin of canonical introns.
Nature. 2002 Sep 19;419(6904):270.
Spliceosomal introns, one of the hallmarks of eukaryotic genomes, were thought to have originated late in evolution and were assumed not to exist in eukaryotes that diverged early -- until the discovery of a single intron with an aberrant splice boundary in the primitive 'protozoan' Giardia. Here we describe introns from a close relative of Giardia, Carpediemonas membranifera, that have boundary sequences of the normal eukaryotic type, indicating that canonical introns are likely to have arisen very early in eukaryotic evolution.

Collins L, Penny D.
Complex spliceosomal organization ancestral to extant eukaryotes.
Mol Biol Evol. 2005 Apr;22(4):1053-66

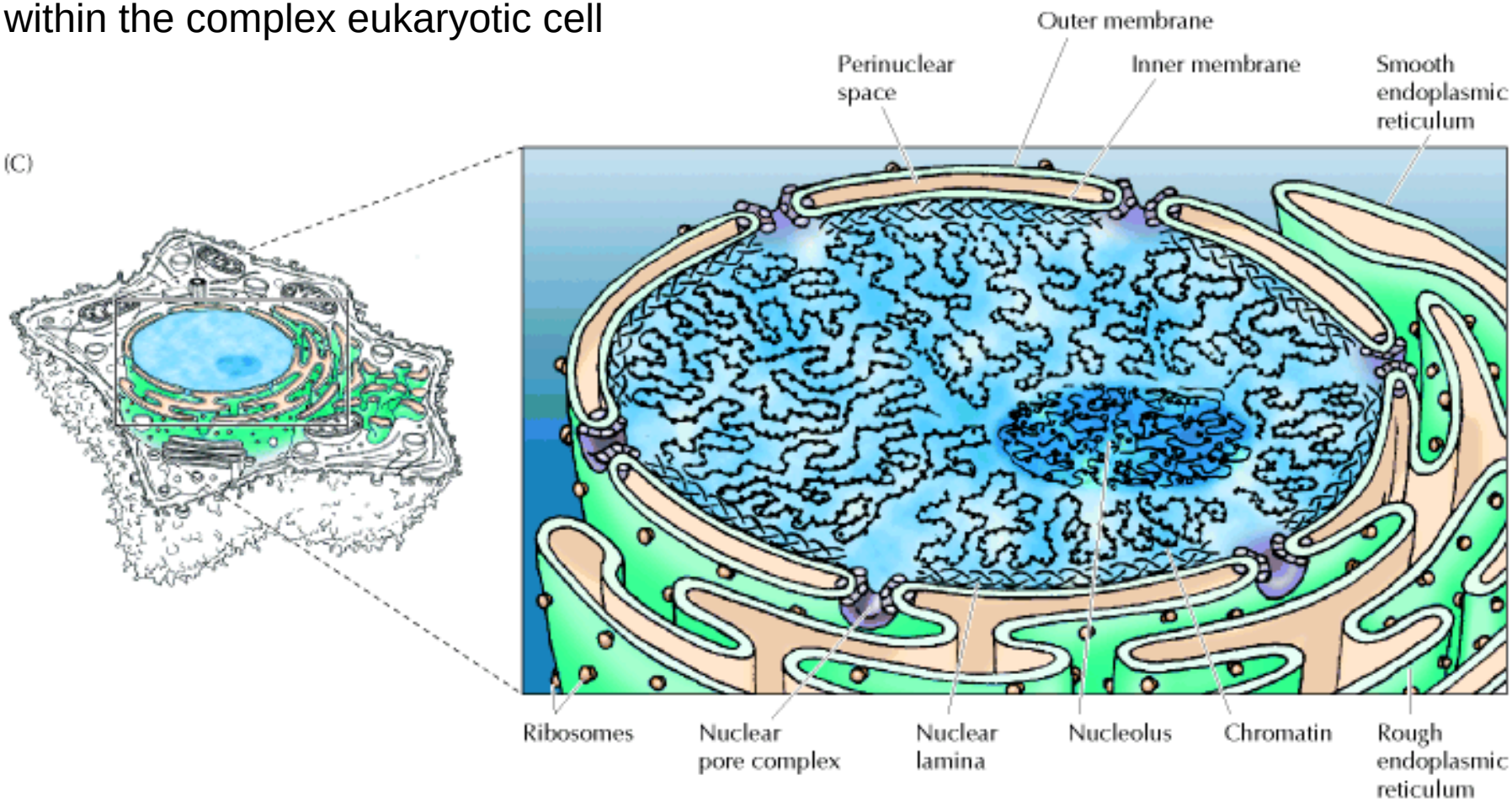
…examination of the distribution of spliceosomal components indicates that not only was a spliceosome present in the eukaryotic ancestor but it also contained most of the key components found in today's eukaryotes. All the small nuclear ribonucleoproteins (snRNPs) protein components are likely to have been present, as well as many splicing-related proteins.

Not only are introns inferred to have been present in the last common ancestor of eukaryotes, but their positions are highly conserved: ~25% human intron positions are shared with Arabidopsis which might mean they come from the last eukaryotic common ancestor.

Rogozin IB, Wolf YI, Sorokin AV, Mirkin BG, Koonin EV.
Remarkable interkingdom conservation of intron positions and massive, lineage-specific intron loss and gain in eukaryotic evolution.
Curr Biol. 2003 Sep 2;13(17):1512-7.

Roy SW, Gilbert W.
Complex early genes.
Proc Natl Acad Sci U S A. 2005 Feb 8;102(6):1986-91.

The ***nucleus*** is a mind-bogglingly complex organelle within the complex eukaryotic cell

