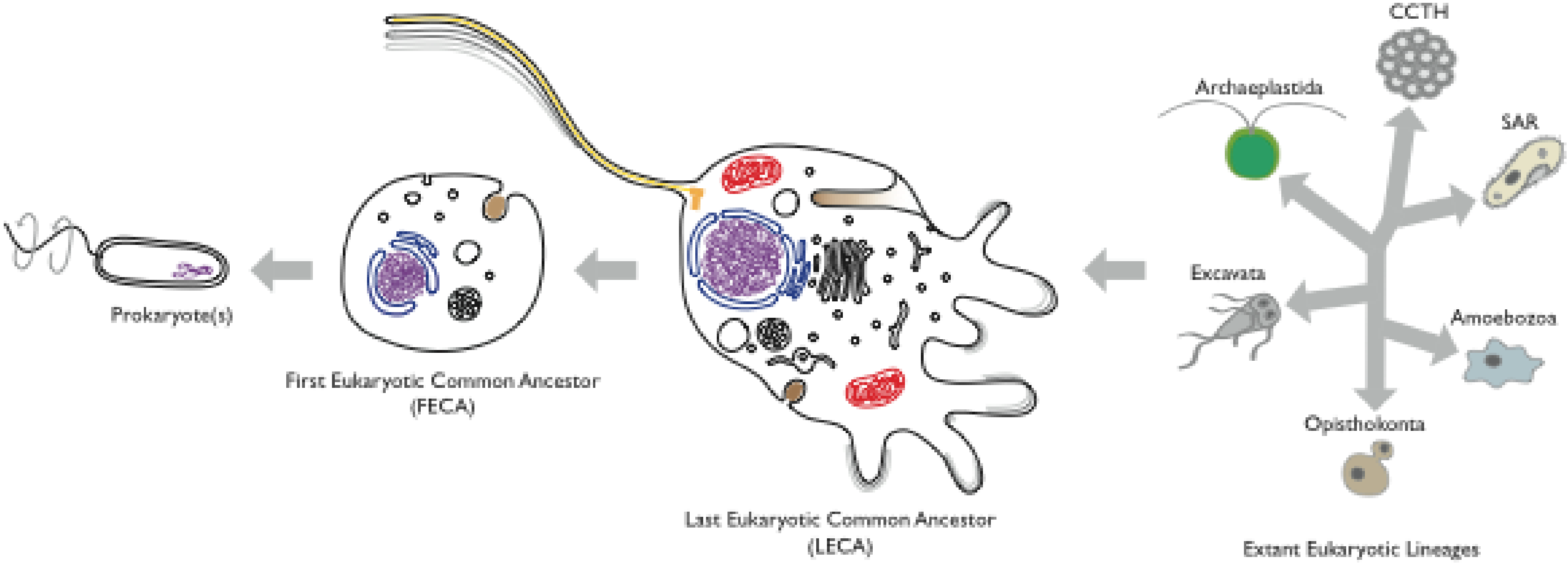
Origin of eukaryotes

Joel Dacks, Univ. Alberta

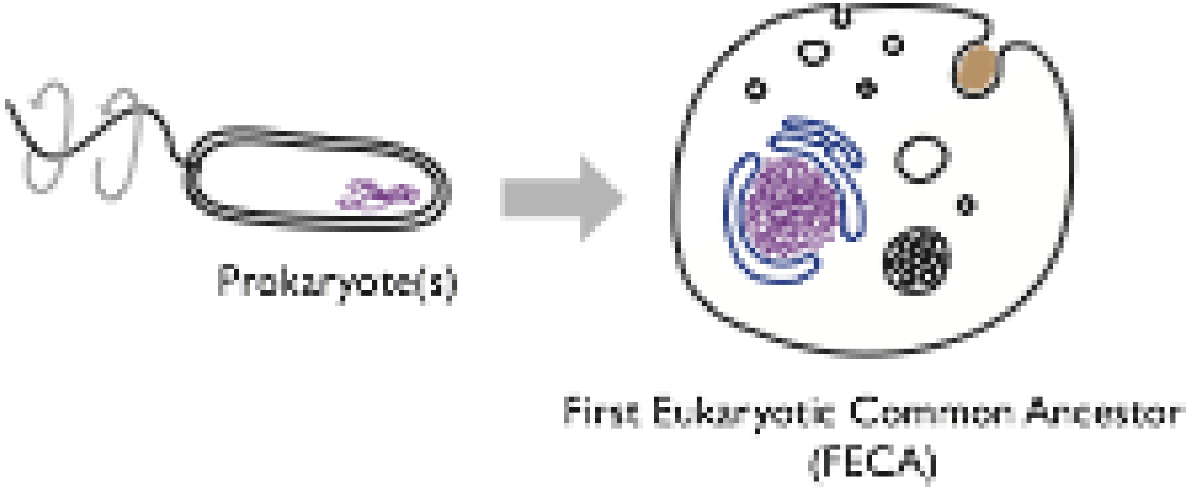


Aug 11, 2015

KITP EvoCell 15

The prokaryote to eukaryote transition

- Origin and timing of diverse eukaryotic features
- It is generally accepted that the eukaryotic cellular configuration needs to be explained from non-eukaryotic state: Microfossils, biomarkers, molecular clocks, chimeric nature of genome and cell biology



