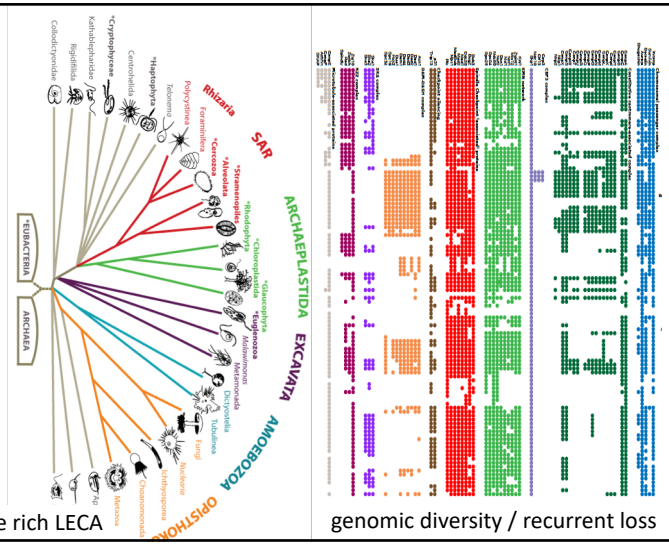
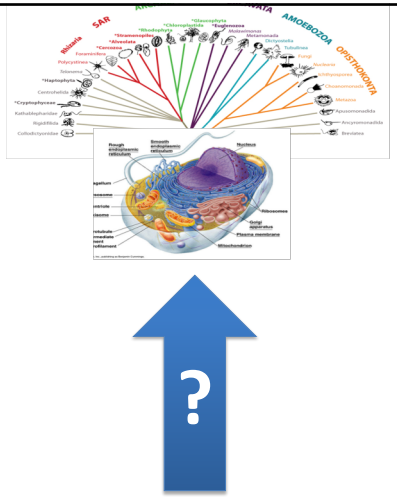
Eukaryogenesis, LECA



gene rich LECA

genomic diversity / recurrent loss

Eukaryogenesis
(how did we get such a complex LECA)



Importance of eukaryogenesis

"the basic divergence in cellular structure, which separates the bacteria and blue-green algae from all other cellular organisms, represents the greatest single evolutionary discontinuity to be found in the present-day world"

(Stanier *et al 1963)*

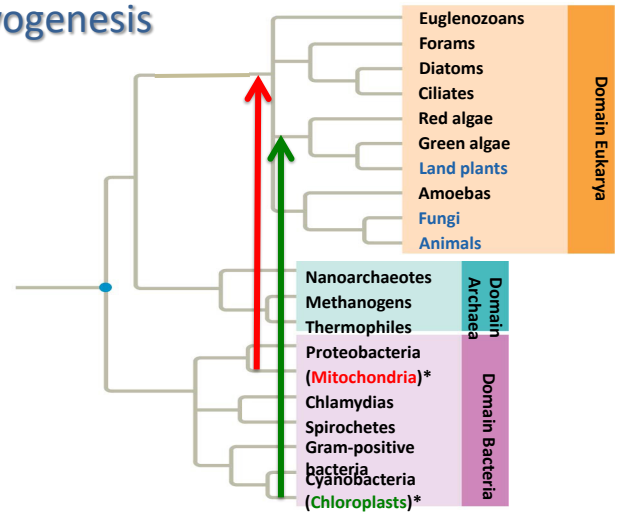
This radical transformation of cell structure (eukaryogenesis) is the most complex and extensive case of quantum evolution in the history of life [2,3,6]. Beforehand earth was a sexless, purely bacterial and viral world. Afterwards sexy, endoskeletal eukaryotes evolved morphological complexity: diatoms, butterflies, corals, whales, kelps, and trees.

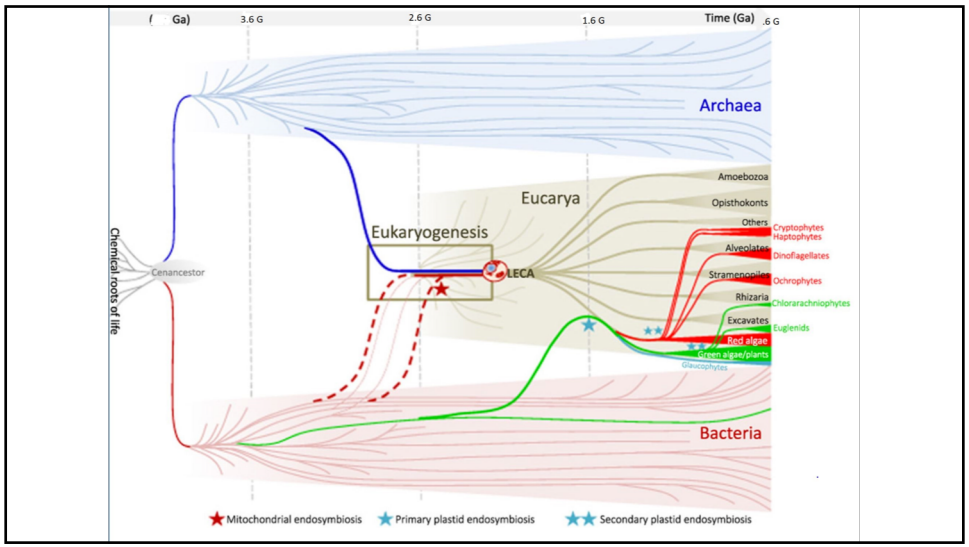
(Cavelier-Smith, 2010)

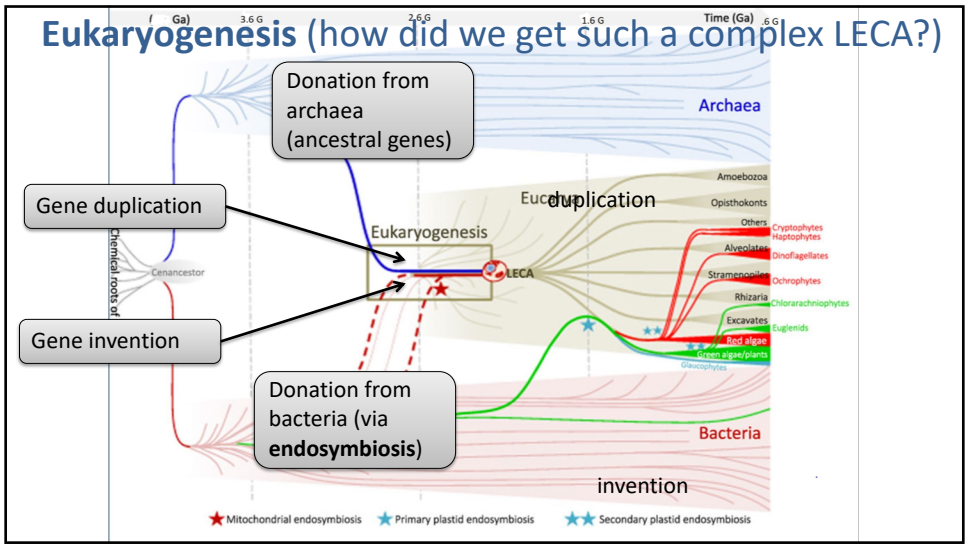
Stuff that changed

- Loss of operons
- Linearization of chromosomes, telomeres
- Uncoupling of transcription and translation
- Introns & splicing
- Nucleus
- Curtailing of horizontal gene transfer
- Membrane based organelles
- Meiosis / sexual reproduction
- New protein complexes and machines (the eukaryotic flagellum)
- New processes (phagocytosis, amoeboid movement, mitosis)
- New folds & functions
- Complexification (/ duplication) of existing complexes and machines (proteaosme, RNA polymerase)
- Mitochondrial endosymbiosis
- Dramatic increase in intrinsically disordered proteins

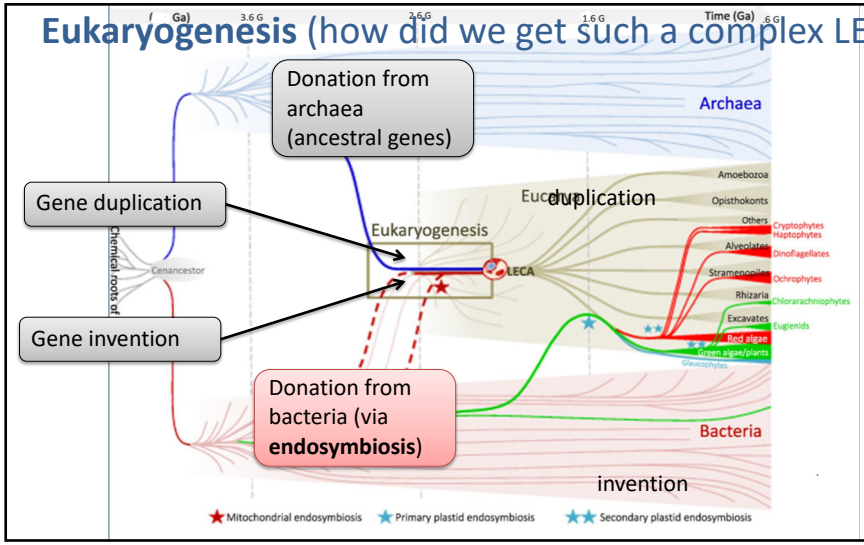
Eukaryogenesis



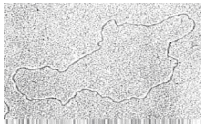




Eukaryogenesis (how did we get such a complex LECA?)



