The Viral Eukaryogenesis Hypothesis

A role for viruses in the emergence of eukaryotes from a prokaryotic world environment

Dr Philip Bell

Salzburg

The VE hypothesis is based on modern observable phenomena

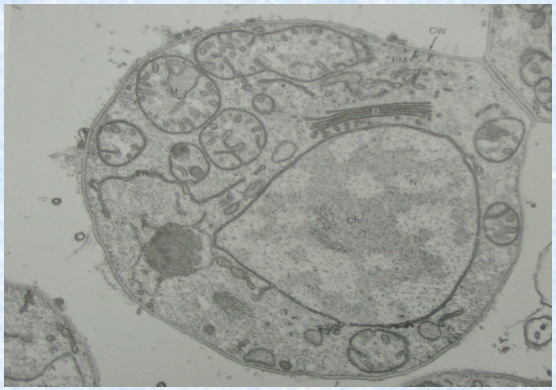
- Eukaryotes arose out of a ‘prokaryotic world’ consisting of bacteria, archaea and viruses
- Archaea, bacteria and viruses were modern in design before eukaryogenesis
- Replication strategies were similar to today
- Evolutionary pressures were similar to today
- No organisms are invoked of which we have no evidence (RNA organisms, progenotes)

Within these boundaries the VE hypothesis can explain the massive gulf in genetic design observed between the prokaryotic and eukaryotic cells

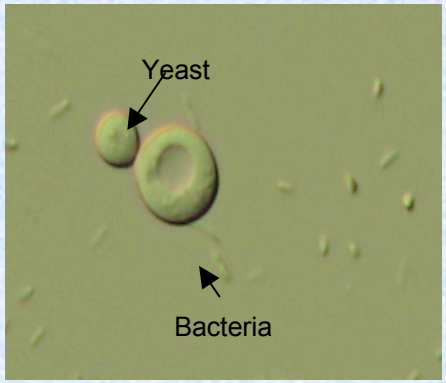
It can also explain the origin of sex and meiosis which were crucial in changing the mode of cellular evolution