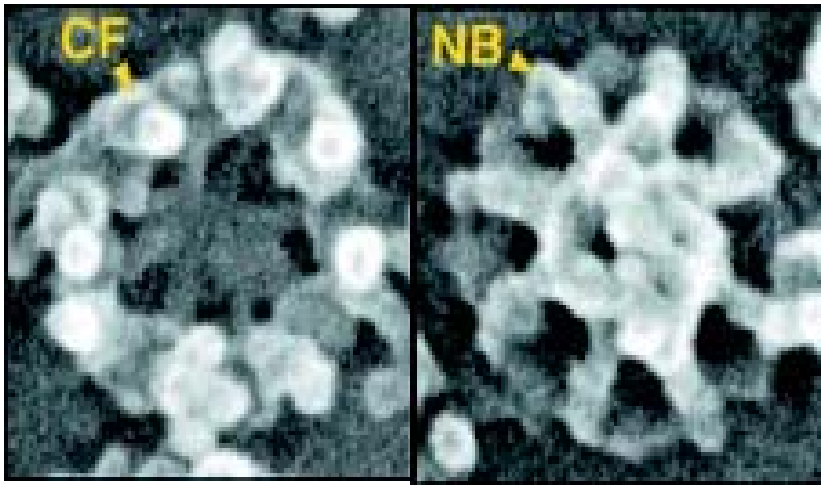


Nuclear matrix, lamina – stability and structural integrity

Bilayer nuclear membrane - containment

Nuclear pore complex – communication and transport

The nuclear pore complex



Diameter – 120 nm

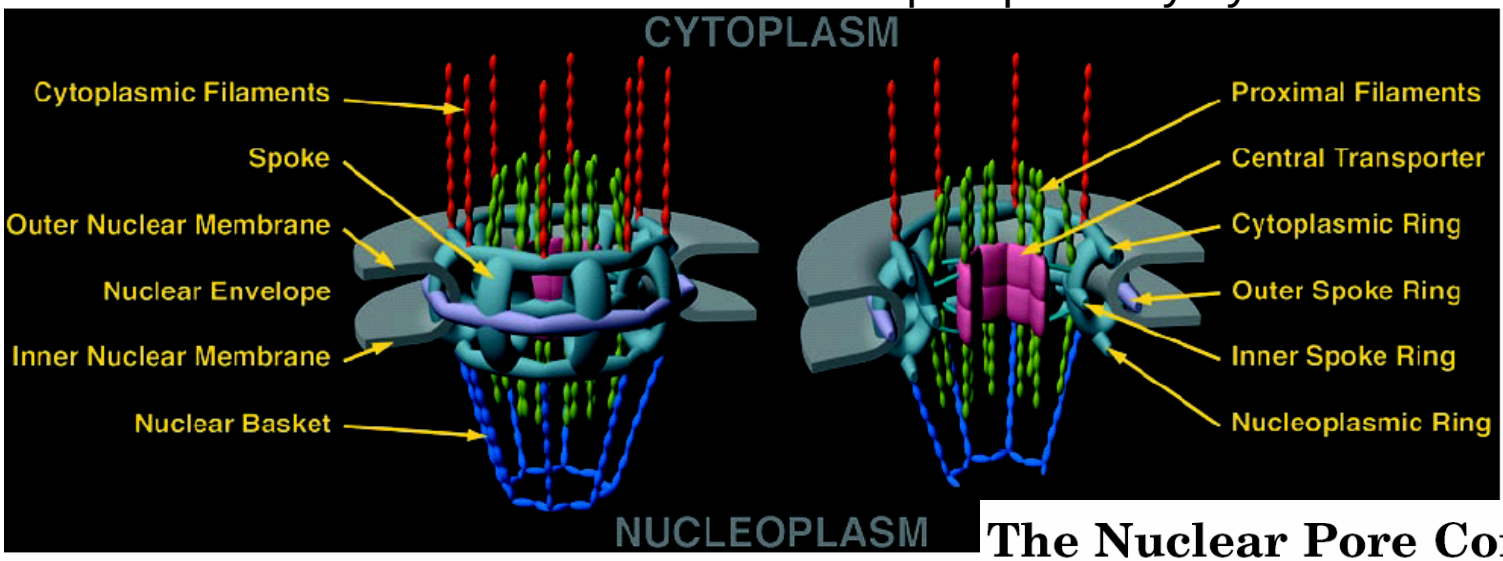
8-fold symmetry

Size – 60-125 MDa (30X ribosome)

Small proteins (<60 kDa) – diffusion

Large proteins and complexes (>60 kDa) – facilitated, energy-dependent

mRNAs - facilitated, energy-dependent, but in part probably by diffusion



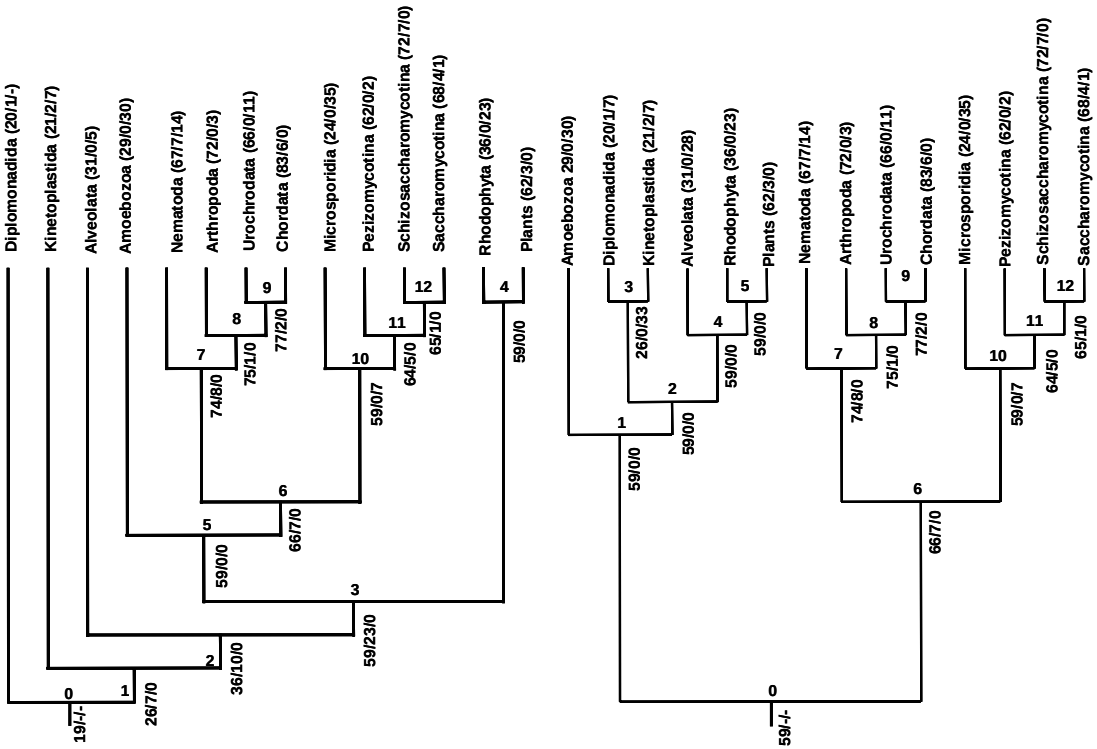
At least >70 integral proteins in animals

The Nuclear Pore Complex as a Transport Machine*

Published, JBC Papers in Press, April 5, 2001,
DOI 10.1074/jbc.R100015200

Michael P. Rout‡ and John D. Aitchison§¶

The nuclear pore apparatus, much like the spliceosome, can be traced back to the last eukaryotic common ancestor in a complex form - perhaps, nearly as complex as the modern versions



‘Crown group’ tree:
substantial increase in complexity during eukaryotic evolution

Bikonth-opisthokonth tree:
almost no increase in complexity

Mans BJ, Anantharaman V, Aravind L, Koonin EV.
Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex.
Cell Cycle. 2004 Dec;3(12):1612-37.