Archael contribution to eukaryotic origins

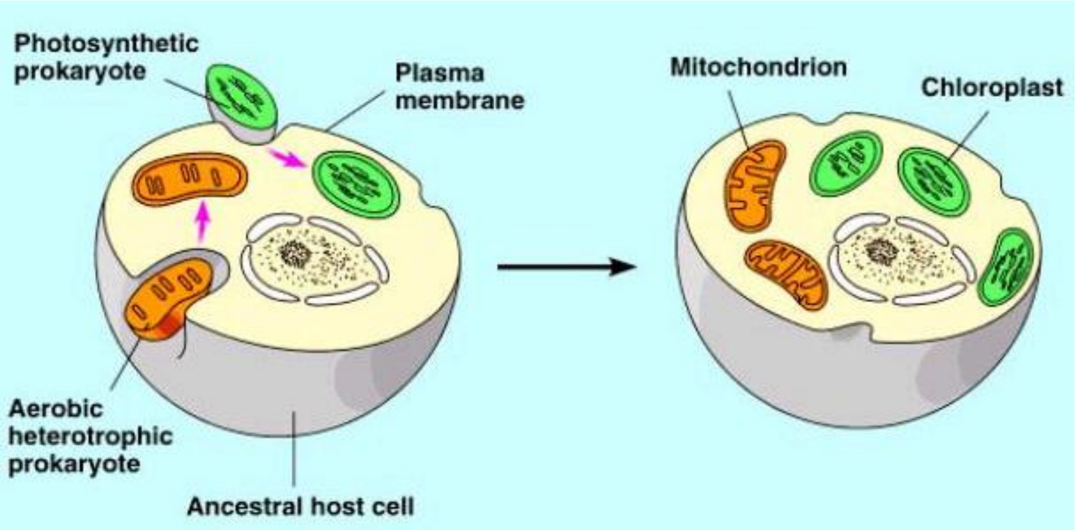
- Common Replication, transcription, translation machinery
- In many trees, archaeal homologues as outgroups
- Concept of proto-eukaryote being archael in nature

Bacterial contributor

- Bacteria and eukaryotes use Fatty ester lipids in their membranes, Archae use Fatty ether
- Some genes from eukaryotes have bacterial not archael ancestry
  - Genomic studies
  - Metabolic genes bacterial, house keeping genes archael
- Genomically, eukaryotes look like both domains
  - Histones: Histone 1-bacterial, Histones 2A, 2B, 3, 4 archaeal

Endosymbiotic origin of mitochondria

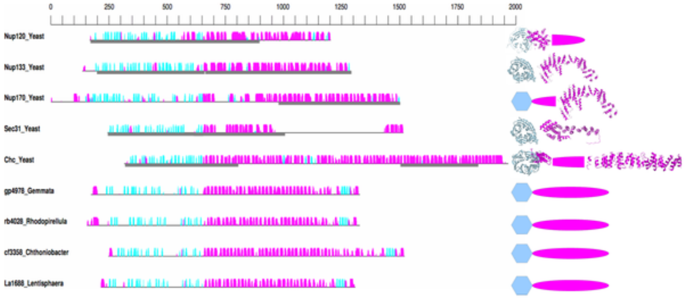
- Mitochondria and plastids have two membranes
- Mitochondria and plastids have their own DNA



- Both are morphological features that suggest these organelles are homologous to bacteria
- Mitochondria= alpha proteobacteria: Chloroplasts=cyanobacteria
- Extensive genomic and phylogenetic data confirmed this

Planctomycetes

- Bacteria with enveloped DNA
- Presence of “homologues/analogues” of protocoatomer proteins