Differences between prokaryotic and eukaryotic design

Eukaryotic microbe

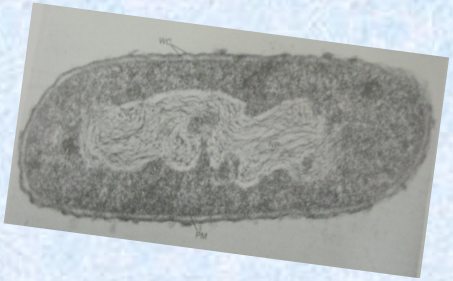


Yeast 6000 genes



Very similar niche

Prokaryotic microbe

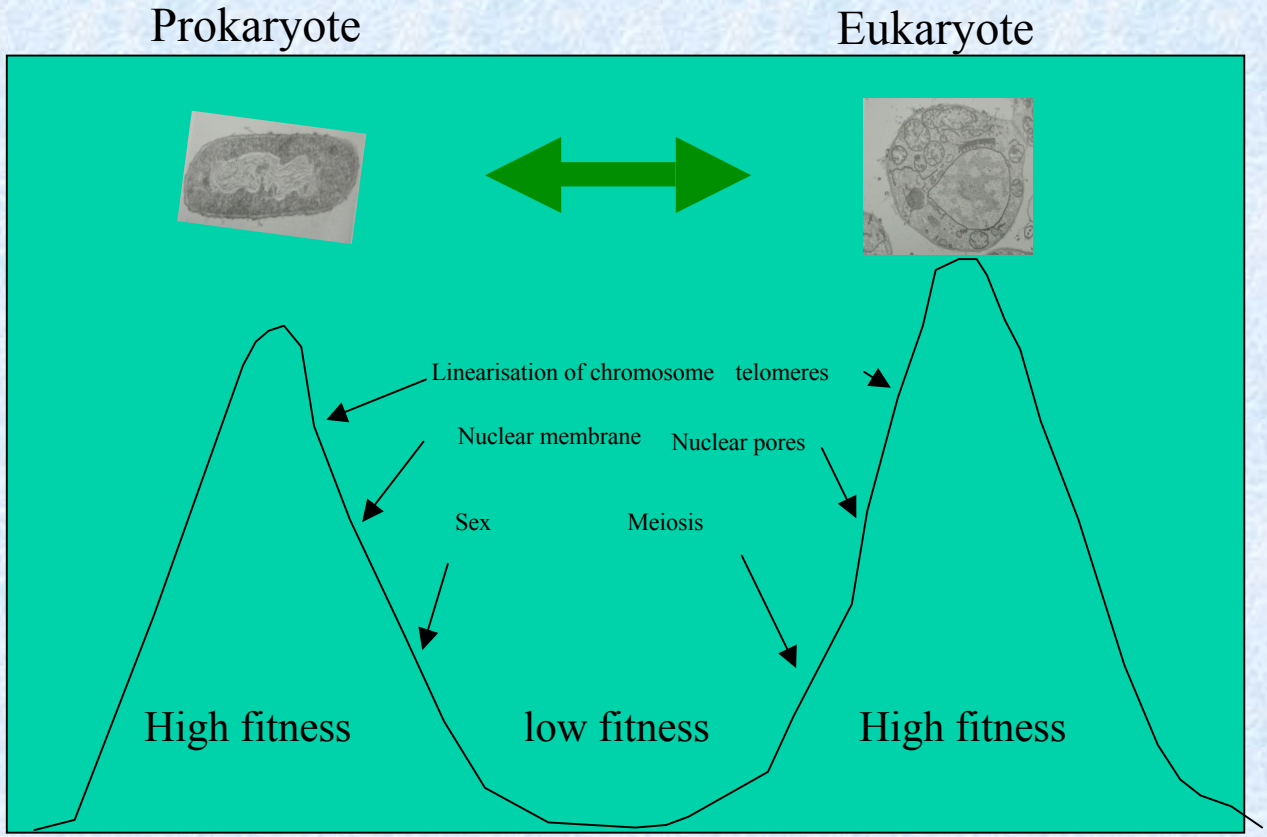


E. coli~4000 genes

Presence of membrane bound nucleus	No nucleus or nuclear membrane
Inside nucleus genome is made of linear chromosomes	Circular chromosomes immersed in cytoplasm
mRNA is processed, capped and transported outside nucleus	Transcription and translation in same compartment
mRNA is transported via the nuclear pore complex	No nuclear pores
Endo-membrane consists of golgi, ER and nuclear membrane	No Endo-membrane system
Specialized organelles with own genetic systems	No organelles
Sex, meiosis and mitosis	Binary cell division

Ribosomes show that all life shares a common ancestory, but the gulf between eukaryotic and prokaryotic design is difficult to explain

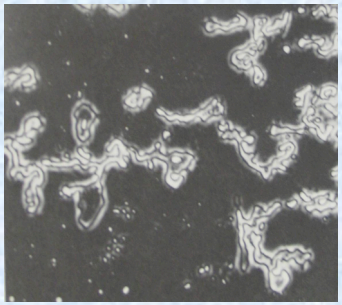
Transitional organisms have not been observed and would have low fitness



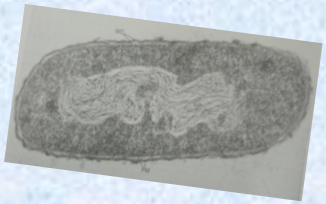
The VE hypothesis solves this problem by using a novel approach to understanding the evolution of the eukaryotes

Viral Eukaryogenesis hypothesis proposes the eukaryotes are not a direct descendent of any individual prokaryote lineage

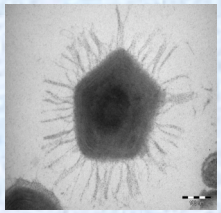
Archaea
(Cytoplasmic ancestor)

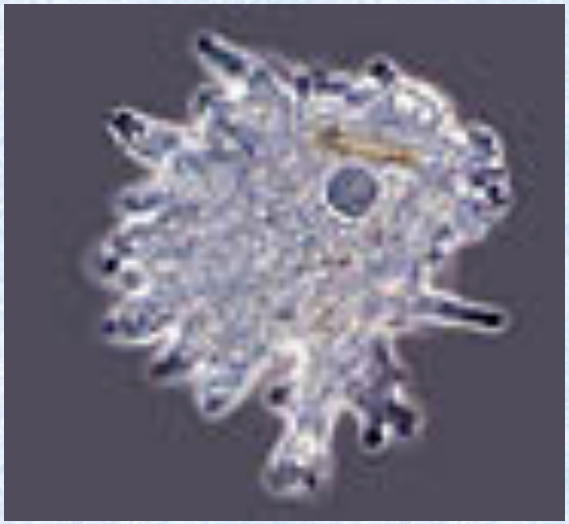


Bacteria
(Mitochondria ancestor)



Virus
(Nuclear ancestor)