Since the nucleus is the eponymous feature of eukaryotes, it might seem "natural" to assume that it emerged before other features of eukaryotic cellular organization

However…given that no one has been able to discover a eukaryote (a nucleate cell) without

- Mitochondria
- Introns/spliceosome

…or, conversely, any kind of cell possessing either mitochondria or Introns/spliceosome but ***no*** nucleus

…**perhaps, such primitive anucleate cells ("Archezoa") never existed?**

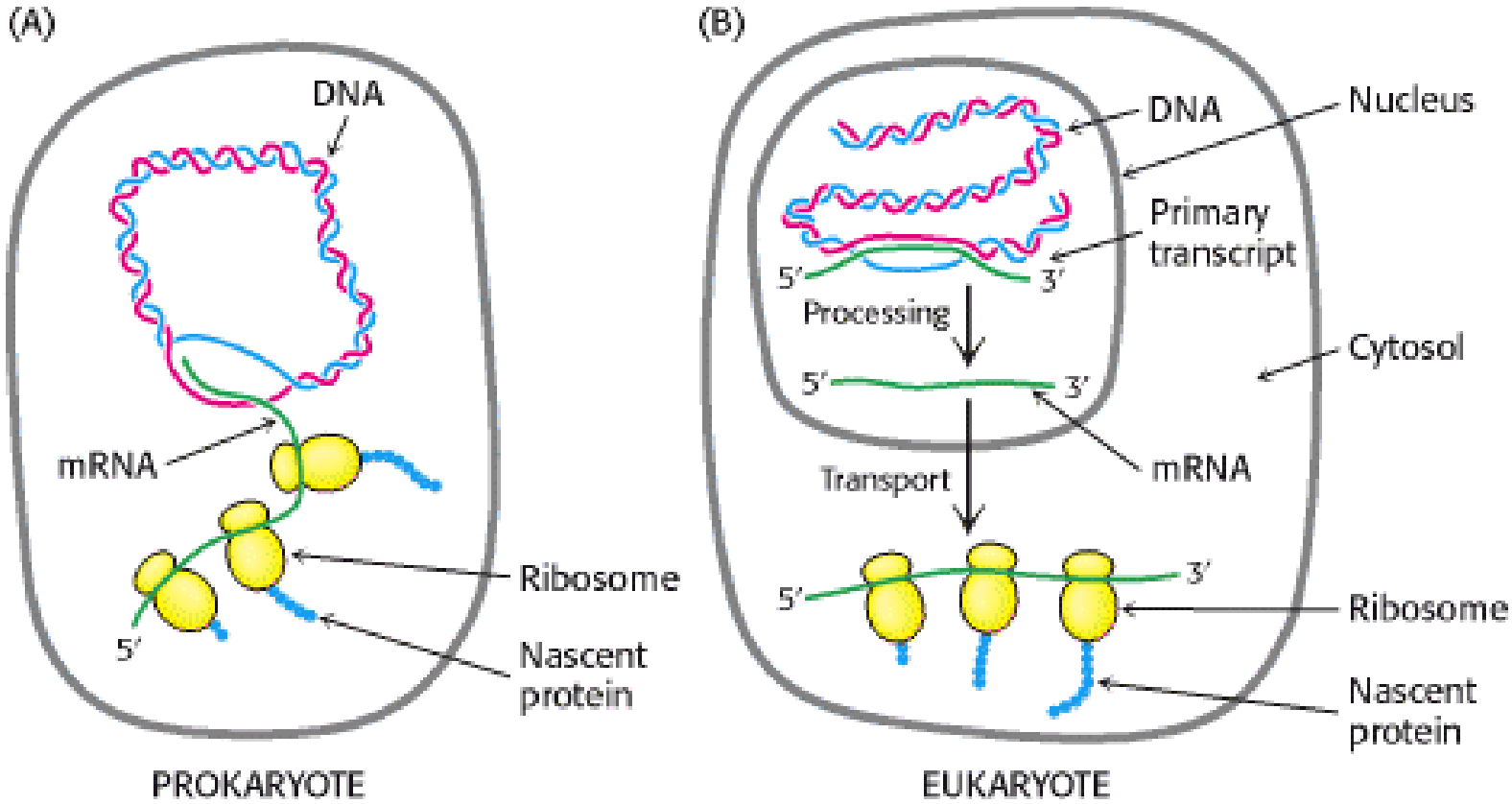
More importantly, is there a causal connection between the emergence of mitochondria, introns, and nucleus?

The nuclear compartment is almost certainly a specifically eukaryotic novelty, but here it is suggested to have arisen in a cell that possessed a facultatively anaerobic, heterotrophic organelle, the common ancestor of mitochondria and hydrogenosomes.

Martin, W. 1999. Proc. R. Soc. B

Why a nucleus?

- The nucleus' raison d'etre is unclear
- Nucleus complicates the cell functioning by eliminating the transcription-translation coupling typical of prokaryotes and requiring a complex export-import machinery
- There should be a powerful driving force behind the evolution of such a complex machine



Prokaryotes: **Closely coupled transcription and translation (cotranscriptional translation)**
Eukaryotes: **spatially and temporally separated transcription and translation**

Darnell, Lodish, Baltimore, *Molecular Cell Biology*

Could there be a causal link between the original spread of introns and the origin of the nucleus?

Hypothesis:

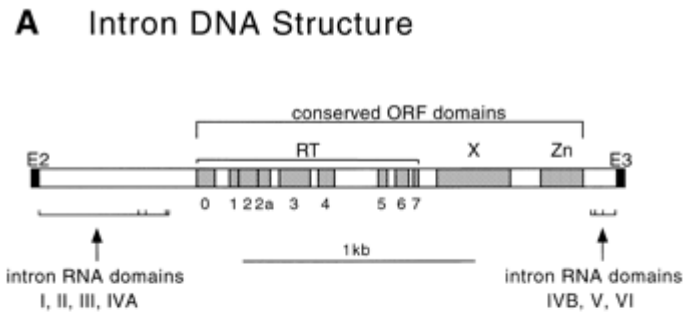
The nucleus evolved as defense against massive invasion of introns into protein-coding genes

Without compartmentalization to separate intron-containing transcripts from the ribosomes and to allow only processed mRNAs to be translated, aberrant proteins would accumulate, probably, to a fatal effect, especially, because **splicing is much slower than translation**

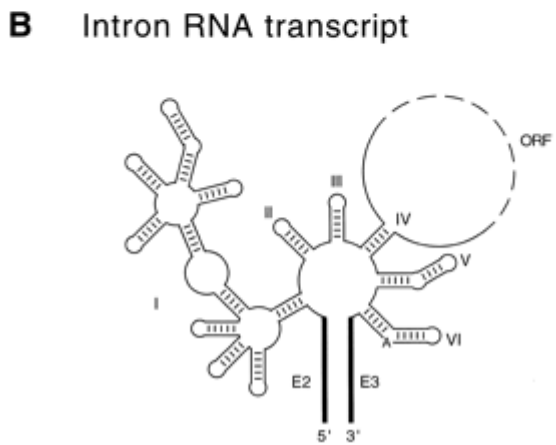
Why, all of a sudden, intron invasion?

Origin of spliceosomal introns: derivation from Group II self-splicing introns

(prokaryotes have introns, too, albeit few and very different ones)