Loki-I

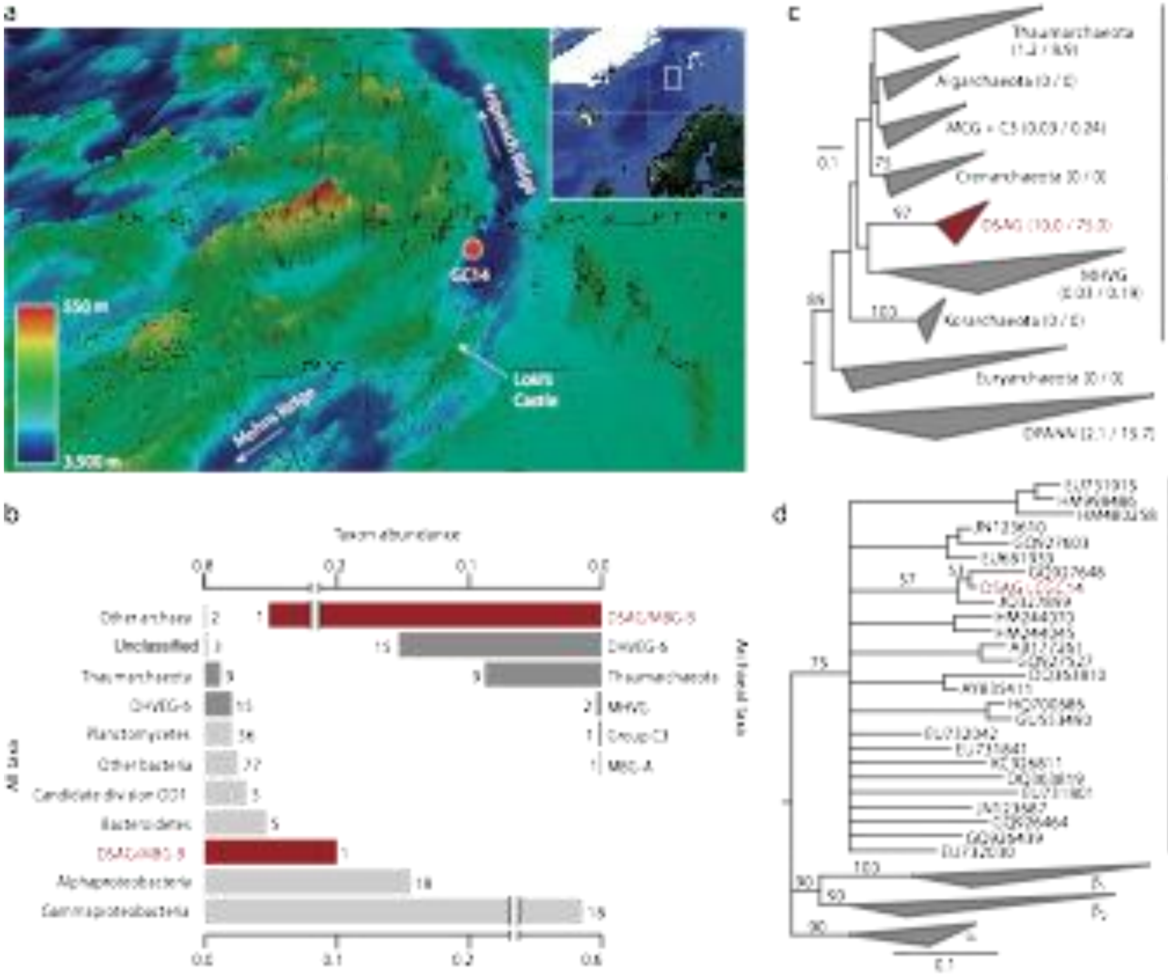


Figure 1 | Identification of a novel archaeal lineage. a, Bathymetric map of the sampling site (GC14; red circle) at the Arctic Mid-Ocean Spreading Ridge, located 15 km from Loki's Castle active vent siteb, 16S rRNA amplicon-based assessment of microbial diversity in GC14. Bars on the left represent the fraction of the respective prokaryotic taxa and bars on the right depict archaeal diversity. Numbers refer to operational taxonomic units for each group. MHVG, Marine Hydrothermal Vent Group; DHVEG-6, Deep-sea Hydrothermal Vent Euryarchaeota Group 6; MBG-A and -B, Marine Benthic Group A and B. c, Maximum likelihood phylogeny of the archaeal 16S rRNA reads (see b), revealing that DSAG sequences cluster deeply in the TACK super-phylum. Numbers between brackets indicate relative abundance (%) of each group relative to total and archaeal reads, respectively. MCG, Miscellaneous Crenarchaeota Group; MHVG, Marine Hydrothermal Vent Group. d, Maximum-likelihood phylogeny of 16S rRNA gene sequences indicating that the DSAG operational taxonomic unit (red font) belongs to the DSAGc cluster. Bootstrap support values above 50 are shown. c, d, Scale indicates the number of substitutions per site.

Lokiarchaeota are the outgroup to eukaryotes

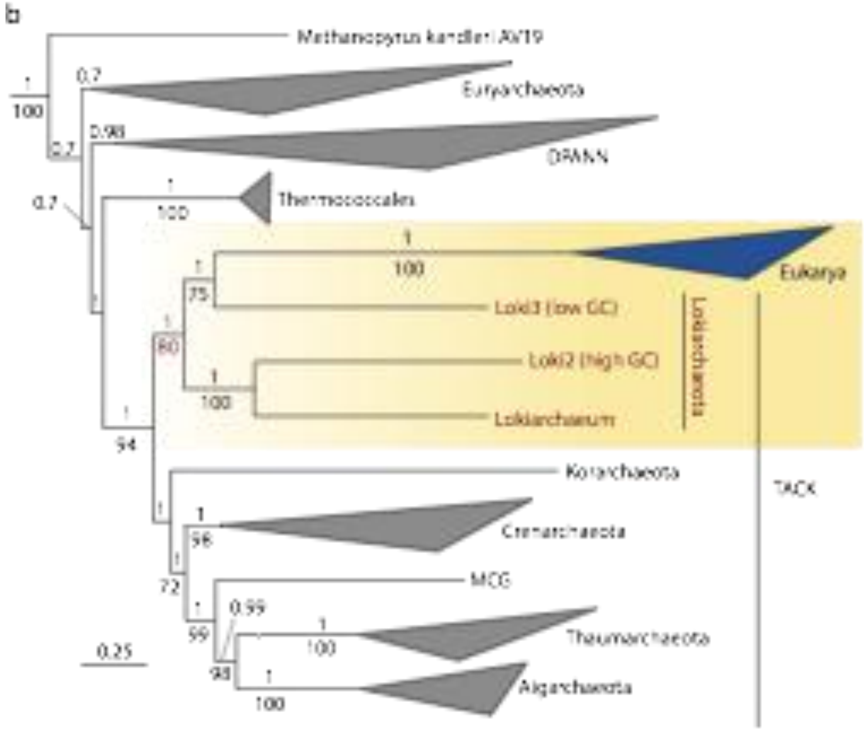


Figure 2 | …phylogenetic analysis of Lokiarchaeum

b, Bayesian phylogeny of concatenated alignments comprising 36 conserved phylogenetic marker proteins using sophisticated models of protein evolution (Methods), showing eukaryotes branching within Lokiarchaeota. Numbers above and below branches refer to Bayesian posterior probability and maximum-likelihood bootstrap support values, respectively. Posterior probability values above 0.7 and bootstrap support values above 70 are shown. Scale indicates the number of substitutions per site.