DNA



Mitchondria have their own chromosome … and this chromosome is circular and not enveloped in a "nucleus"

prokaryotes
Circular chromosomes, no organelles

eukaryotes
Linear chromosomes
organelles

"Theory of endosymbiosis"

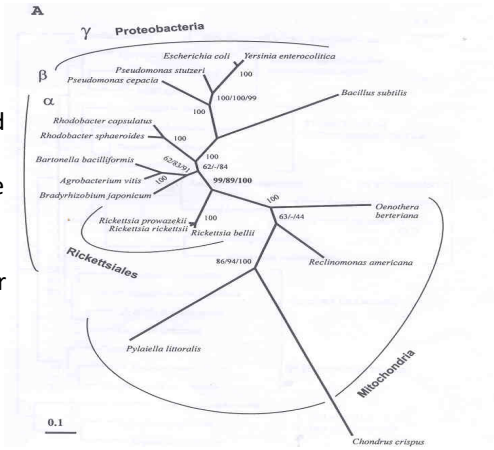
http://home.nc.rr.com/ambiient/site/mtdna.htm

Phylogenetic trees

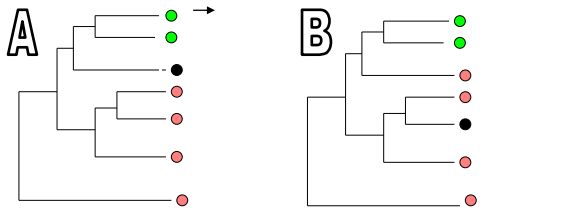
•Mitochondrial chromosome, genes, rRNA

•Similarity according to an established model of sequence change. Determine how organisms / genes are related: tree

•Tree: eukaryotic mitochondria cluster <u>within</u> bacteria, <u>within</u> alpha - proteobacteria, next to rickettsia, obligate intracellular parasites of eukaryotic cells



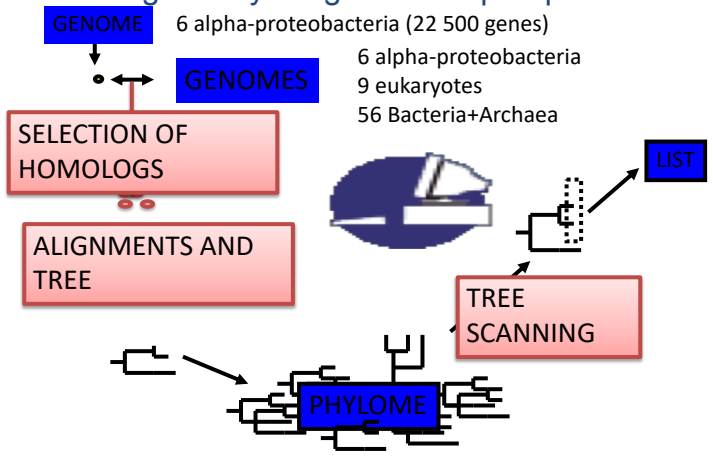
Identifying eukaryotic proteins with an alpha-proteobacterial origin based on their phylogeny



Eukaryotic + alpha-proteobacteria in the same branch

Alpha-proteobacterial proteins with the rest of the bacteria and archaea

Detecting eukaryotic genes of alpha-proteobacterial ancestry



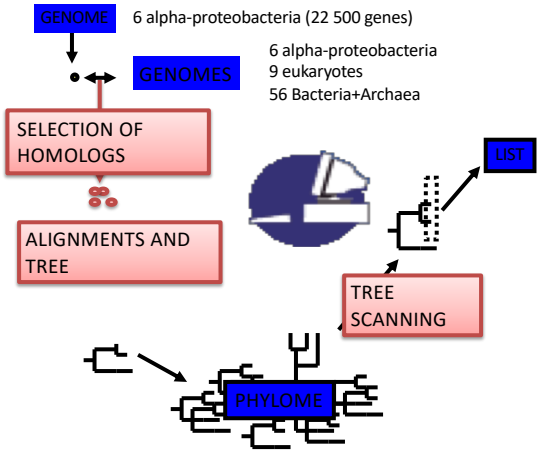
So quite a lot of proteins from alpha-prot: the vast majority of these are no longer encoded in the mitochondrial genome, endosymbiotic gene transfer

Species	Proteins[a]	NJ-Set[b]	ML-Set[b]
Reclinomonas americana (mitochondrion)	67	71.6	62.70
Deinococcus radiodurans	3,084	1.1	0.03
Agrobacterium tumefaciens (Cereon)	5,392	13.03	7.59
Agrobacterium tumefaciens (Washington)	5,298	13.11	11.76
Bradyrhizobium japonicum	8,257	11.18	8.25
Brucella melitensis	3,186	16.1	11.08
Brucella suis	3,247	15.86	9.67
Caulobacter crecentus	3,718	13.23	8.85
Magnetococcus magnetotacticum	4,280	11.36	8.74
Rhizobium loti	7,259	13.08	8.94
Rhizobium meliloti	6,149	13.67	9.17
Rickettsia conorii	1,374	20.3	16.59
Rickettsia prowazekii	834	25.06	19.78
Total selected[c]		1,026	842

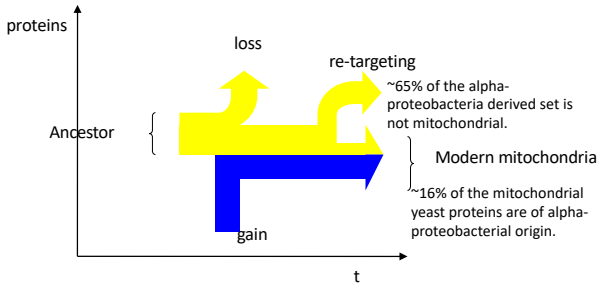
[a]Number of protein-coding genes per genome.
[b]Percentage of selected proteins in each genome by each approach.
[c]Total number of selected OGs.
An analysis that was based on six alpha-proteobacterial genomes [9] detected 630 proteins of alpha-proteobacterial origin that were in eukaryotes, and retrieved 49% of the *R. americana* mitochondrial genes and 1.3% of *D. radiodurans*. Increasing the number of analyzed genomes has thus substantially increased the number of proteins and the "completeness" of the proteome, while reducing the number of potential false positives.
doi:10.1371/journal.pcbi.0030219.t001

Number of proteins of alpha-proteobacterial descent

Detecting eukaryotic genes of alpha-proteobacterial ancestry



From endosymbiont to organell, not only loss and gain of proteins but also "retargeting":



Gabaldon and Huynen, Science 2004

LETTER

doi:10.1038/nature16941

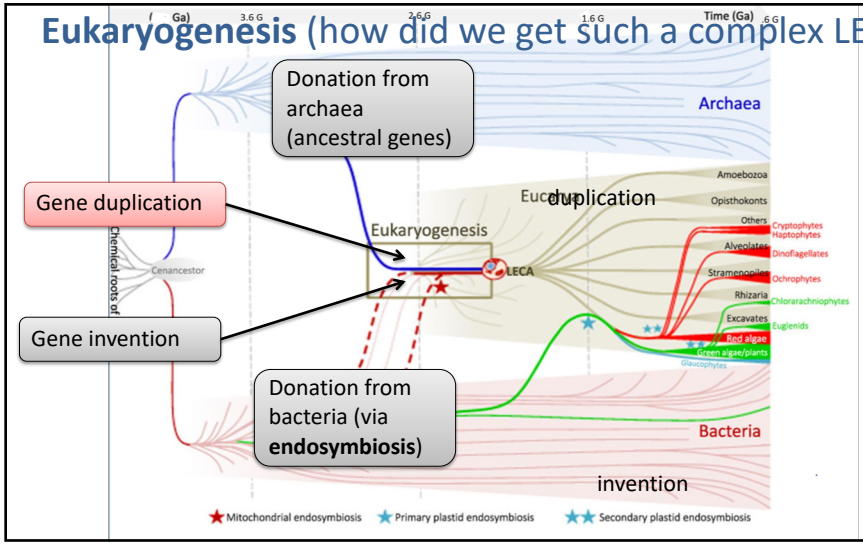
Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry

Alexandros A. Pittis[1,2] & Toni Gabaldón[1,2,3]

The origin of eukaryotes stands as a major conundrum in biology[1]. Current evidence indicates that the last eukaryotic common ancestor already possessed many eukaryotic hallmarks, including a complex subcellular organization[1–3]. In addition, the lack of evolutionary intermediates challenges the elucidation of the relative order of emergence of eukaryotic traits. Mitochondria are ubiquitous organelles derived from an alphaproteobacterial endosymbiont[4]. Different hypotheses disagree on whether mitochondria were acquired early or late during eukaryogenesis[5]. Similarly, the nature and complexity of the receiving host are debated, with models ranging from a simple prokaryotic host to an already complex proto-eukaryote[1,3,6,7]. Most competing scenarios can be roughly grouped into either mito-early, which consid… eukaryogenesis to be mi… host…