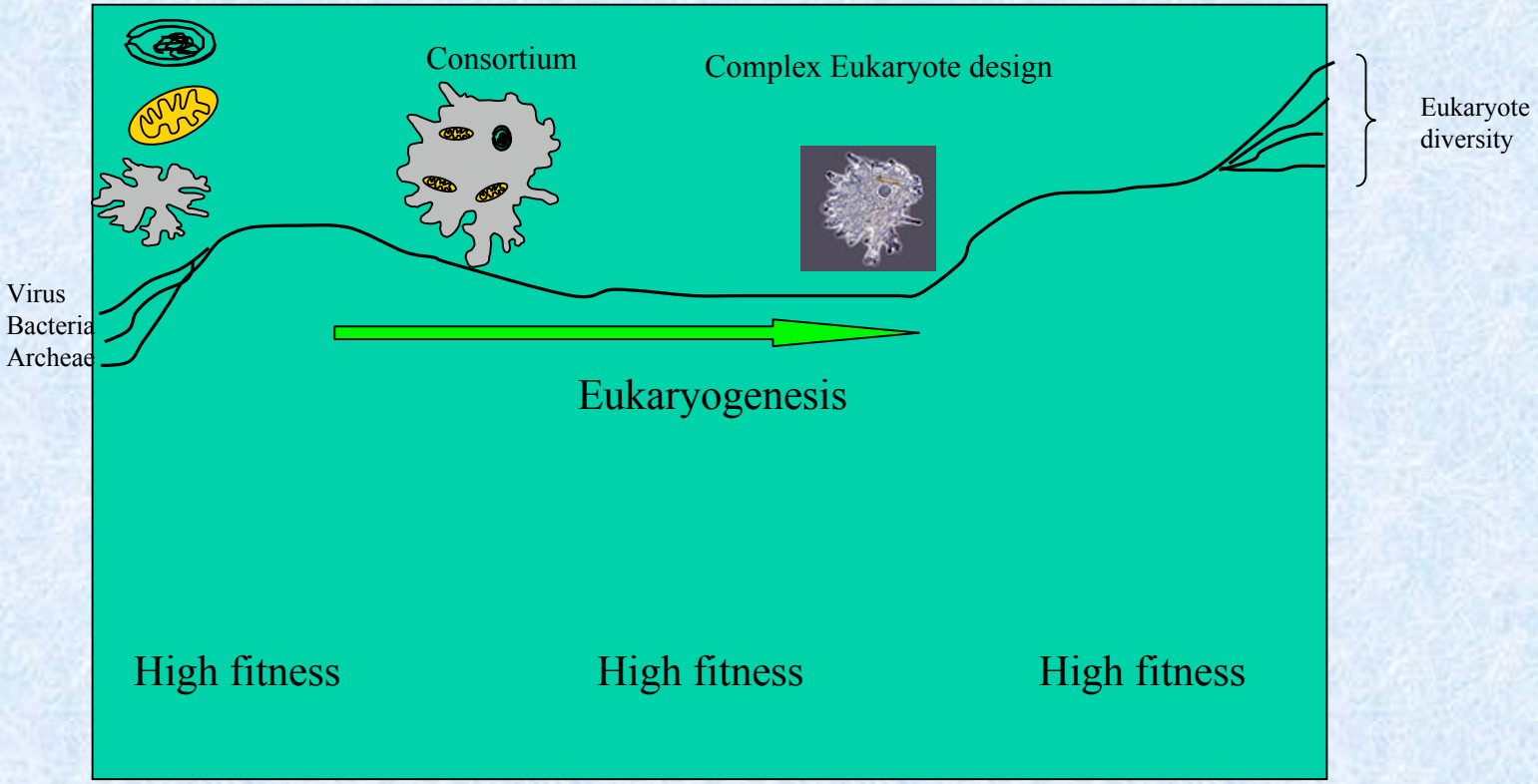




Eukaryote
(consortium)

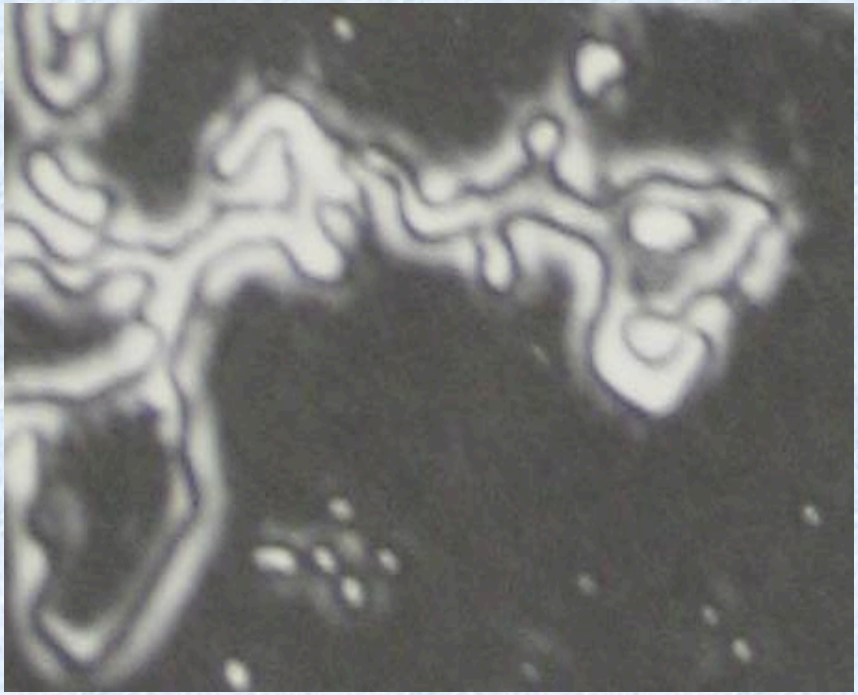
Eukaryotes are an integrated consortium of three lineages from a 'prokaryotic world'