Lokiarchaeota possess features previously considered Eukaryote-specific

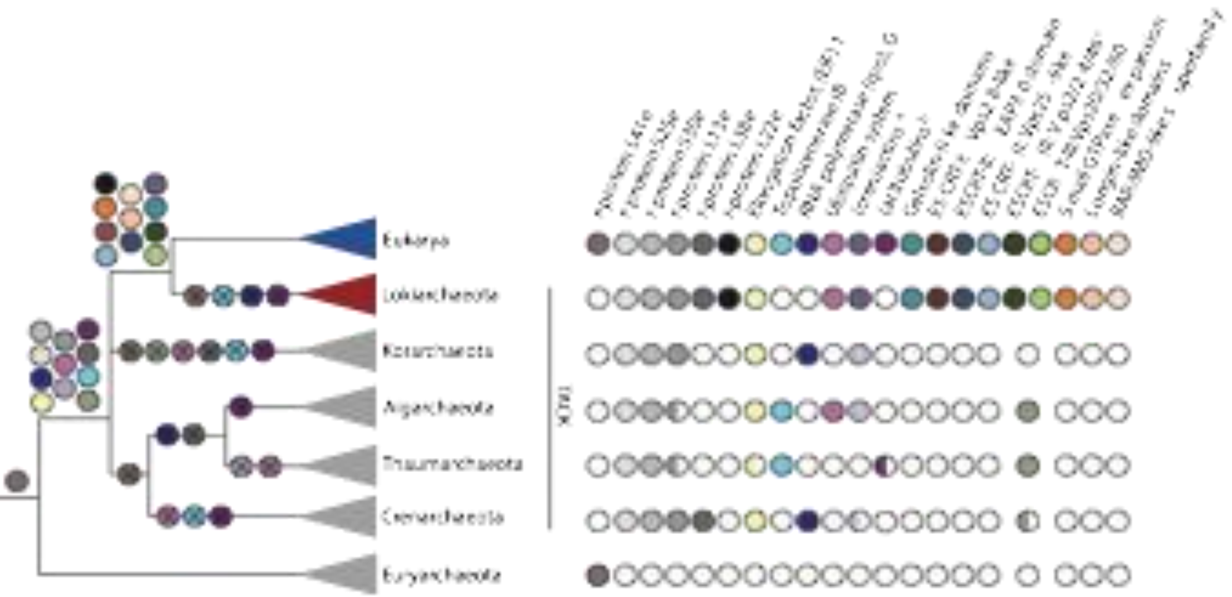


Figure 5 | The complex archaeal ancestry of eukaryotes. Schematic overview of the distribution of ESPs in major archaeal lineages across the tree oflife. Each ESP is depicted as a coloured circle and losses are indicated with a cross. Patchy distribution and absence of a particular ESP in archaeal phyla is indicated by half-shaded and white circles, respectively. While eukaryotes and Lokiarchaeota contain bona fide actins, other archaea encode the more distantly related Crenactins. [b]Only few members of the Thaumarchaeota contain distantly related homologs of tubulins (ar-tubulins). [c]Thaum-, Aig- and some Crenarchaeota contain distant homologues of ESCRT-III (SNF7 domain proteins).

Explanation of origin of eukaryotes

- Must account for…
  - Cell Biology
    - Organelles;genome architecture;
  - Physiological capacities of proto-eukaryote
  - Genome signatures
  - Environment
  - Inertia
    - Selective pressure or catastrophy

How might population genetics add to this debate?

- Add a dimension of rigour to an area with a great deal of hand-waving
  - E.g. Energetics of genome complexity and cost of a gene
- A priori questions
- Provide constraints or ranges on existing theories? Can anything be ruled out?

How might population Genetics add to this debate?

- Add a dimension of rigour to an area with a great deal of hand-waving
  - E.g. Energetics of genome complexity and cost of a gene
- A priori questions
- Provide constraints or ranges on existing theories? Can anything be rules out?

A priori questions

- Gene flow?
  - When did sex evolve?
  - What type of genome architecture (X-somes)?
- In what order did the ESPs evolve?
  - Order of endomembrane compartments?
  - Genome size leads to mitosis leads to meiosis?
- What population growth rate, population size say about the types of ecological scenarios?
  - Implications of a small Ne initially, then large
  - What types of environments?
  - Virgin niche and major mutational jump?

How might population Genetics add to this debate?