LECA family and connecting it to the last prokaryotic ancestor shared with its closest prokaryotic relatives (raw stem length; Fig. 1a). Branch lengths indicate the number of inferred substitutions per site, which reflect both divergence time and evolutionary rate. To disentangle time from rates, which may vary across families, we normalized the raw stem length by taking into account the median of the branch lengths within the LECA family (see Methods for further details). We used this measurement (hereafter referred to as stem length) as a proxy for the phylogenetic distance between a given LECA protein family… shared ancestor with prokaryotes. Com… hypotheses nat…

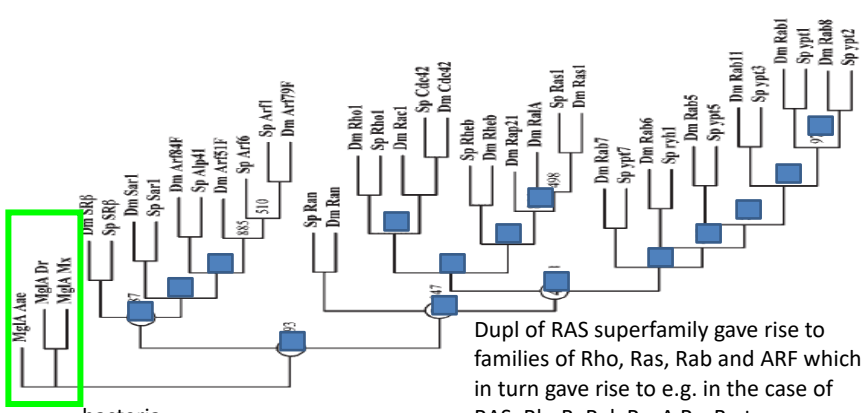
Eukaryogenesis (how did we get such a complex LECA?)



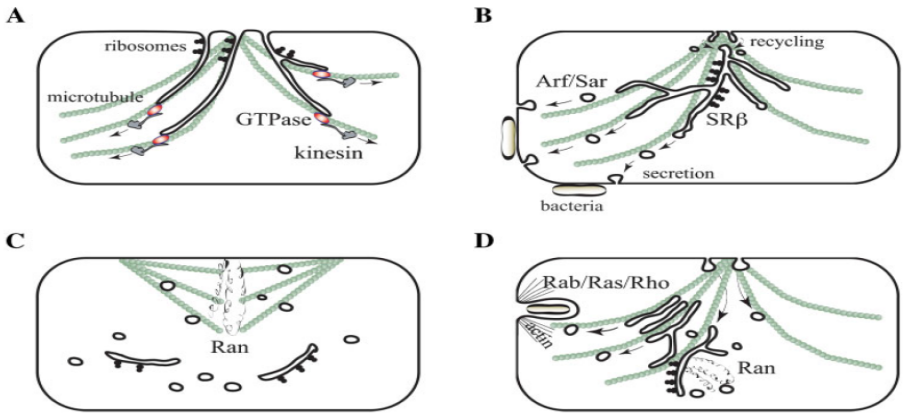
duplications: e.g. small GTPases

Small GTPases and the evolution of the eukaryotic cell

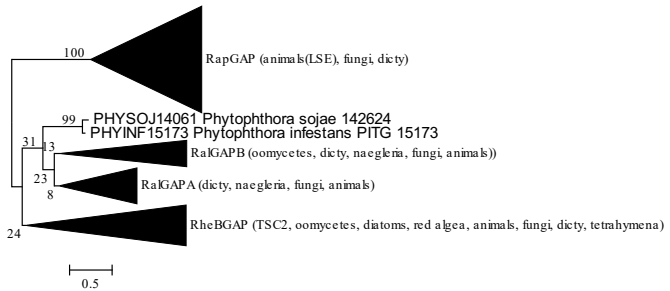
Gáspár Jékely[1,2]

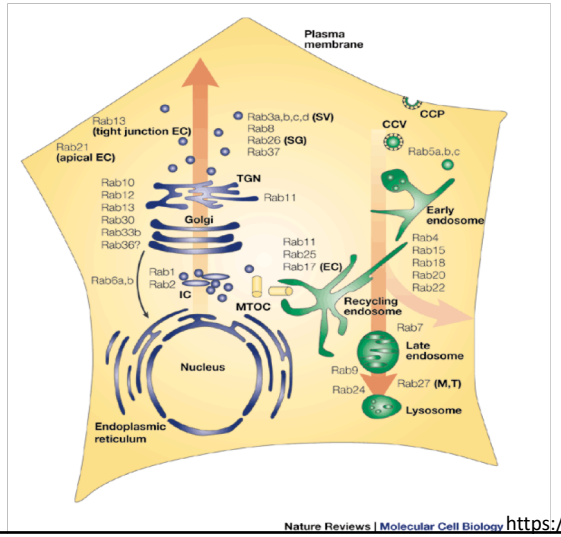


Dupl of RAS superfamily gave rise to families of Rho, Ras, Rab and ARF which in turn gave rise to e.g. in the case of RAS, RheB, Ral, RasA RasB etc.

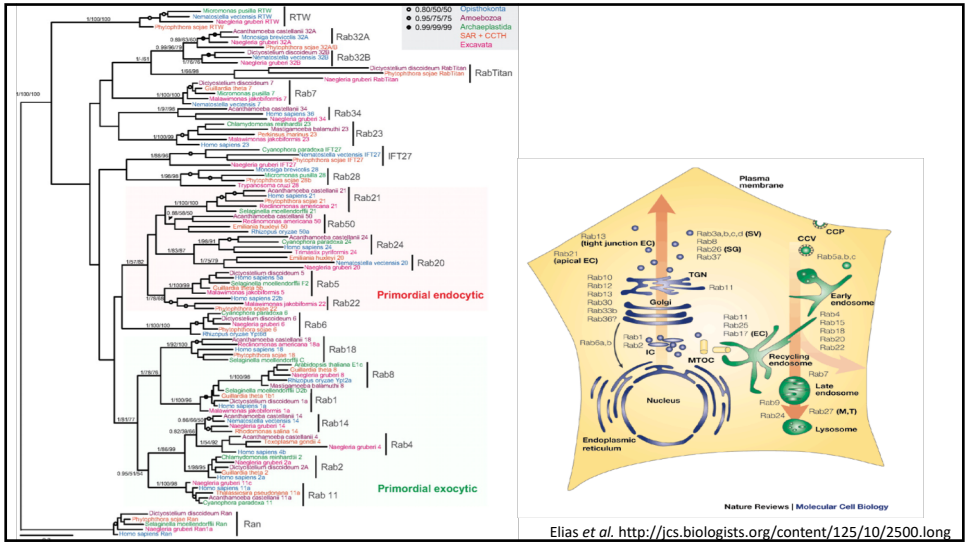


Not just the gtpases, also their activating proteins: Rap/Ral/RheB GAP tree: events from before the LECA



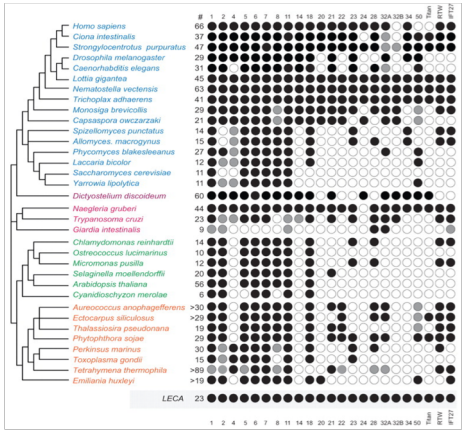


https://www.nature.com/articles/35052055/

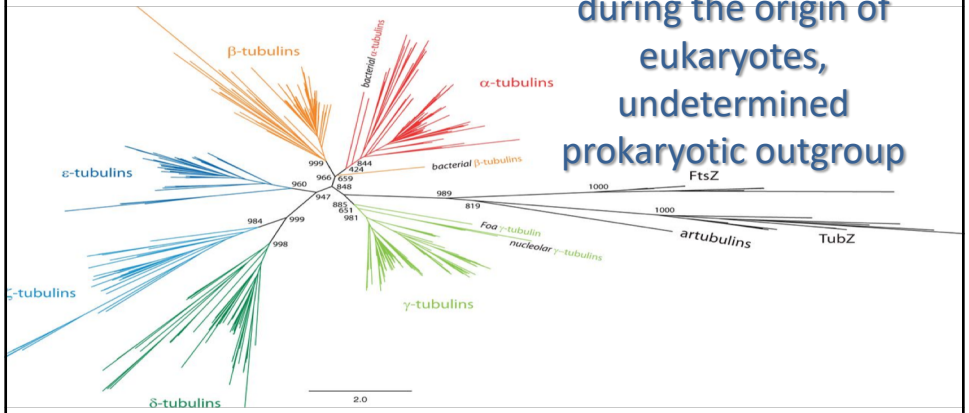


Elias *et al.* http://jcs.biologists.org/content/125/10/2500.long

NB1 These 23 LECA Rab's were differentially retained in present-day eukaryotes



Duplication of tubulins during the origin of eukaryotes, undetermined prokaryotic outgroup