Consortium allows eukaryotes to arise from a prokaryotic world without passing through a low fitness intermediate



Eukaryotic replication = Co-ordinated replication of viral, archaeal and bacterial lineages

Lineage 1: an archaeal ancestor of the eukaryotic cytoplasm



'Methanoplasma elizabethii.'

a methanogenic 'mycoplasma' (Rose and Pirt 1981)
produces methane from carbon dioxide and hydrogen produced by bacterial syntrophs

Archaea chosen because

•Strong phylogenetic support for a common ancestor of the archaeal and eukaryotic translation apparatus

'Mycoplasma' chosen because

•Ancestor eukaryote had no cell wall

•Absence of cell wall allows infection by viruses using membrane fusion

Lineage 1, like *M.elizabethii,* was in a syntrophic relationship with a bacterium which became lineage 2

Lineage 2: an alpha proteo-bacterial ancestor of the mitochondrion

Emelyanov VV. Common evolutionary origin of mitochondrial and rickettsial respiratory chains.Arch Biochem Biophys. 2003 Dec 1;420(1):130-41.

Alpha-proteobacteria chosen because

- Strong phylogenetic support
- Similarity in structure and function
- Electron transport chain homology

← Mitochondria

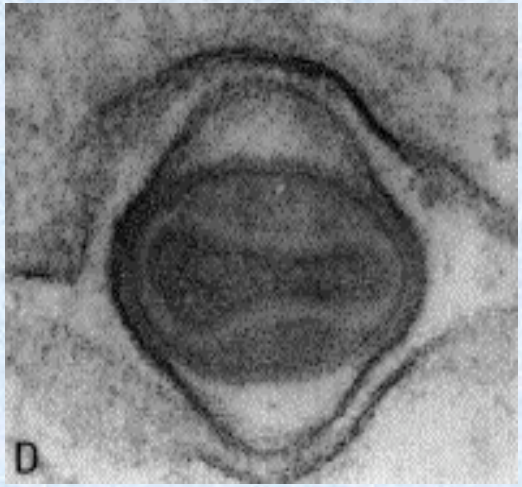
← Endoparasite (eg Rickettsia)