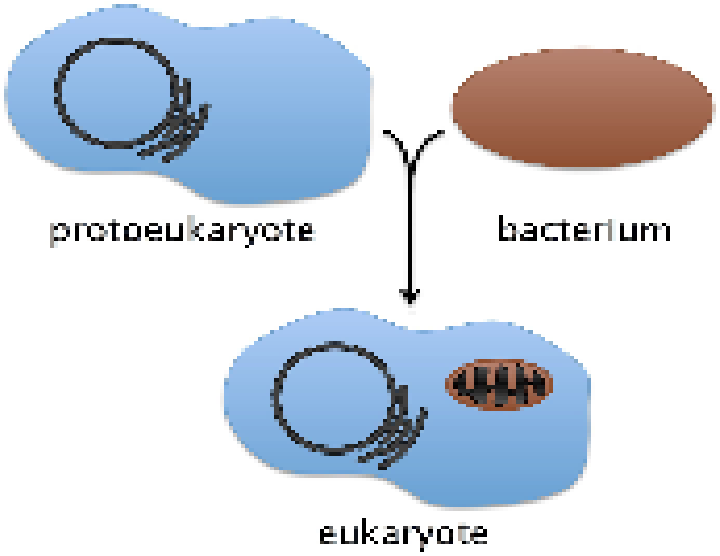
- Add a dimension of rigour to an area with a great deal of hand-waving
  - E.g. Energetics of genome complexity and cost of a gene
- A priori questions
- Provide constraints or ranges on existing theories? Can anything be ruled out?

2 major classes of existing theories

- Phagotrophic and syntrophic
- What type of organism took on the mitochondrial endosymbiont: Phagotrophic (eukaryote-like), Syntrophic (2 prokaryotes interact)
- Also about origin was initially a fusion (endosymbiosis) or autogenous event

Phagotrophic origins

- The cell that engulfed the mitochondria already possessed a nucleus.
- The nucleus (and other endomembrane compartments) evolved autogenously and gradually

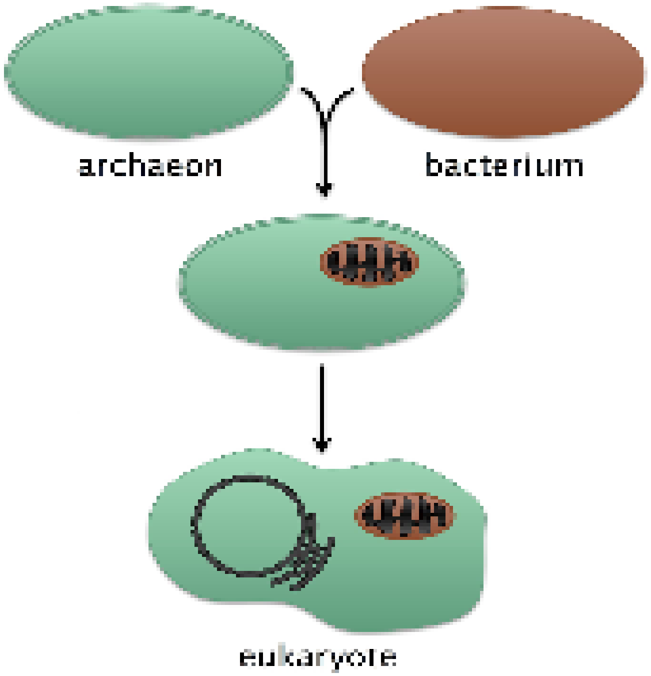


Phagotrophic origins-Cavalier-Smith circa 2002

- Euks and Archae common (neomuran) ancestor
  - 850 million years ago (discounts *Bangiomorpha*)
- Loses bacterial cell wall
  - Archaea gain S-layer
  - Eukaryotes gain cytoskeleton
- Enables phagocytosis
- This allows eventual mitochondrial endosymbiosis, event of secondary importance

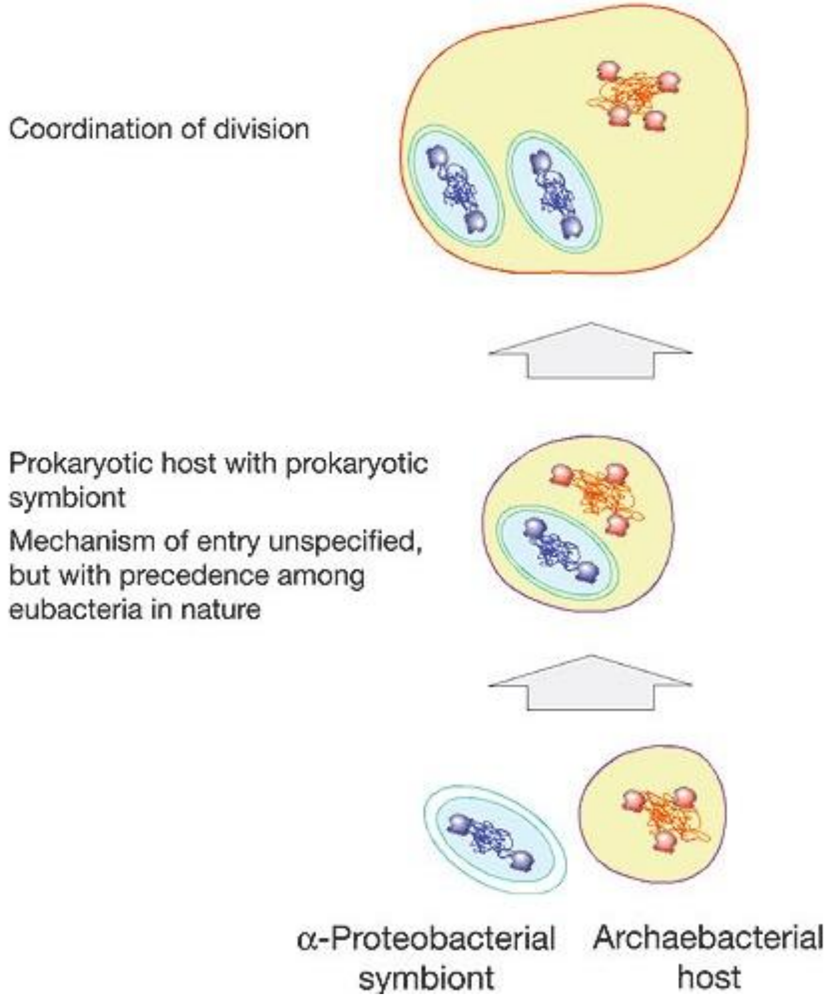
Syntrophic

- Metabolic interdependence between two prokaryotes
- Eventual fusion into a eukaryote