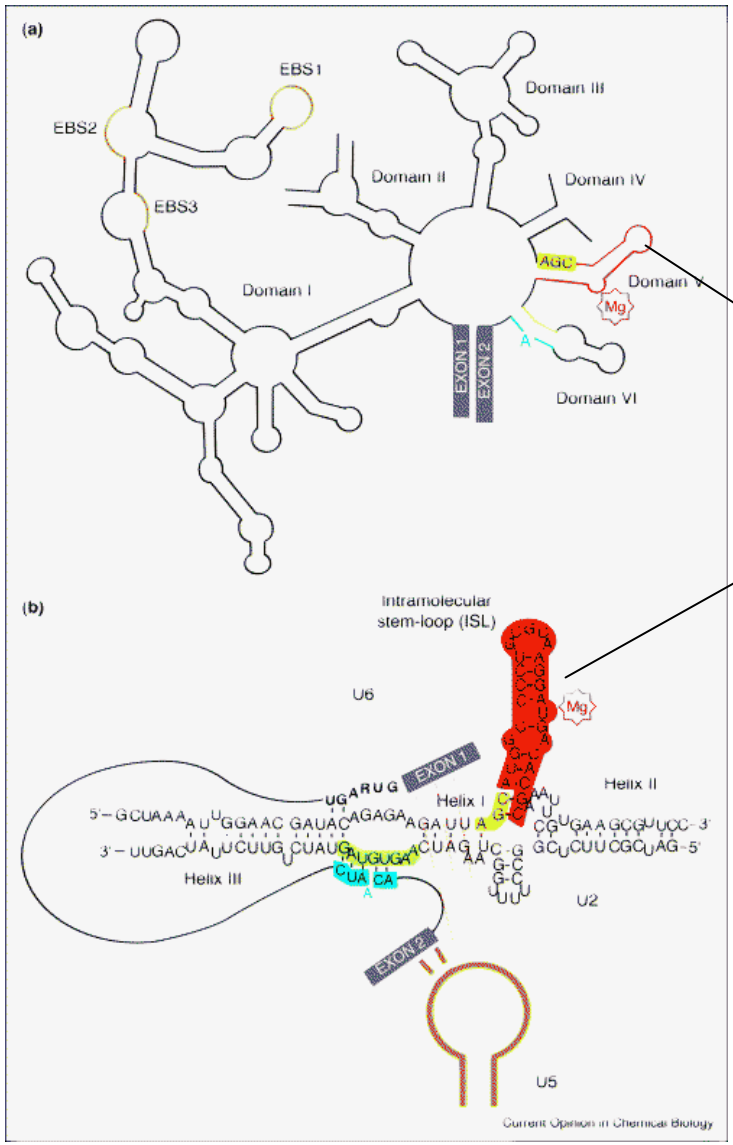


Typical organization of
A Group II itnron



Steven Zimmerly*, Georg Hausner and Xu-chu Wu
Phylogenetic relationships among group II intron ORFs
Nucleic Acids Research, 2001, Vol. 29, No. 5 **1238-1250**

There are striking similarities between spliceosomal introns junctions and snRNAs (catalytic moieties of the spliceosome) and structural elements of Group II introns



Valadkhan S.
snRNAs as the catalysts of pre-mRNA splicing.
Curr Opin Chem Biol. 2005 Dec;9(6):603-8.

In bacteria, group II introns are retroelements that are deleterious when inserted into important protein-coding genes

Lixin Dai and Steven Zimmerly
Compilation and analysis of group II intron insertions in bacterial genomes: evidence for retroelement behavior.
Nucleic Acids Research, 2002, Vol. 30, No. 5 **1091-1102**

Bacterial introns are not inserted into conserved genes, are often inserted outside of genes altogether and are frequently fragmented, suggesting a high rate of intron gain and loss. Some introns have multiple natural homing sites while others insert after transcriptional terminators. All bacterial group II introns identified to date encode reverse transcriptase open reading frames and are either active retroelements or derivatives of retroelements. Together, these observations suggest that group II introns in bacteria behave primarily as retroelements rather than as introns …

α–Proteobacteria (the ancestors of the mitochondria) have a relatively large number of group II introns, and these can invade new sites

Munoz E, Villadas PJ, Toro N.
Ectopic transposition of a group II intron in natural bacterial populations.
Mol Microbiol. 2001 Aug;41(3):645-52.

Self-splicing group II introns are thought to be the evolutionary progenitors of eukaryotic spliceosomal introns. The invasion of novel (ectopic) sites by group II introns is considered to be a key mechanism by which spliceosomal introns may have become widely dispersed.
……

We found that ectopic transposition of RmInt1 to the oxi1 site occurred in this natural bacterial population. This ectopic transposition was also the most frequent genetic event observed. This work provides further evidence that the ectopic transposition of group II introns is an important mechanism for their spread in natural bacterial populations.

A crucial speculation: in the wake of mitochondrial endosymbiosis, the symbiont's group II introns went on a rampage – probably, through a combination of weak purifying selection due to the small Ne of the chimera and absence of specific control mechanisms in the host archaeon

There is independent, convergent evidence:
The mitochondrial endosymbiont not only could have unleashed Group II introns but also supplied essential building blocks for the nucleus

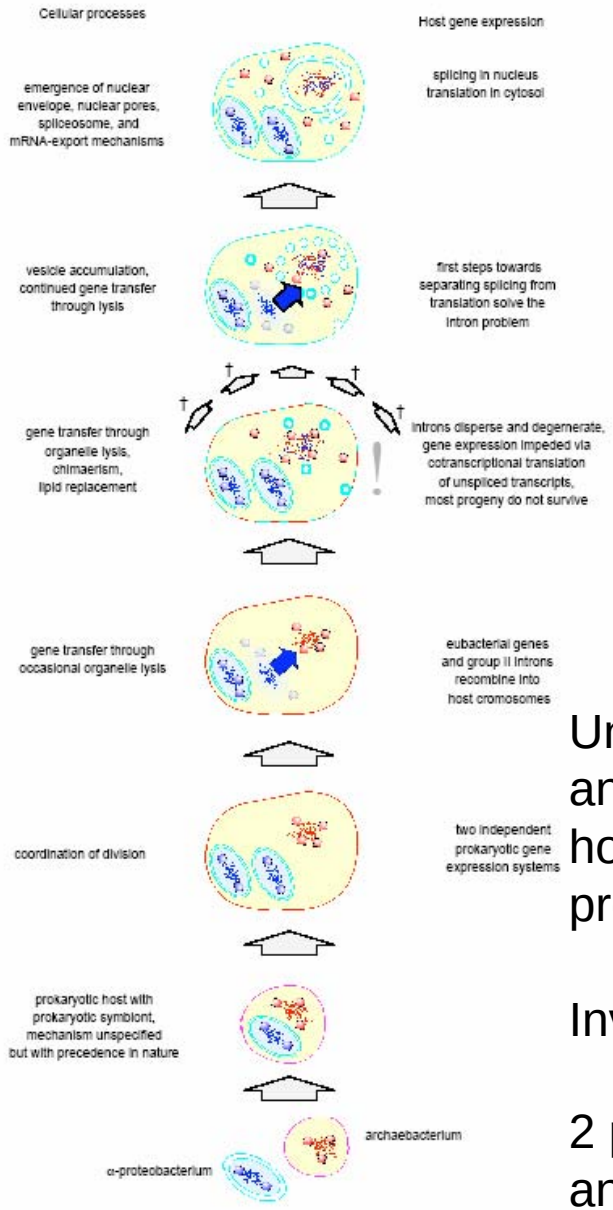
Likely endosymbiont contributions to the origin of nucleus

Mans BJ, Anantharaman V, Aravind L, Koonin EV.
Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. Cell Cycle. 2004 Dec;3(12):1612-37.