← Free living /syntrophic bacteria

Syntrophy links evolution of the archaeal and bacterial lineages together

Lineage 3: a viral ancestor of the nucleus

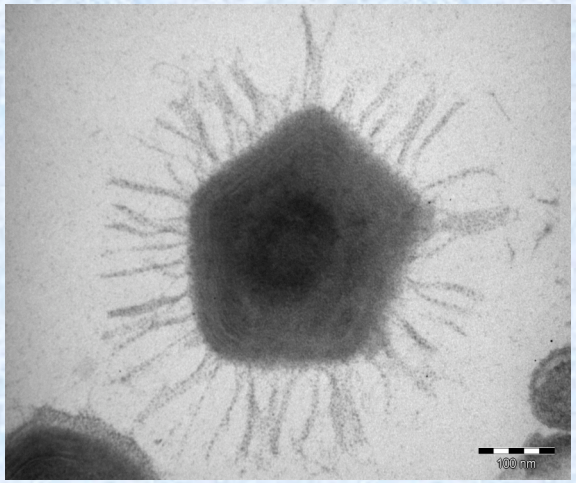
Pox virus



Animal host

200 kb linear

Mimivirus



Amoeba host

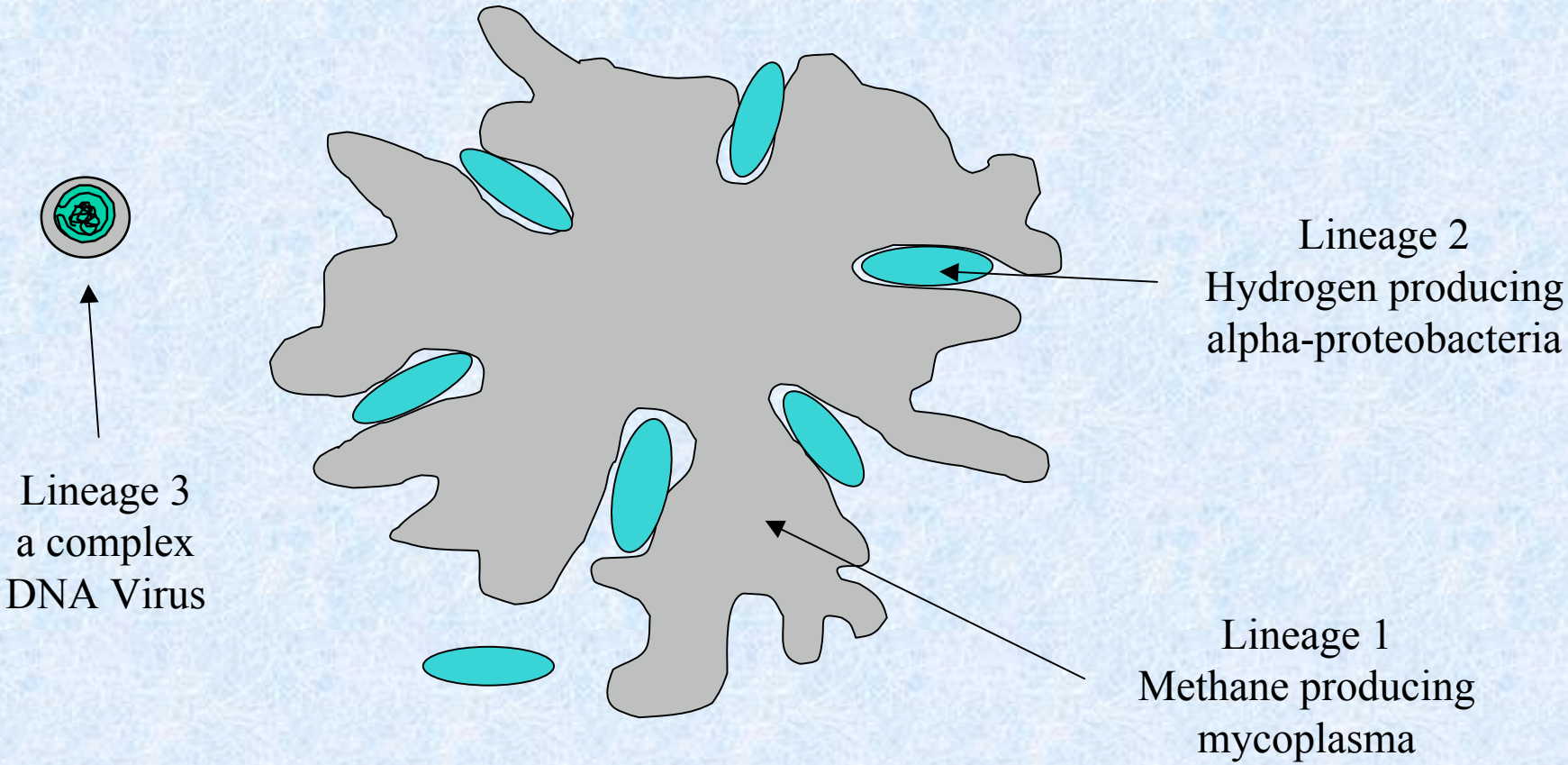
1200 kb linear

Pox-like member of NCLDV viruses (Iyer et al 2001) **chosen because of similarities to the eukaryotic nucleus**

- DNA based genome
- Large linear chromosome
- Tandem repeats at telomere
- transcription apparatus
- cytoplasmic replication
- membrane bound
- capped mRNA