Hydrogen hypothesis V.2.0

- Eukaryogenesis by fusion of archaebacterial host and 〈-proteobacterial endosymbiont
- Integration of symbiosis but with two different gene expression systems



A syntrophic theory to explain a phagotrophic host

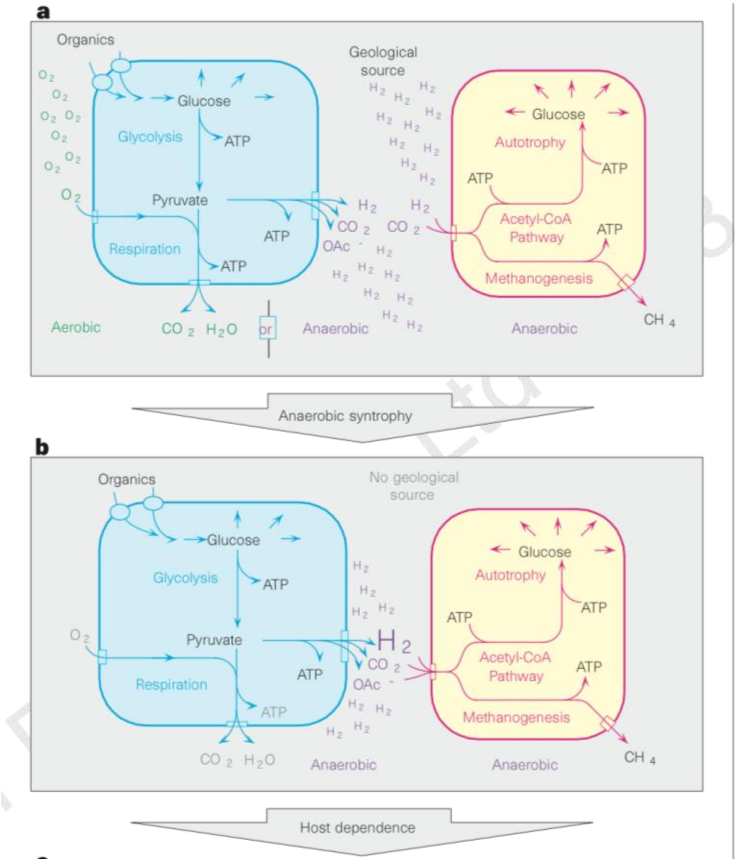
- Forterre 2010
- “A new fusion hypothesis for the origin of Eukarya: better than previous ones, but probably also wrong”
- Planctomycete plus Thaumarchaea plus viral contribution
- Not an explanation about mitochondria, but before
- *Might be onto something…*

Putting bounds on existing theories

- Can population genetics provide ranges on growth rate, population size, mutation rate, gene flow?
- Can these be used to temper existing theories?
- Can these be used to rule any out?

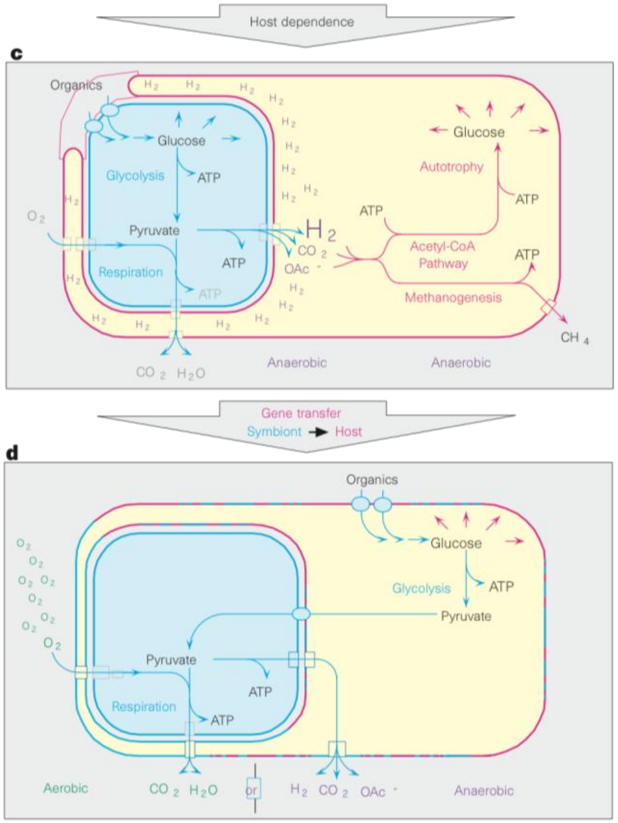
Hydrogen hypothesis

- Bill Martin and Miklos Muller, 1998
- Emphasis on mitochondria having a hydrogen producing capacity
- Hydrogen producing alpha proteo bacteria plus anaerobic autotrophic archael methanogen
- Syntrophic hydrogen replaces geological source