It is shown that several central components of the NPC, in particular, the RanGDP import factor NTF2, the HEH domain of Src1p-Man1, and, probably, also the key domains of karyopherins and nucleoporins, the HEAT/ARM and WD40 repeats, ***have a bacterial, most likely, endosymbiotic origin.***

……

…several NPC proteins containing super-structure-forming alpha-helical and beta-propeller modules are most closely related to corresponding proteins in the cytoplasmic vesicle biogenesis and coating complexes. From these observations, we infer an autogenous scenario of nuclear evolution in which ***the nucleus emerged in the primitive eukaryotic ancestor (the "prekaryote") as part of cell compartmentalization triggered by archaeo-bacterial symbiosis.***



Origin of nucleus and spliceosome

Dispersal of introns, population bottleneck

Unidirectional flow of genes and introns from symbiont to host – ratchet due to propagation/lysis of symbiont

Invasion

2 prokaryotes: archaeon and α-proteobacterium

Martin, Koonin, 2006, Introns and the origin of nucleus-cytosol compartmentalization. Nature 440: 41-5

On the origin of the spliceosome

- Inactivation of RT in the invading group II introns – loss of mobility – prevention of further damage, but also impairment of self-splicing

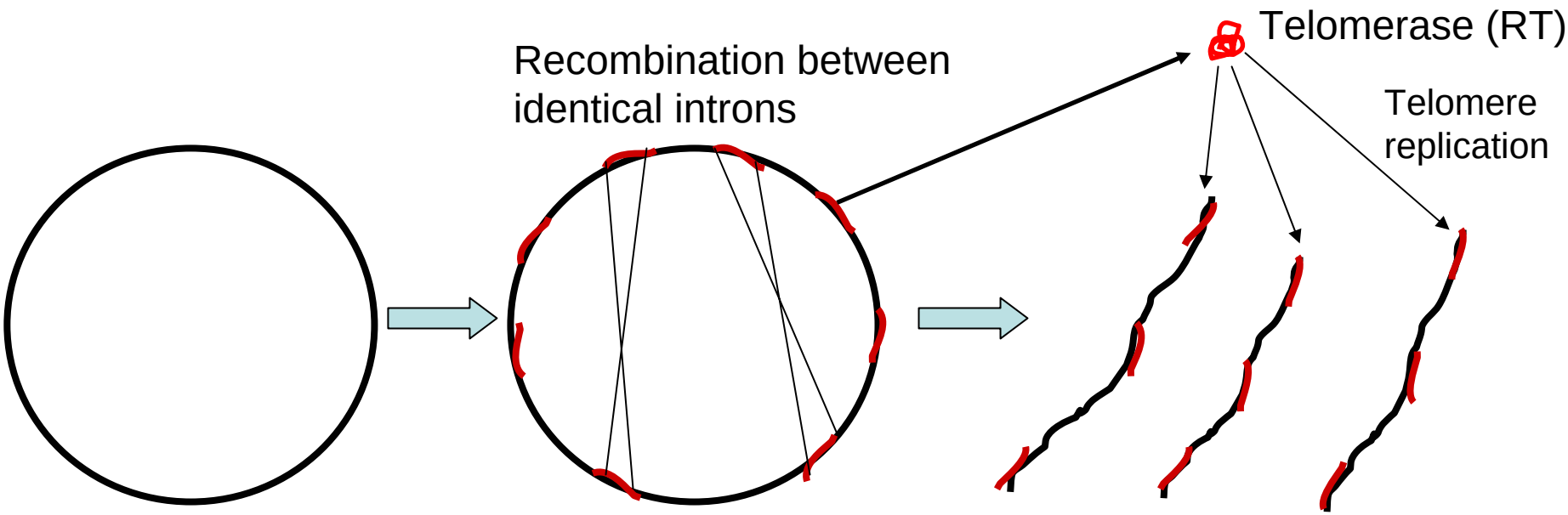
- RT-mediated splicing in-trans inefficient

- Selection for an alternative mechanism

- Recruitment of Sm protein

- Early origin of splicing-export coupling

From circular (prokaryotic) to linear (eukaryotic) chromosomes



Circular archaeal chromosome

Circular chromosome with multiple inserted, mobile, group II introns

Linear eukaryotic chromosome

Salient features of the model:

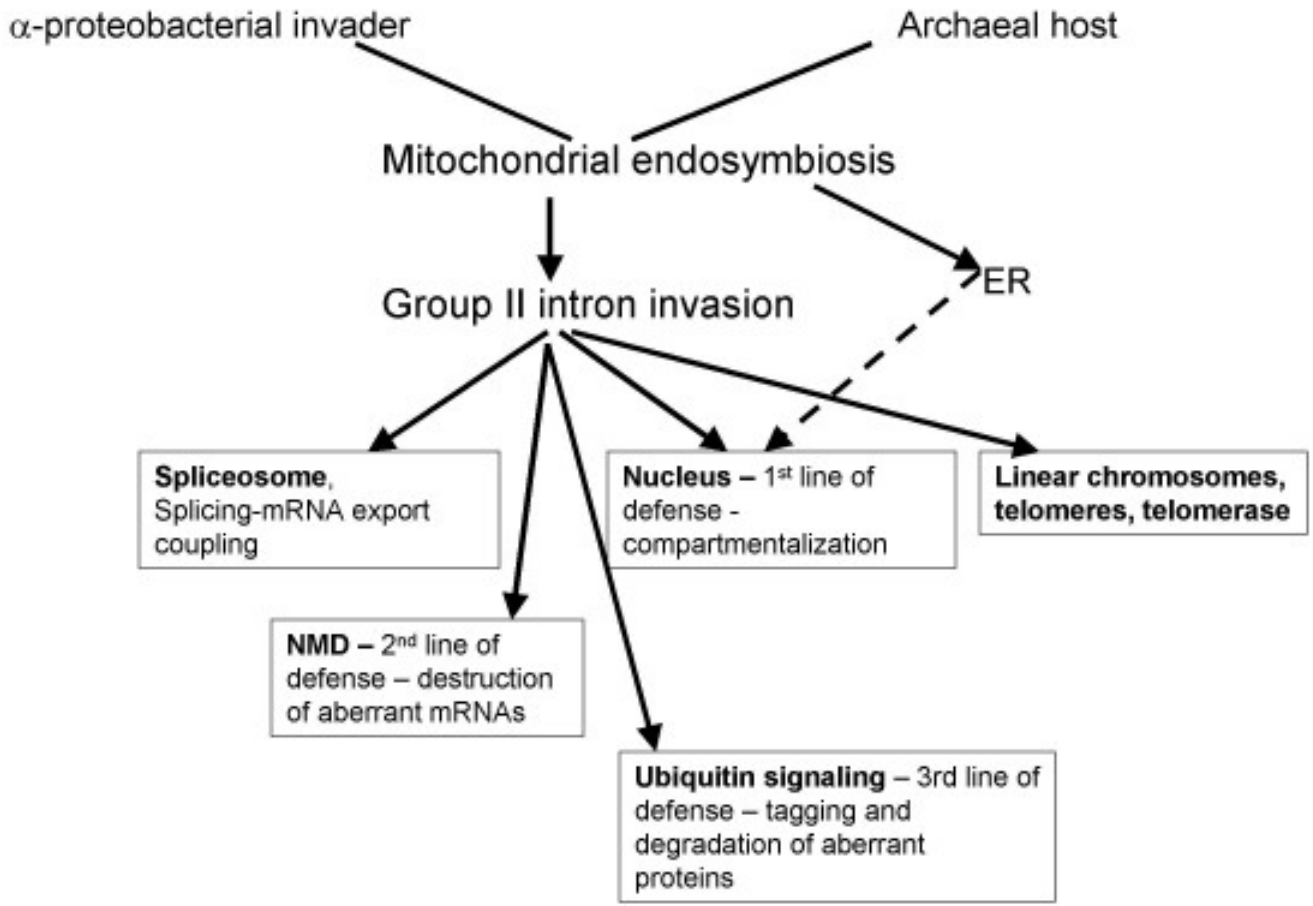
•Mystery of nuclear origin receives a plausible explanation – it was ***the only way to survive the intron invasion***