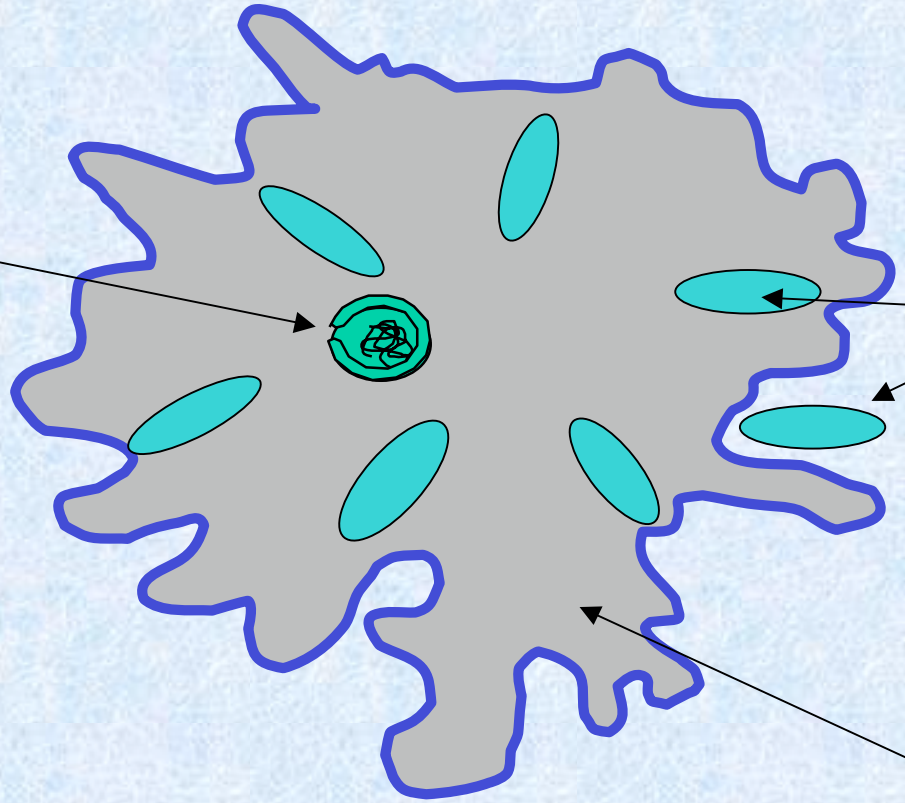
Host/parasite relationship links evolution of the archaeal and viral lineages together

Syntrophic ecology consists of archaeal, bacterial and viral lineages



Virus and archaea co-evolve (host/parasite relationship)
Bacteria and archaea co-evolve (syntrophic relationship)
All three elements are evolutionarily linked via the archaeon

Eukaryogenesis begins when virus lysogenises the host methanogen and a bacterium is internalised