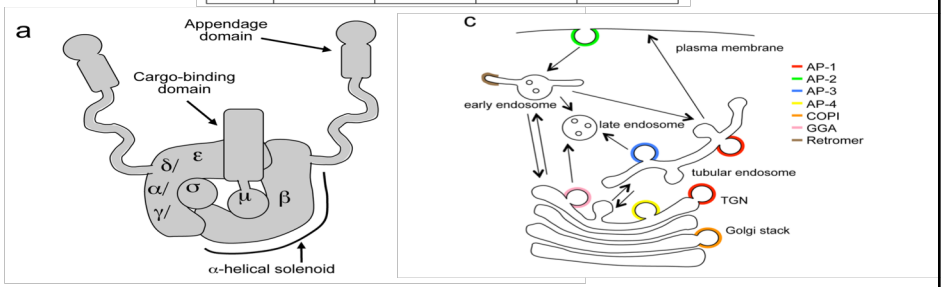


http://gbe.oxfordjournals.org/content/6/9/2274.long

- Adaptor protein (AP) complexes sort cargo into vesicles for transport from one membrane compartment of the cell to another
- Evolution of specificity in the eukaryotic endomembrane system.Dacks JB, Peden AA, Field MC. Int J Biochem Cell Biol. 2009 Feb;41(2):330-40.

Adaptor proteins arose via feca-2-leca duplications: *organellar paralogy*

Complex	Large chain 1	Large chain 2	Medium chain	Small chain
AP-1	γ	β1	μ1	σ1
AP-2	α	β2	μ2	σ2
AP-3	δ	β3	μ3	σ3
AP-4	ε	β4	μ4	σ4
F-COPI	γ-COP	β-COP	δ-COP	ζ-COP



Neo or sub functionalization … for membrane identity?

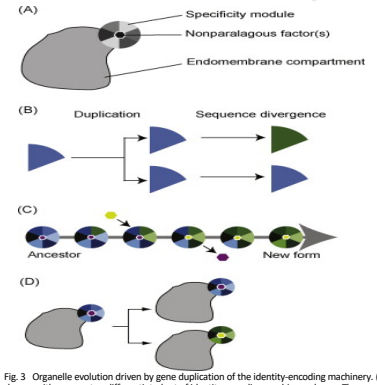
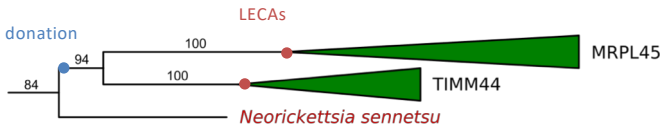


Fig. 3 Organelle evolution driven by gene duplication of the identity-encoding machinery. (A) An initial endomembraneous compartment is shown, with an as-yet undifferentiated set of identity-encoding machinery shown. The segmented circle indicate…

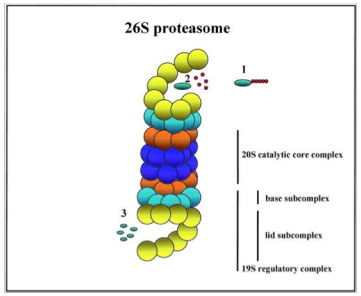
Parallels discussion for protein Complexes, e.g. stepwise duplication model

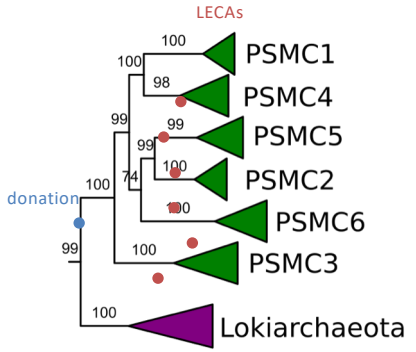
Also alpha-proteobacterial genes duplicated: TIMM44-MRPL45: 1 duplication

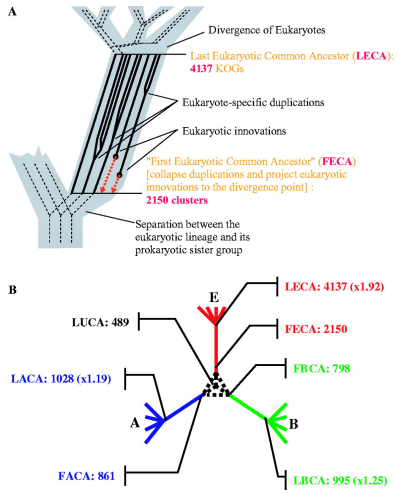


Duplications also for "complexification" of existing complexes: PSMC1-6, 5 duplications

- Proteasomal base: hexameric AAA ring



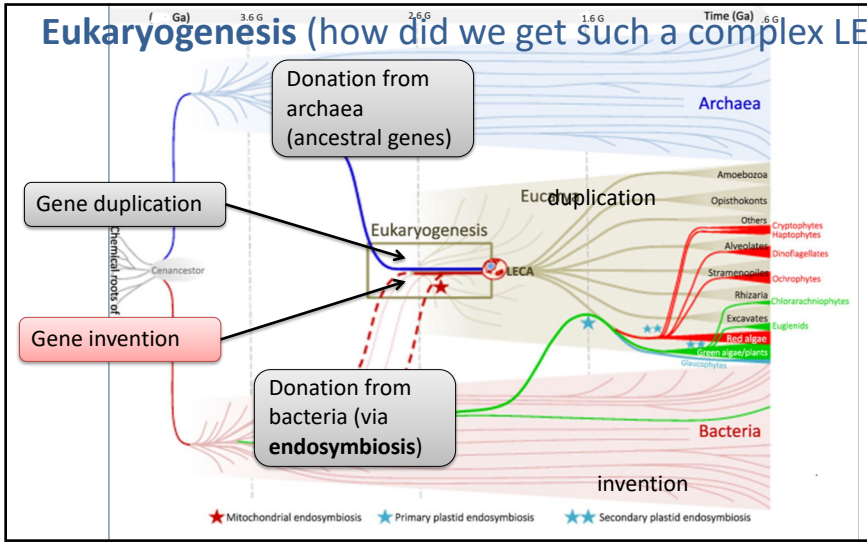




Quantitative estimates of duplication effect, almost a doubling, but no update for ±14 years?

Nucleic Acids Res. 2005 Aug 16;33(14):4626-38
Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell.
Makarova KS, Wolf YI, Mekhedov SL, Mirkin BG, Koonin EV.

Eukaryogenesis (how did we get such a complex LECA?)



Genes / Folds invented during eukaryogenesis

Received: 30 July 2017 | Revised: 17 April 2018 | Accepted: 18 April 2018

DOI: 10.1002/prot.25517

WILEY PROTEINS

RESEARCH ARTICLE

Eukaryote specific folds: Part of the whole

Anni Kauko | Kirsi Lehto

Department of Biochemistry, University of Turku, Turku, Finland

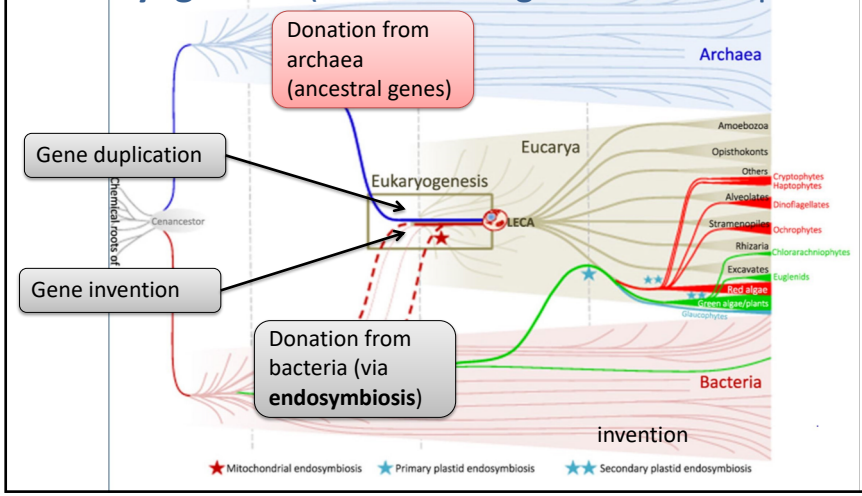
Abstract

The origin of eukaryotes is one of the central transitions in the history of life; without eukaryotes there would be no complex multicellular life. The most accepted scenarios suggest the endosymbiosis of a mitochondrial ancestor with a complex archaeon, even though the details regarding the host and the triggering factors are still being discussed. Accordingly, phylogenetic analyses have

Genes / Folds "invented" during eukaryogenesis

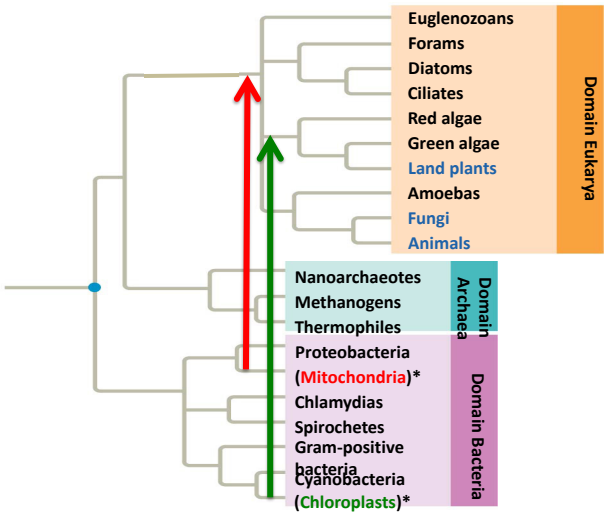
- ±500 new folds!
- Many of them are all-alpha
- Regulatory or interaction function, not enzymatic functions
- A lot these new folds subsequently duplicated!