•Single chain of causation – acquisition of mitochondria triggers ***all other pivotal events*** in the emergence of the eukaryotic cell

Corollary:

The phase of evolution from acquisition of mitochondria to bona fide eukaryotic cell was extremely rapid and turbulent (“inflationary phase”), and might have been one long population bottleneck, which allowed fixation of many otherwise deleterious “macromutations” by drift

The proposed chain of causes and effects in eukaryogenesis – the pivotal roles of mitochondrial endosymbiosis and intron invasion

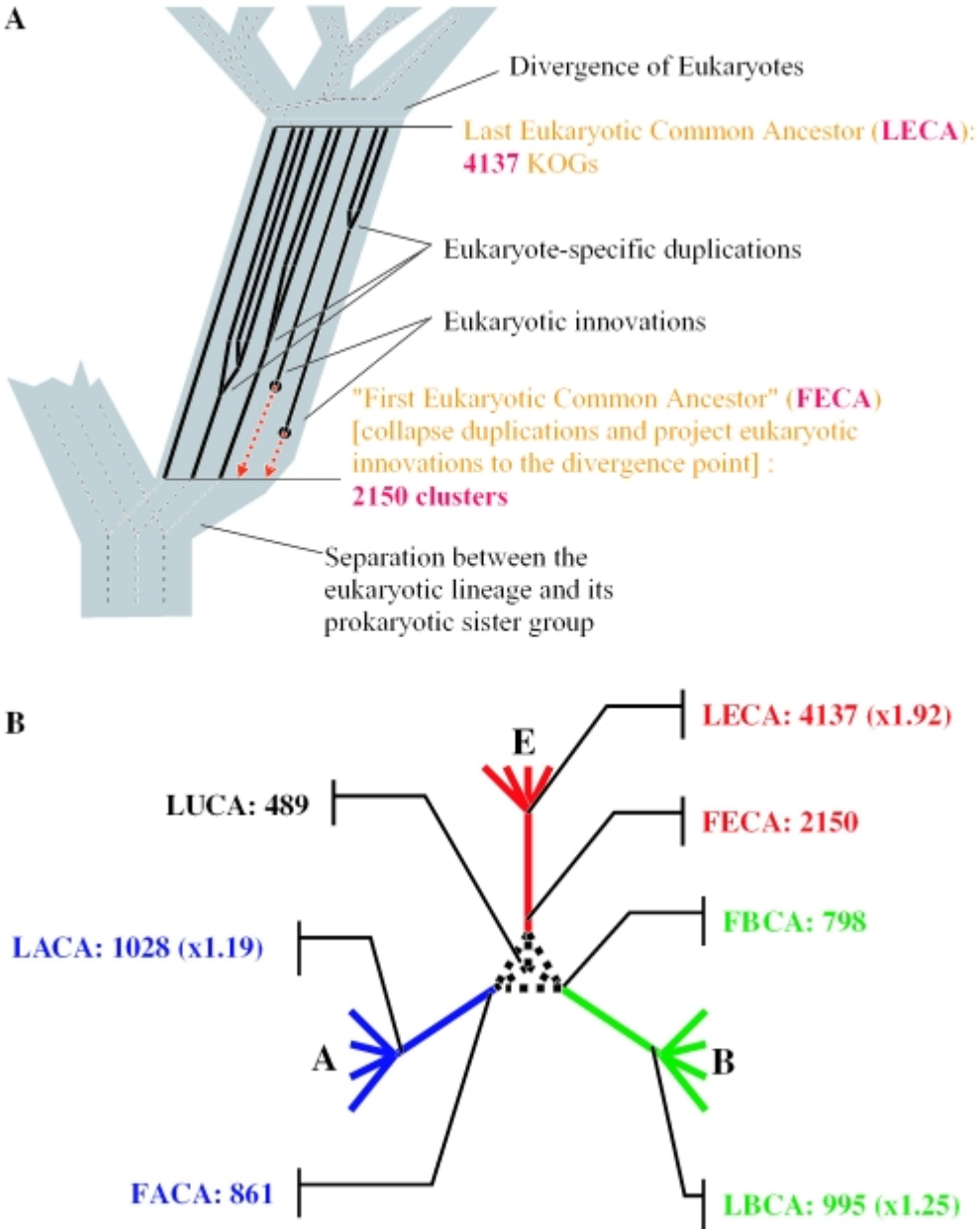


Koonin EV. The origin of introns and their role in eukaryogenesis: a compromise solution to the introns-early versus introns-late debate? Biol Direct. 2006 Aug 14;1:22

Possible falsification: identify

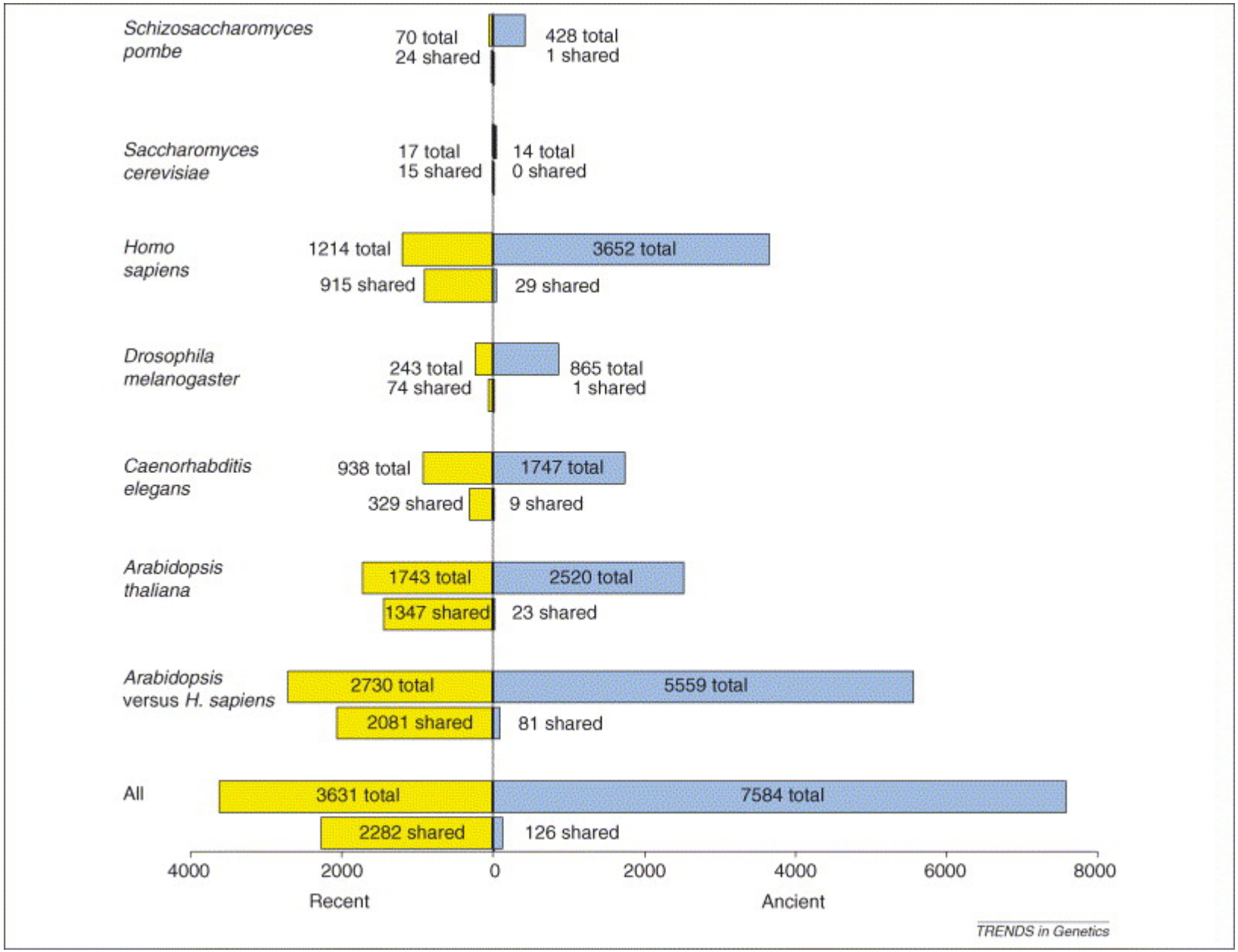
- a eukaryote without traces of the mitochondrion (archezoon)
- an organism with spliceosomal introns and spliceosome but without nucleus or mitochondria
- extensive transcription-translation coupling (nuclear translation) in eukaryotes