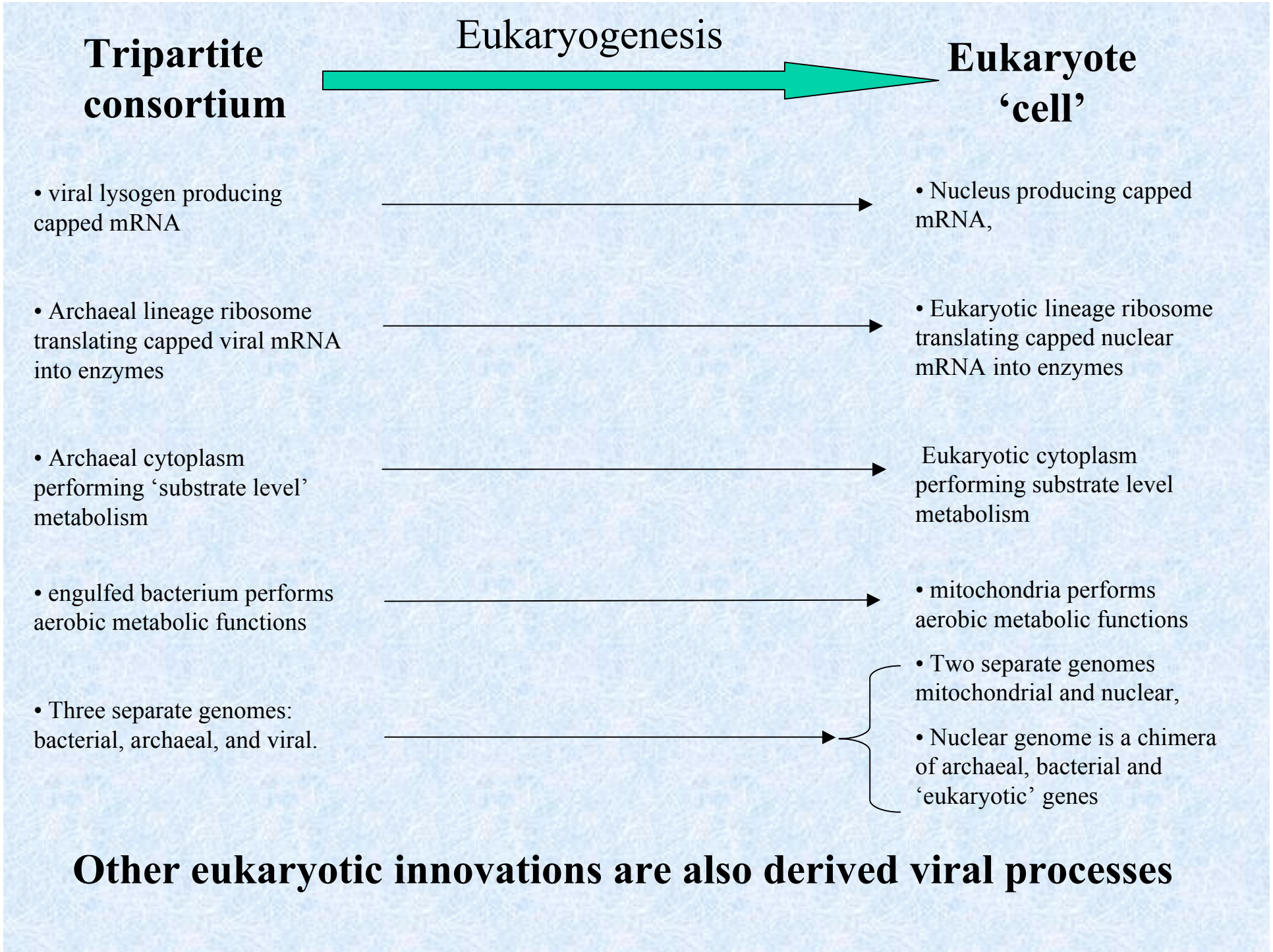


Virus stably replicates in host cytoplasm

Hydrogen producing bacteria internalised either via membrane fusion process or bacterial invasion

Methane producing mycoplasmal host supports replication of the virus and syntroph

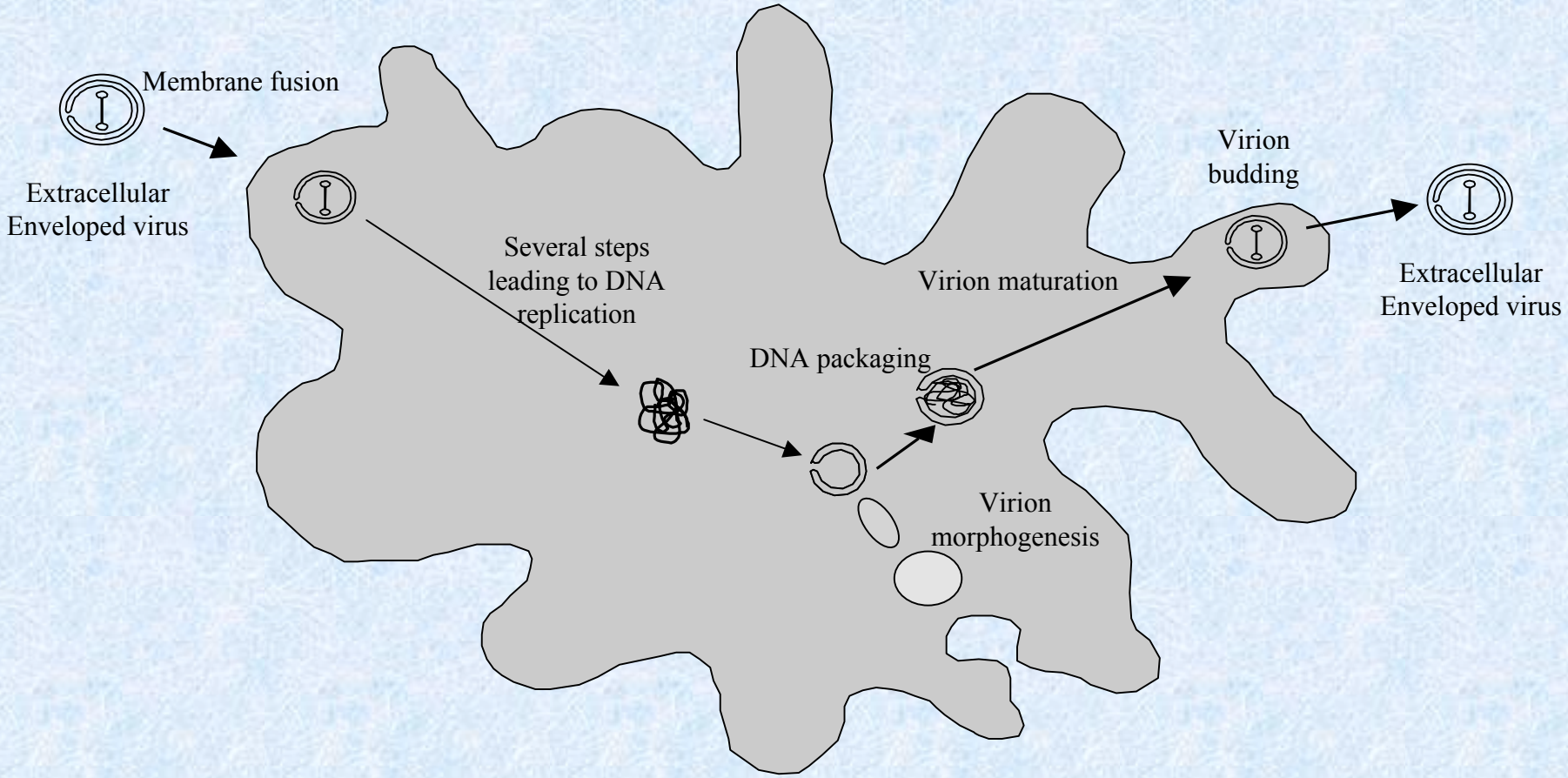
Tripartite consortium is ancestor of ‘eukaryotic cell’