Eukaryogenesis (how did we get such a complex LECA?)

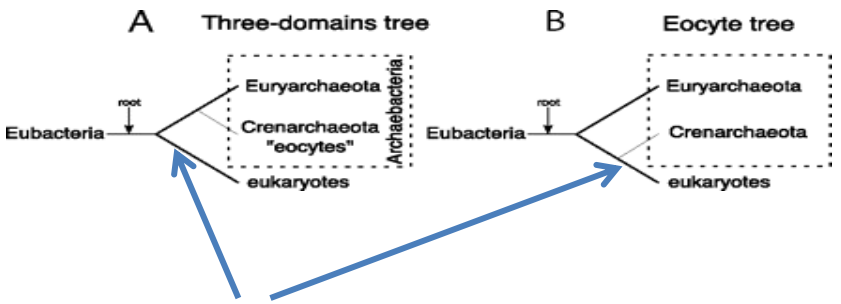


Where do eukaryotes go in the tree of life?

- Text book 3 domains of life

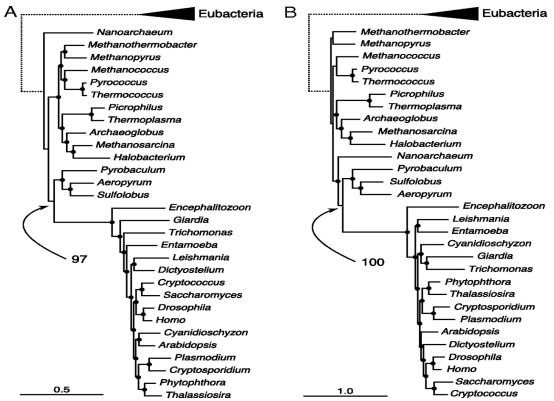


… but there is the old Eocyte hypothesis

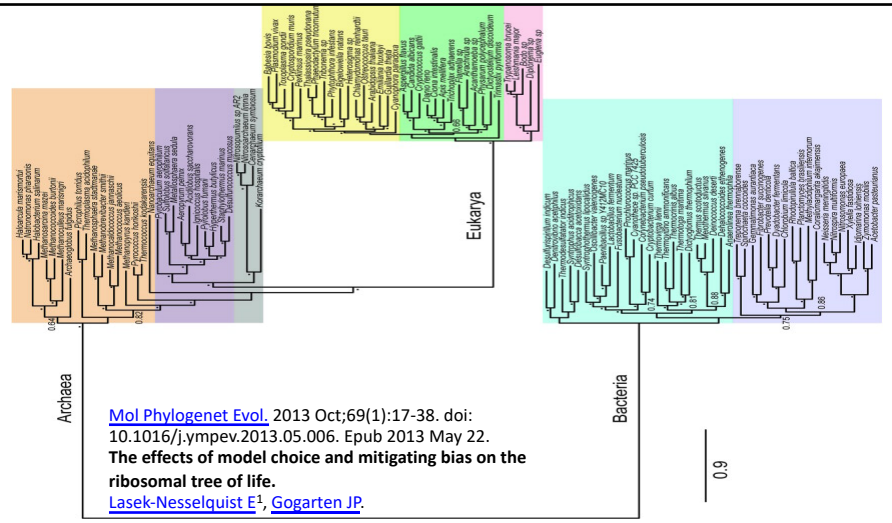


First Eukaryotic Common Ancestor (FECA) / proto eukaryote

some trees suggest eocyte (2 kingdoms, not 3)



Proc Natl Acad Sci U S A. 2008 Dec 23;105(51):20356-61. Epub 2008 Dec 10.
The archaebacterial origin of eukaryotes.
Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM.

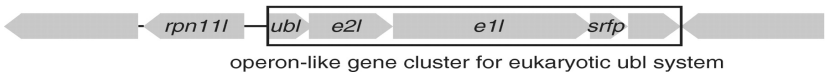


Mol Phylogenet Evol. 2013 Oct;69(1):17-38. doi: 10.1016/j.ympev.2013.05.006. Epub 2013 May 22.
The effects of model choice and mitigating bias on the ribosomal tree of life.
Lasek-Nesselquist E[1], Gogarten JP.

Novel archaea has operon with UBQ system

- Insights into the evolution of **Archaea** and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. Nunoura T, Takaki Y, Kakuta J, Nishi S, Sugahara J, Kazama H, Chee GJ, Hattori M, Kanai A, Atomi H, Takai K, Takami H. Nucleic Acids Res. 2011 Apr;39(8):3204-23

The gene cluster of the Ub-like protein modifier system in *C. subterraneum.* , eukaryotic "type" ubiquitin



operon-like gene cluster for eukaryotic ubl system

Nunoura T et al. Nucl. Acids Res. 2011;39:3204-3223

Nucleic Acids Research

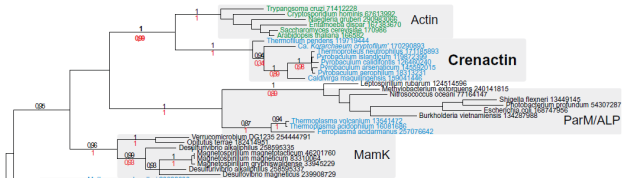
Molecular Microbiology (2011) 80(4), 1052–1061 ■

doi:10.1111/j.1365-2958.2011.07635.x
First published online 6 April 2011

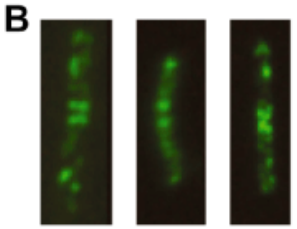
An actin-based cytoskeleton in archaea

Thijs J. G. Ettema,*† Ann-Christin Lindås† and Rolf Bernander
Department of Molecular Evolution, Evolutionary Biology Center, Uppsala University, Norbyvägen 18C, SE-752 36, Uppsala, Sweden.

ments twisted around one another to form a right-handed double helix, which constitute well-characterized central components of cytoskeleton-dependent processes (Pollard and Cooper, 2009; van den Ent *et al.*, 2001). As a result of the involvement in pivotal processes, the primary structure of actin has been extremely well conserved



1:1 Orthologous to eukaryotic actin with limited phylogenetic dsitrubution in archaea



Examples of subpopulation of cells displaying centrally located band-like structures.

Cytokinesis?

Eukaryotic features in archaea are present in subclade of archaea (TACK) where also now some ToL places the eukaryotes

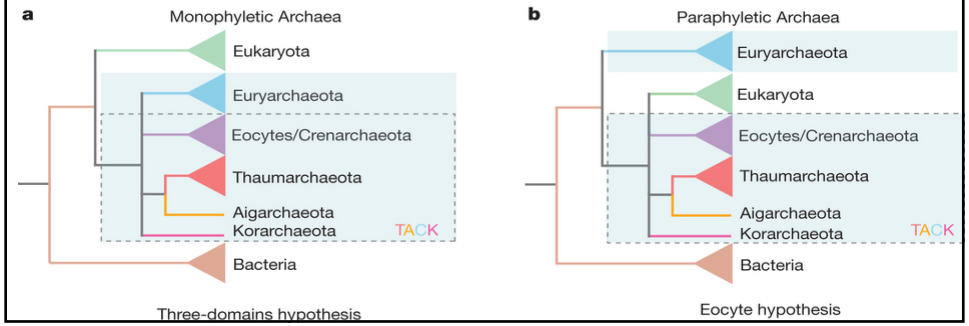
Proto-eukaryote (FECA) is getting a little bit more complex as more archaeal diversity is sequenced and bioinformatically and biochemically characterized