- Can we get a glimpse of the earliest, turbulent phase of the eukaryotic evolution – *the epoch of intron invasion?*

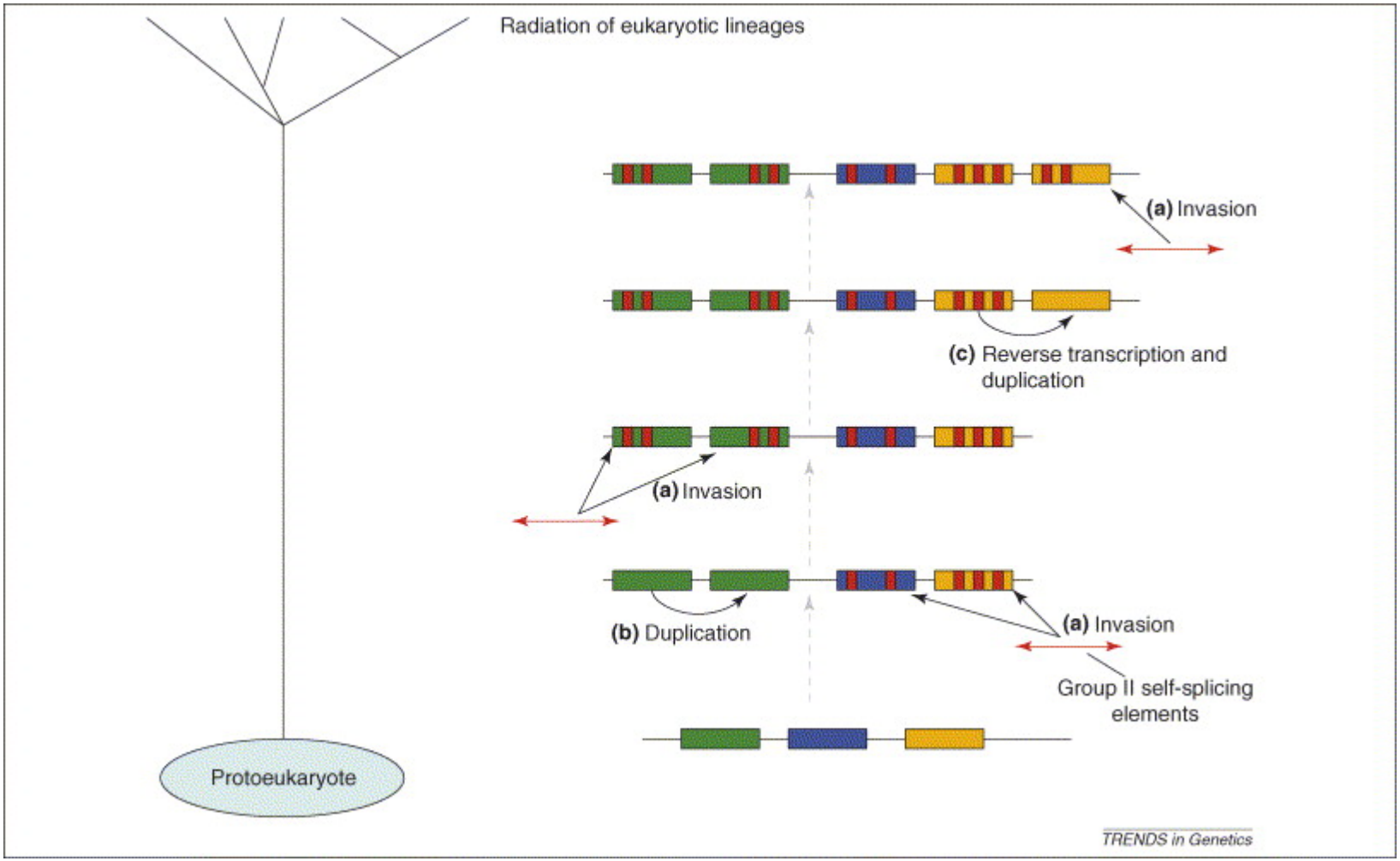


A burst of gene duplication associated with the emergence of eukaryotes

Makarova et al. Nucleic Acids Res. 2005 Aug 16;33(14):4626-38



Conservation – and lack thereof – of intron positions in ancient and recent eukaryotic paralogs