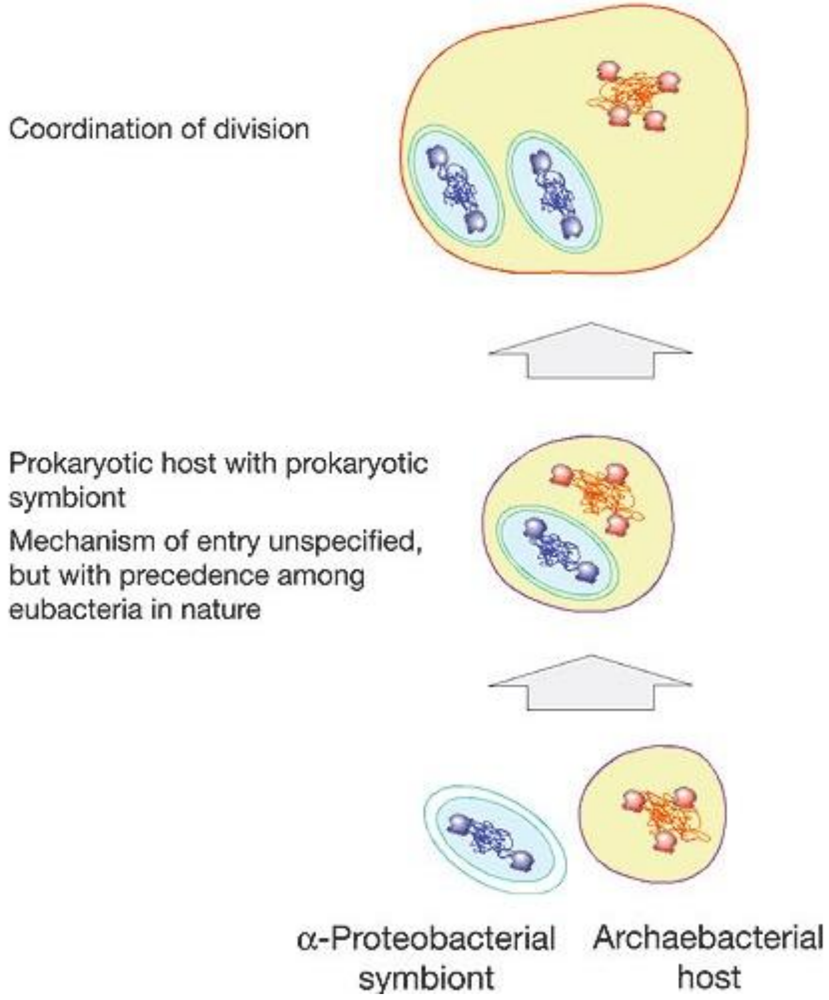
Hydrogen hypothesis

- Dependence is selective for engulfment
- Eventually gene transfer turns into organelle
- Integration of two genetic systems



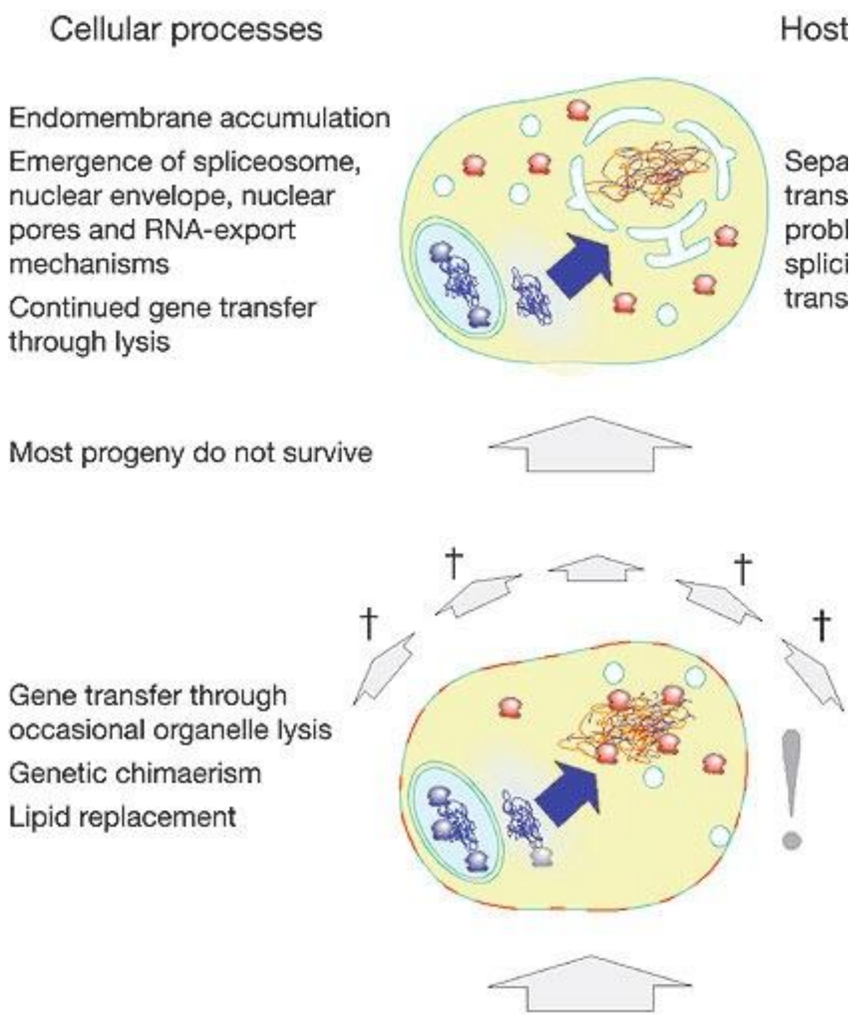
Hydrogen hypothesis V.2.0

- Eukaryogenesis by fusion of archaebacterial host and 〈-proteobacterial endosymbiont
- Integration of symbiosis but with two different gene expression systems