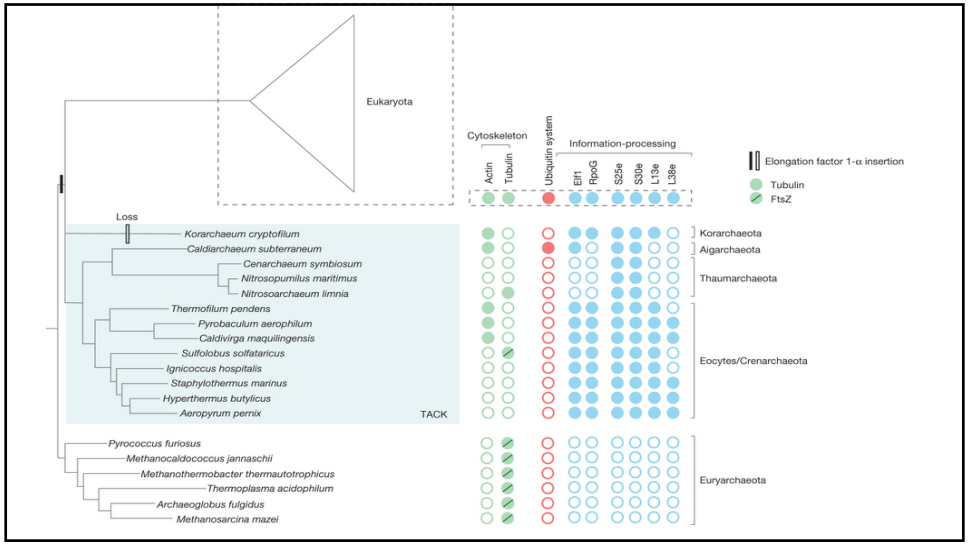
REVIEW

doi:10.1038/nature12779

An archaeal origin of eukaryotes supports only two primary domains of life

Tom A. Williams[1], Peter G. Foster[2], Cymon J. Cox[3] & T. Martin Embley[1]





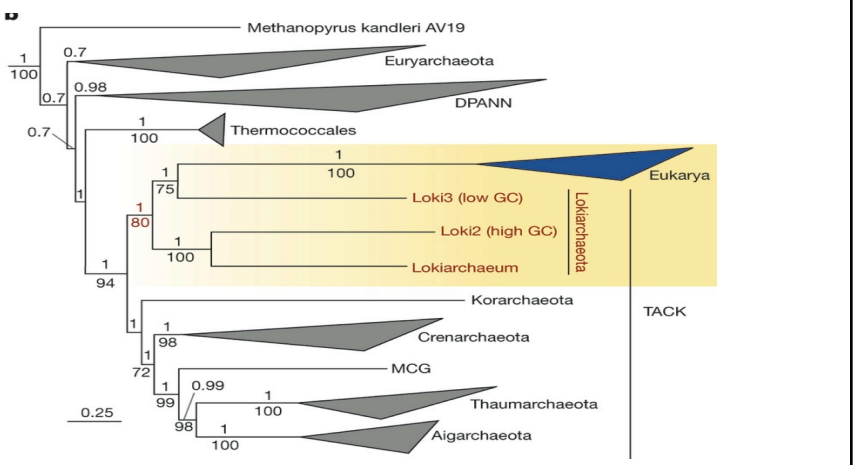
ARTICLE

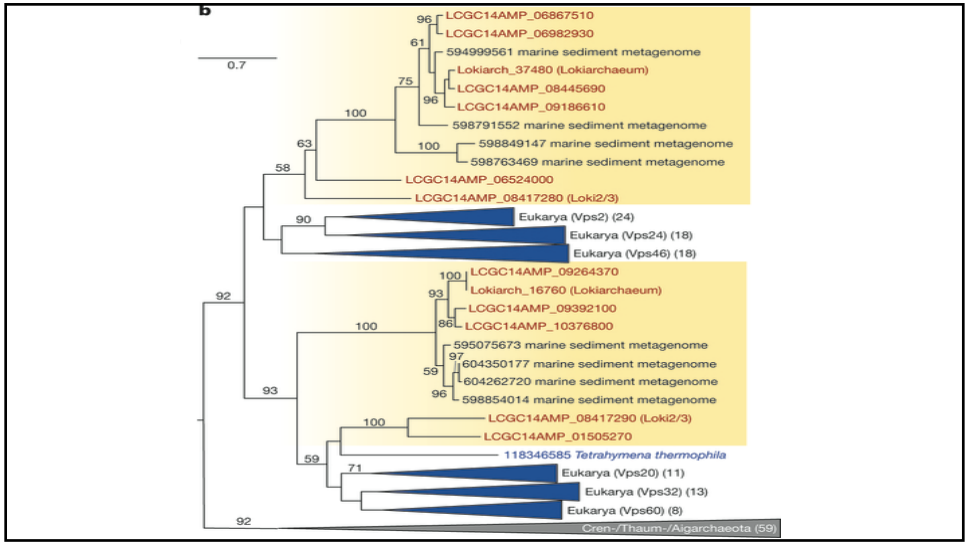
doi:10.1038/nature14447

Complex archaea that bridge the gap between prokaryotes and eukaryotes

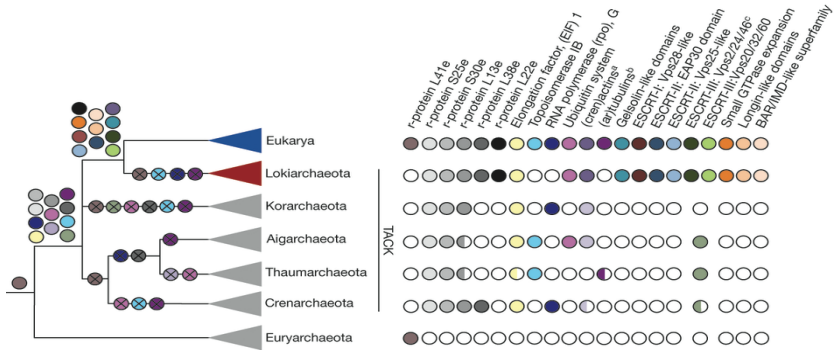
Anja Spang[1]*, Jimmy H. Saw[1]*, Steffen L. Jørgensen[2]*, Katarzyna Zaremba-Niedzwiedzka[1]*, Joran Martijn[1], Anders E. Lind[1], Roel van Eijk[1]†, Christa Schleper[2,3], Lionel Guy[1,4] & Thijs J. G. Ettema[1]

The origin of the eukaryotic cell remains one of the most contentious puzzles in modern biology. Recent studies have provided support for the emergence of the eukaryotic host cell from within the archaeal domain of life, but the identity and nature of the putative archaeal ancestor remain a subject of debate. Here we describe the discovery of 'Lokiarchaeota', a novel candidate archaeal phylum, which forms a monophyletic group with eukaryotes in phylogenomic analyses, and whose genomes encode an expanded repertoire of eukaryotic signature proteins that are suggestive of sophisticated membrane remodelling capabilities. Our results provide strong support for hypotheses in which the eukaryotic host evolved from a bona fide archaeon, and demonstrate that many components that underpin eukaryote-specific features were already present in that ancestor. This provided the host with a rich genomic 'starter-kit' to support the increase in the cellular and genomic complexity that is characteristic of eukaryotes.

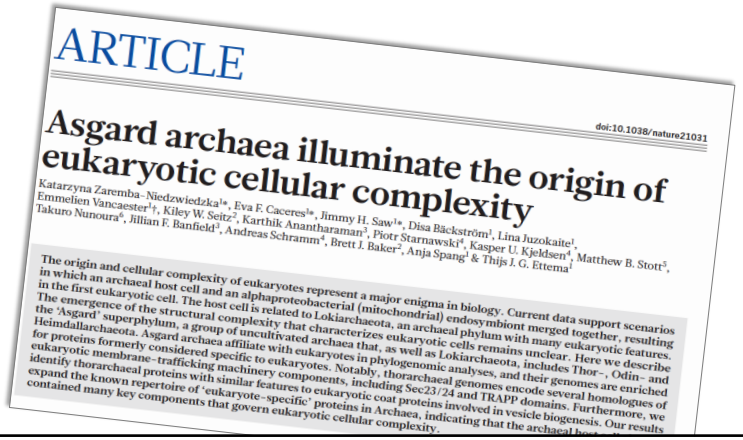




“eukaryotic signature proteins”