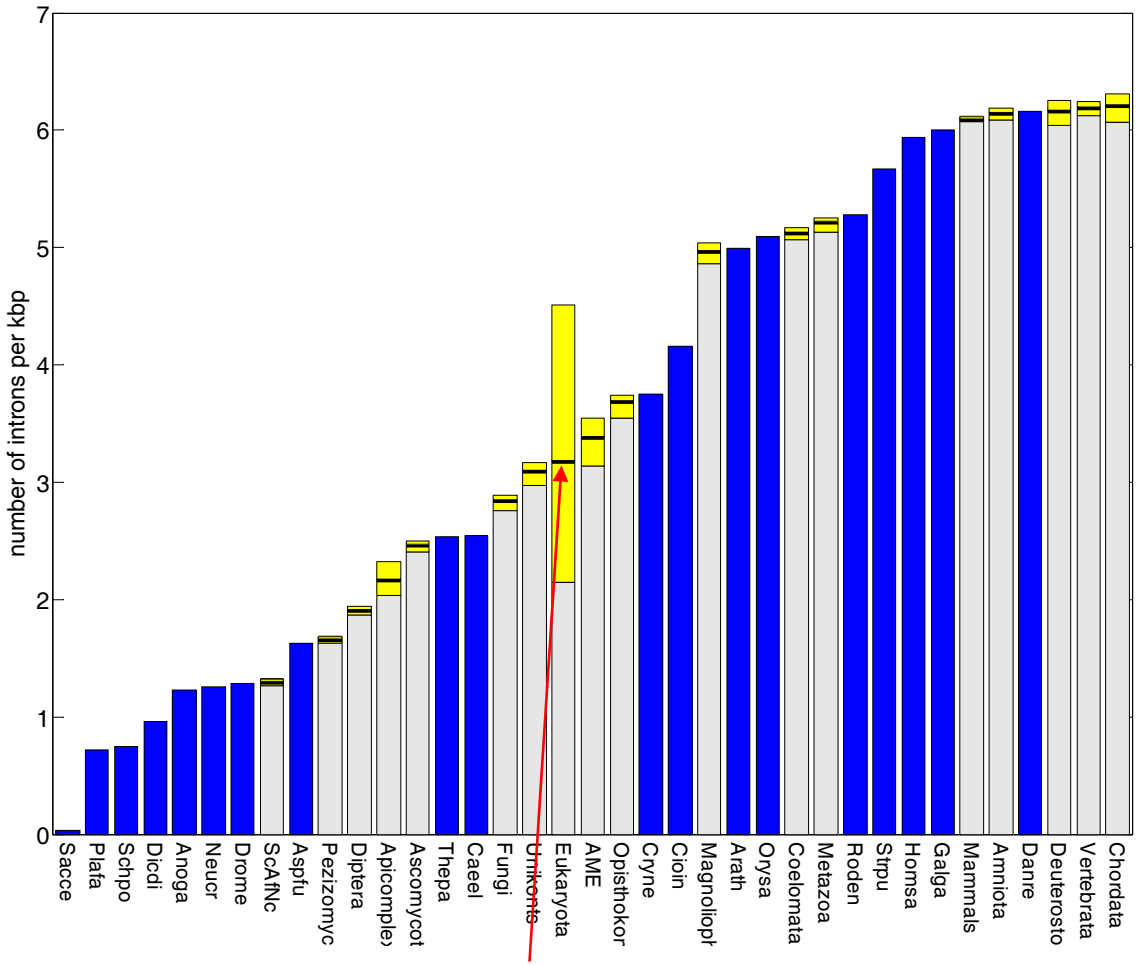


Early duplication of intronless genes + reverse-transcription-mediated duplication account for lack of intron position conservation in ancient paralogs

Sverdlov et al.
A glimpse of a putative pre-intron phase of eukaryotic evolution.
Trends Genet. 2007 Mar;23(3):105-8

Results of ML reconstruction of intron density in ancestral eukaryotes

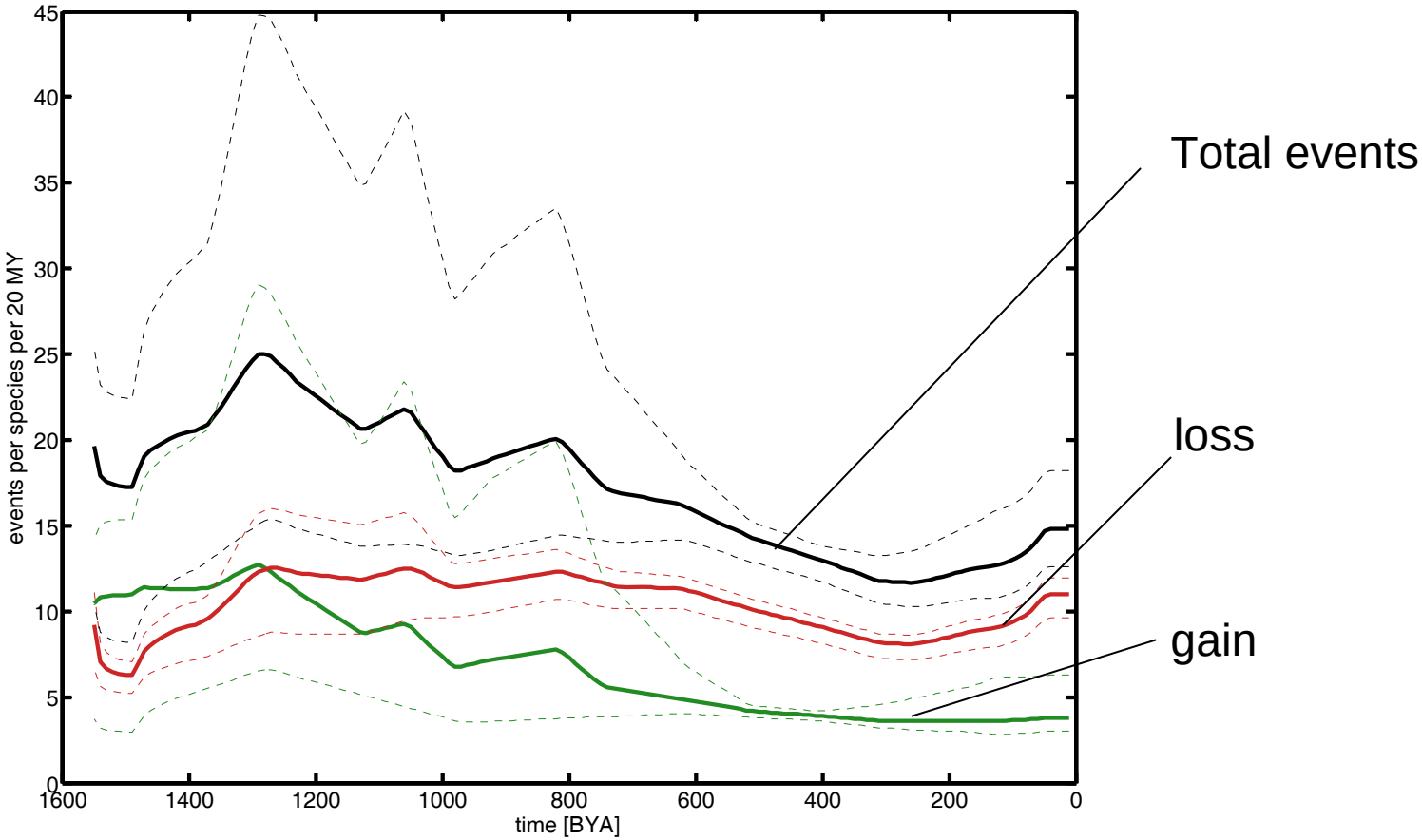
High intron density was reached early on in eukaryotic evolution



The last common ancestor of all extant eukaryotes (LECA)
had intron density comparable to the median of the modern forms

Carmel, Wolf, Rogozin, Koonin, submitted

Temporary dynamics of intron gain and loss from ML reconstructions



Evolution of eukaryotic genes after ~1.5 GYA is, largely, a story of intron loss

Carmel, Wolf, Rogozin, Koonin, submitted

Take home messages

- Eukaryogenesis was an incredible leap in the organizational complexity of the cell; no intermediates; mystery remains
- The most parsimonious current scenario is the invasion of an archaeon by an α-proteobacterium, the proto-mitochondrion
- Rapid invasion of the host genome by group II introns from the symbiont might have precipitated many if not all major events of eukaryogenes: emergence of nucleus, linear chromosomes, NMD, ubiquitin signaling system…?

ACKNOWLEDGEMENTS

Bill Martin (Univ. Duesseldorf)

Major collaborators in relevant research (NCBI): L. Aravind
Vivek Anantharaman
Liran Carmel
Kira Makarova
Ben Mans (currently NIAID)
Igor Rogozin
Alex Sverdlov (currently Columbia Univ)
Yuri Wolf