Origin of the nuclear envelope-Martin and Koonin 2006

- EGT sends mitochondrial genes to nuclear genome
- These include lipid synthesis genes
- Two types of lipids made, so vesicles in cytoplasm
- Fuse around nucleus
- Selection for NE to decouple transcription and translation

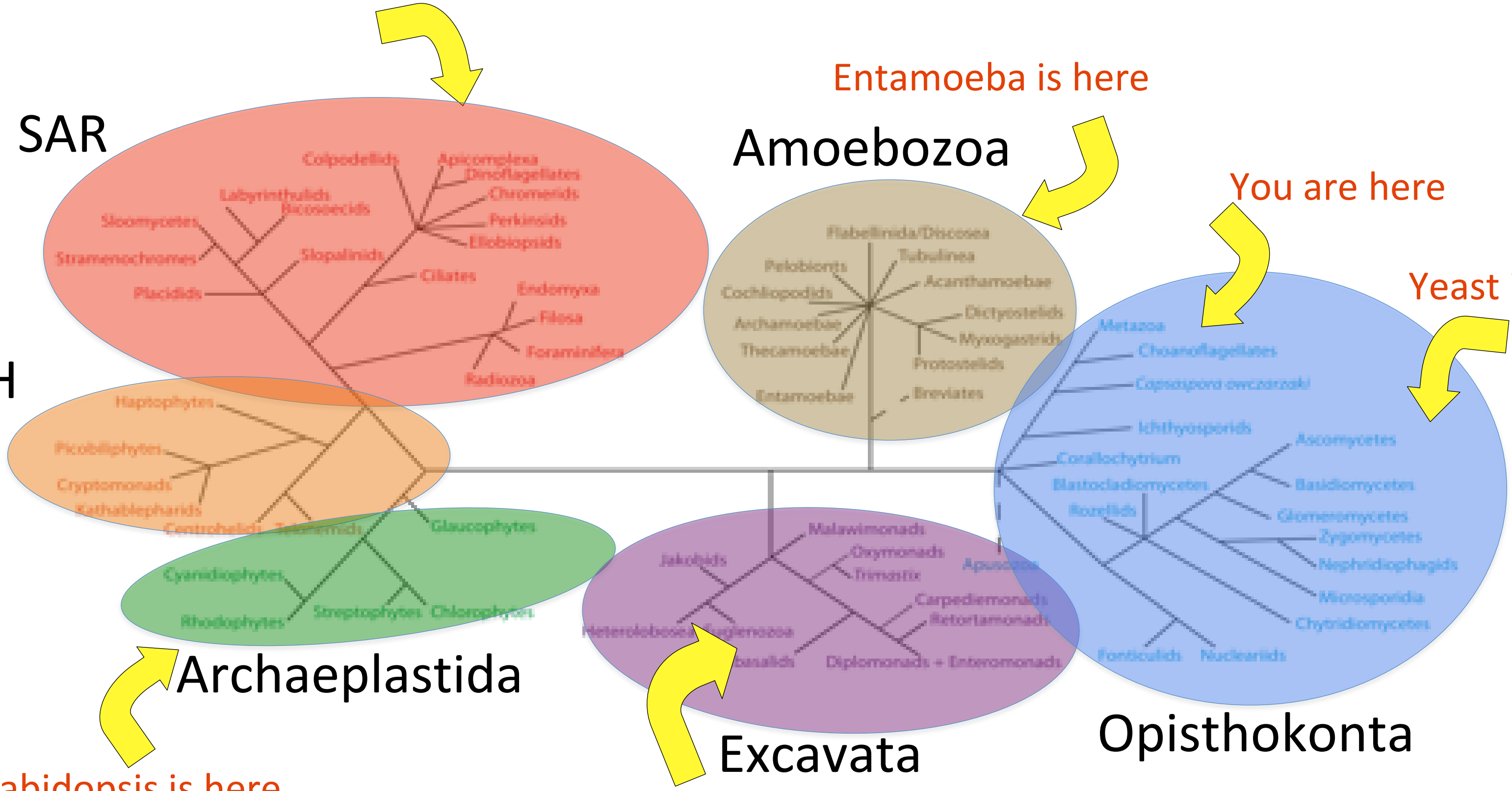


Aspects of the theories

- Phagotrophic
  - Prioritize cell biology
  - Prioritize internal factors
  - Autogenous
  - Prioritize plausibility, probability
  - Gradual
  - Cumulative
- Syntrophic
  - Prioritize genomic/biochemical data
  - Prioritize external factors
  - Exogenous
  - Prioritize the unique event
  - Stochastic
  - Cataclysmic

Both aim to maximize coherence with evidence, logical consistency

The diversity of Eukaryotes



Plasmodium is here

SAR

Entamoeba is here

Amoebozoa

You are here

Yeast is here

CCTH

Emiliania
is here

Archaeplastida

Arabidopsis is here

Excavata

Trypanosoma is here

Opisthokonta