Even more / better …



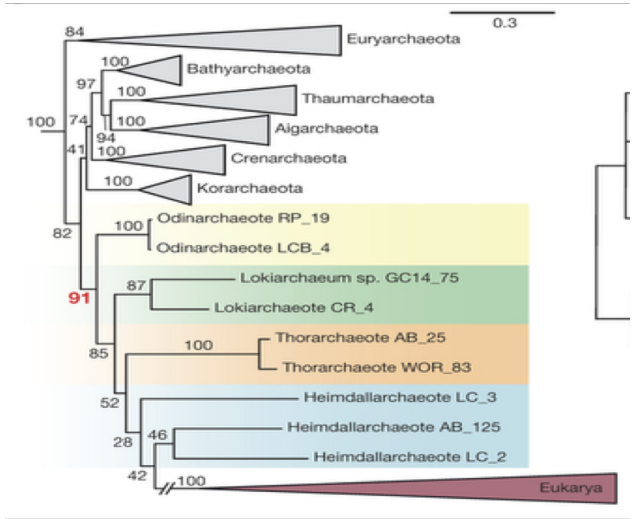
ARTICLE

doi:10.1038/nature21031

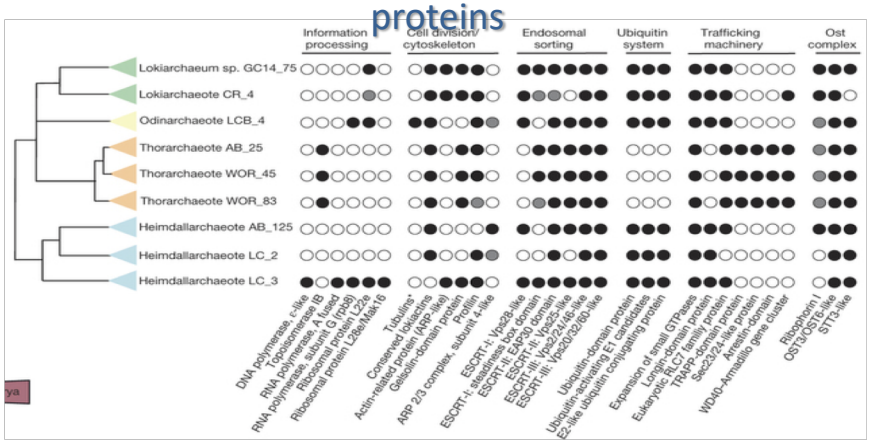
Asgard archaea illuminate the origin of eukaryotic cellular complexity

Katarzyna Zaremba-Niedzwiedzka[1]*, Eva F. Caceres[1]*, Jimmy H. Saw[1]*, Disa Bäckström[1], Lina Juzokaite[1], Emmelien Vancaester[1†], Kiley W. Seitz[2], Karthik Anantharaman[3], Piotr Starnawski[4], Kasper U. Kjeldsen[4], Matthew B. Stott[5], Takuro Nunoura[6], Jillian F. Banfield[3], Andreas Schramm[4], Brett J. Baker[2], Anja Spang[1] & Thijs J. G. Ettema[1]

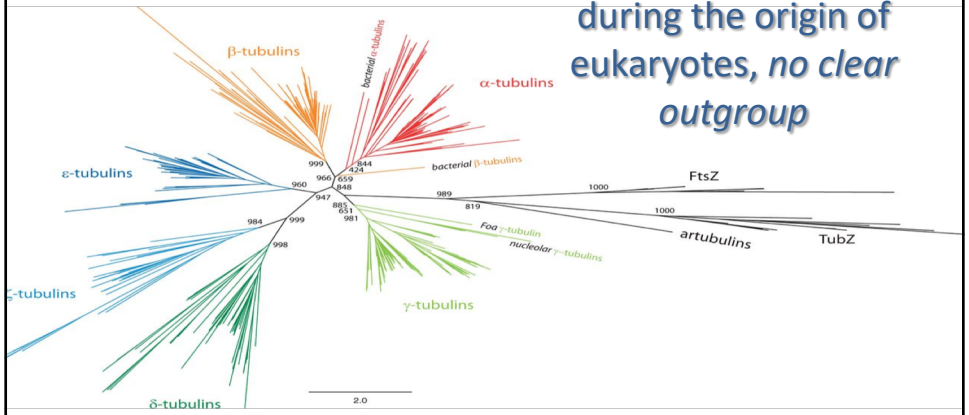
The origin and cellular complexity of eukaryotes represent a major enigma in biology. Current data support scenarios in which an archaeal host cell and an alphaproteobacterial (mitochondrial) endosymbiont merged together, resulting in the first eukaryotic cell. The host cell is related to Lokiarchaeota, an archaeal phylum with many eukaryotic features. The emergence of the structural complexity that characterizes eukaryotic cells remains unclear. Here we describe the ‘Asgard’ superphylum, a group of uncultivated archaea that, as well as Lokiarchaeota, includes Thor-, Odin- and Heimdallarchaeota. Asgard archaea affiliate with eukaryotes in phylogenomic analyses, and their genomes are enriched for proteins formerly considered specific to eukaryotes. Notably, thorarchaeal genomes encode several homologues of eukaryotic membrane-trafficking machinery components, including Sec23/24 and TRAPP domains. Furthermore, we identify thorarchaeal proteins with similar features to eukaryotic coat proteins involved in vesicle biogenesis. Our results expand the known repertoire of ‘eukaryote-specific’ proteins in Archaea, indicating that the archaeal host contained many key components that govern eukaryotic cellular complexity.



Even more (previously) eukaryotic signature proteins

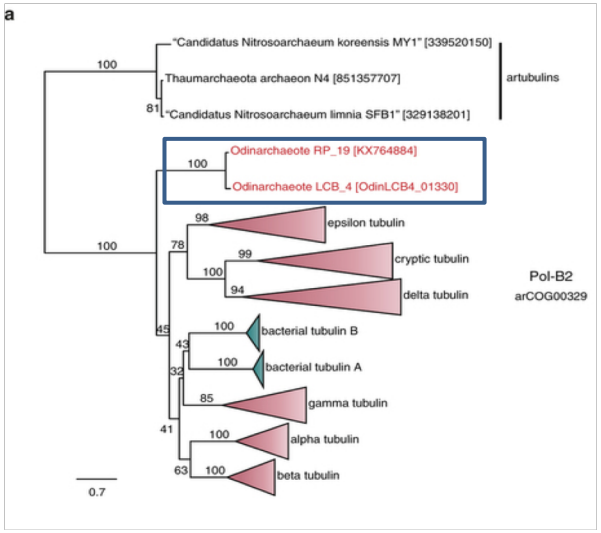


Duplication of tubelins during the origin of eukaryotes, *no clear outgroup*



http://gbe.oxfordjournals.org/content/6/9/2274.long

Asgard archaea provide clear outgroup for tubulin feca-2-leca duplications



biochemical study of asgard ESP proteins

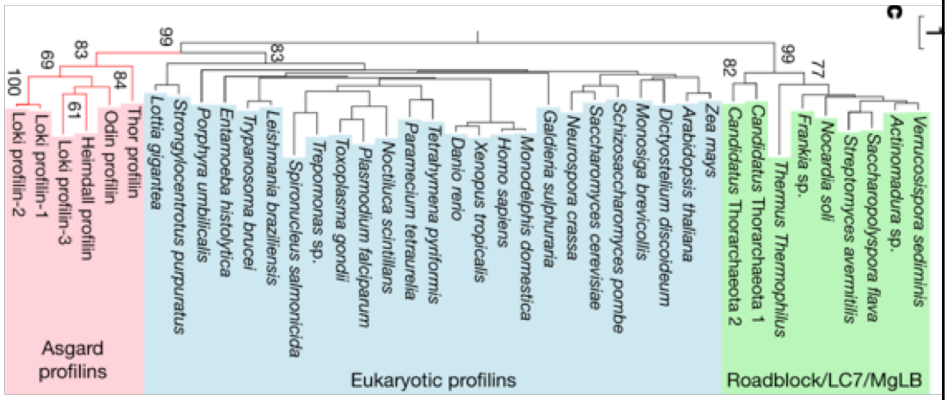
LETTER

https://doi.org/10.1038/s41586-018-0548-6

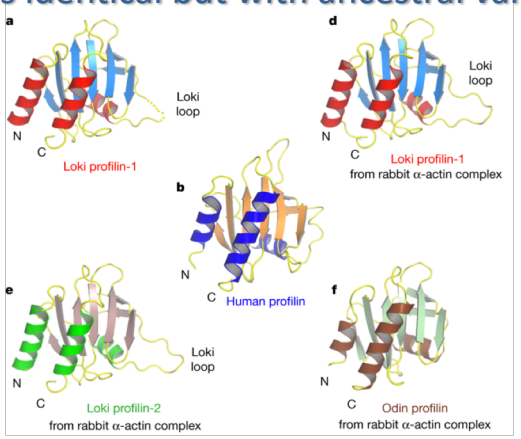
Genomes of Asgard archaea encode profilins that regulate actin

Caner Akıl[1,7] & Robert C. Robin

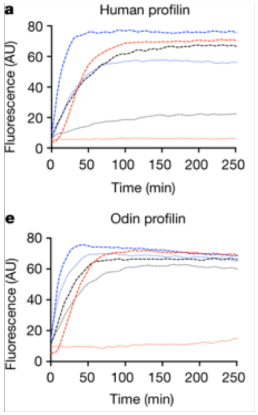
"Here we show that Asgard archaea encode functional profilins and thereby establish that this archaeal superphylum has a regulated actin cytoskeleton, one of the hallmarks of the eukaryotic cell."

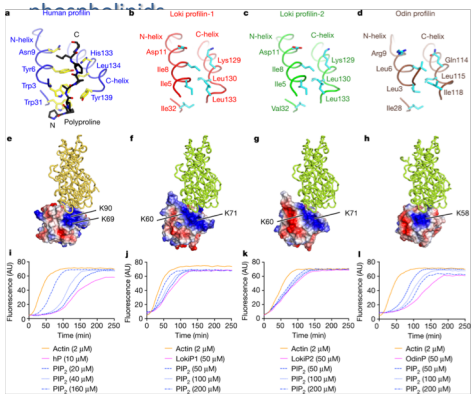


Structure is identical but with ancestral variations

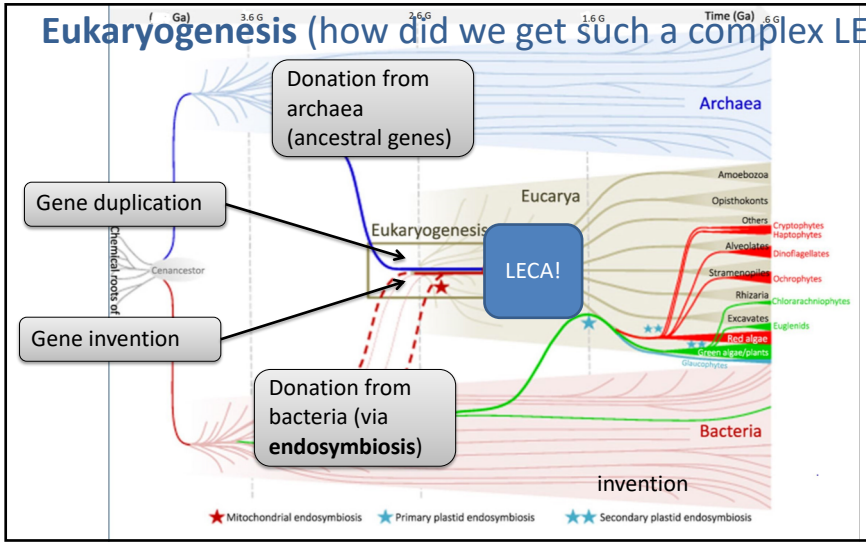


Asgard profilins modulate polymerization of mammalian actin in vitro yet Asgard profilins do not bind to polyproline motifs but are sensitive to





Eukaryogenesis (how did we get such a complex LECA?)



Importance of eukaryogenesis, what was the end result?

This radical transformation of cell structure (eukaryogenesis) is the most complex and extensive case of quantum evolution in the history of life [2,3,6]. Beforehand earth was a sexless, purely bacterial and viral world. Afterwards sexy, endoskeletal eukaryotes evolved morphological complexity: diatoms, butterflies, corals, whales, kelps, and trees.
(Cavelier-Smith, 2010)

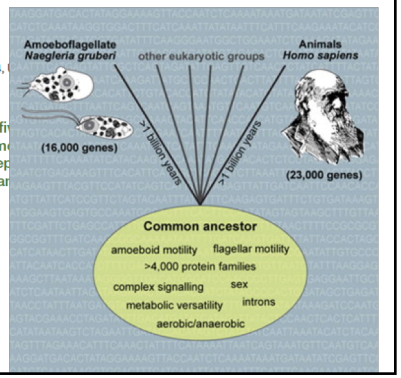
Cell

Leading Edge
Previews

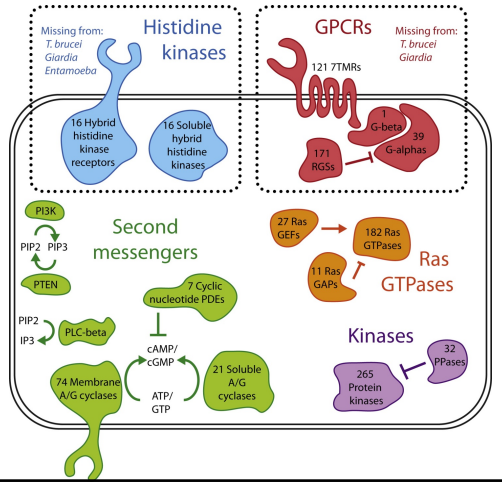
The Incredible Expanding Ancestor of Eukaryotes

Eugene V. Koonin[1,*]
[1]National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD 20894,
*Correspondence: koonin@ncbi.nlm.nih.gov
DOI 10.1016/j.cell.2010.02.022

Comparing the genome sequences of free-living organisms in the fi
enables predictions to be made about the genome of the last comm
The genome sequence of the amoeboflagellate *Naegleria gruberi* rep
(2010) reveals the surprising complexity of this unicellular organism ar
common eukaryotic ancestor.



Signalling complexity



End result; gene content of LECA

Published online 29 October 2018 *Nucleic Acids Research, 2019, Vol. 47, Database issue* **D271–D279**
doi: 10.1093/nar/gky1009

Ancestral Genomes: a resource for reconstructed ancestral genes and genomes across the tree of life

Xiaosong Huang[1,2], Laurent-Philippe Albou[2], Tremayne Mushayahama[2], Anushya Muruganujan[2], Haiming Tang[2] and Paul D. Thomas[2,*]

[1]School of Life Sciences, Guangzhou University, Guangzhou 510006, China and [2]Division of Bioinformatics, Department of Preventive Medicine, Keck School of Medicine of USC, University of Southern California, Los Angeles, CA 90033, USA

Received August 16, 2018; Revised September 27, 2018; Editorial Decision October 09, 2018; Accepted October 10, 2018

ABSTRACT

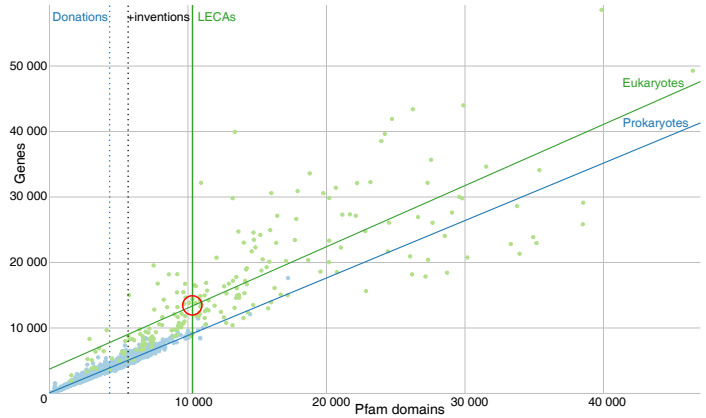
A growing number of whole genome sequencing projects, in combination with development of phylogenetic methods for reconstructing gene evolution, have provided us with a window into genomes that existed millions, and even billions, of years ago. Ancestral Genomes (http://ancestralgenomes.org) is a resource for comprehensive reconstructions of these 'fossil genomes'. Comprehensive sets of protein-coding genes have been reconstructed for 78 genomes of now-extinct species that were the common ancestors of extant species from across the tree of life. The reconstructed genes are based on the extensive library of over 15 000 gene fam-

organisms that are inferred to derive from a single gene in the common ancestor of those organisms—have been used to infer the gene content of ancient ancestral genomes, such as that of the last universal common ancestor (1) (~4 billion years ago) and the last eukaryotic common ancestor (2) (~1.8 billion years ago). More recently, the development of tree reconciliation methods opens the door to ancestral genome reconstruction at any branch point in the tree of life (3,4).

Reconciled trees combine the information in the gene tree, usually obtained from protein sequences of related genes in different organisms, with prior knowledge of the species tree that relates those organisms. Because of this property, each node in a reconciled gene tree can be labeled with the evolutionary 'event' type that separated related genes: speciation, gene duplication, and horizontal

"7178 genes"

LECA: a 'normal' eukaryote



After LECA major fate is loss

ARTICLE
doi:10.1038/nature14963

Endosymbiotic origin and differential loss of eukaryotic genes

Chuan Ku[1], Shijulal Nelson-Sathi[1], Mayo Roettger[1], Filipa L. Sousa[1], Peter J. Lockhart[2], David Bryant[3], Einat Hazkani-Covo[4], James O. McInerney[5,6], Giddy Landan[7] & William F. Martin[1,8]

Chloroplasts arose from cyanobacteria, mitochondria arose from proteobacteria. Both organelles have conserved their prokaryotic biochemistry, but their genomes are reduced, and most organelle proteins are encoded in the nucleus. Endosymbiotic theory posits that bacterial genes in eukaryotic genomes entered the eukaryotic lineage via organelle ancestors. It predicts episodic influx of prokaryotic genes into the eukaryotic lineage, with acquisition corresponding to endosymbiotic events. Eukaryotic genome sequences, however, increasingly implicate lateral gene transfer, both from prokaryotes to eukaryotes and among eukaryotes, as a source of gene content variation in eukaryotic genomes, which predicts continuous, lineage-specific acquisition of prokaryotic genes in divergent eukaryotic groups. Here we discriminate between these two alternatives by clustering and phylogenetic analysis of eukaryotic gene families having prokaryotic homologues. Our results indicate (1) that gene transfer from bacteria to eukaryotes is episodic, as revealed by gene distributions, and coincides with major evolutionary transitions at the origin of chloroplasts and mitochondria; (2) that gene inheritance in eukaryotes is vertical, as revealed by extensive topological comparison, sparse gene distributions stemming from differential loss; and (3) that continuous, lineage-specific lateral gene transfer, although it sometimes occurs, does not contribute to long-term gene content evolution in eukaryotic genomes.

After LECA major fate is loss