Tripartite consortium
Eukaryogenesis
Eukaryote 'cell'
• viral lysogen producing capped mRNA
• Nucleus producing capped mRNA,
• Archaeal lineage ribosome translating capped viral mRNA into enzymes
• Eukaryotic lineage ribosome translating capped nuclear mRNA into enzymes
• Archaeal cytoplasm performing 'substrate level' metabolism
Eukaryotic cytoplasm performing substrate level metabolism
• engulfed bacterium performs aerobic metabolic functions
• mitochondria performs aerobic metabolic functions
• Three separate genomes: bacterial, archaeal, and viral.
• Two separate genomes mitochondrial and nuclear,
• Nuclear genome is a chimera of archaeal, bacterial and 'eukaryotic' genes
Other eukaryotic innovations are also derived viral processes

Viral ability to bend host plasma membrane to produce internal vesicles leads to the evolution of endomembrane system

- Nuclear membrane
- Endoplasmic reticulum
- Pinocytosis and endocytosis

} Viral morphogenesis systems



The PRD1 phage is a prokaryotic model for this process

Prd1 virus causes formation of spherical lipid vesicles in the bacterial host cytoplasm

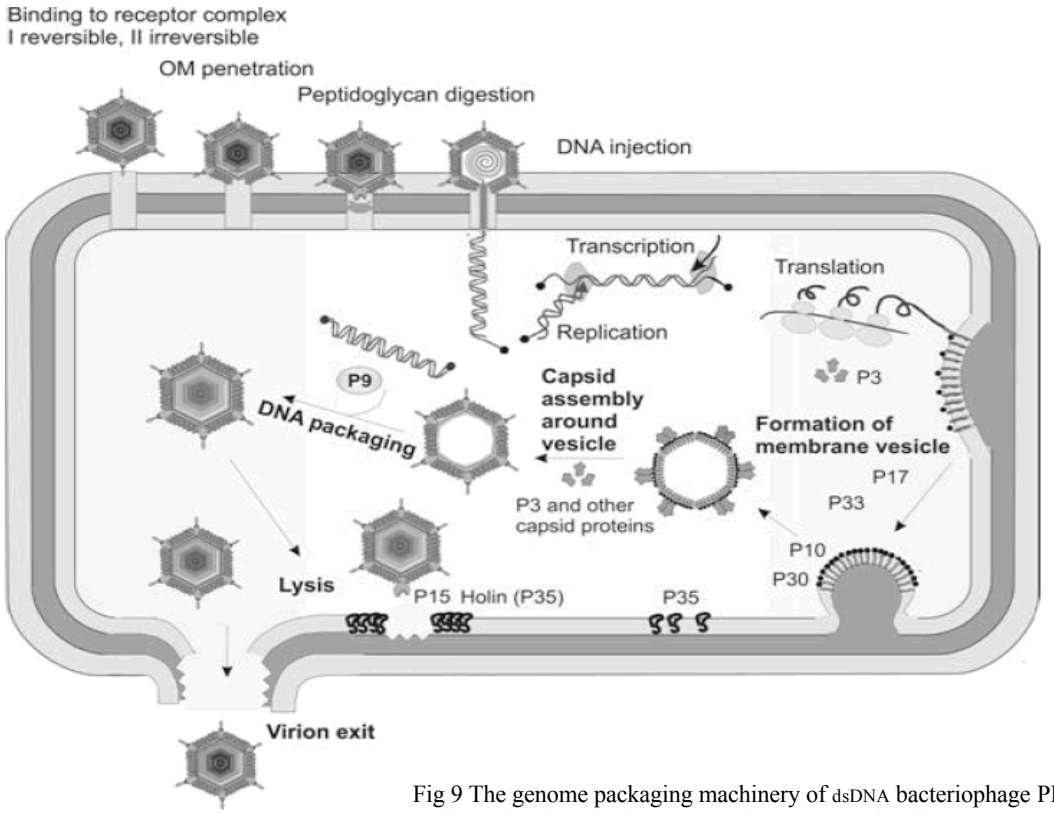


Fig 9 The genome packaging machinery of dsDNA bacteriophage PRD1
Nelli Karhu Ph.D thesis, 2006, University of Helsinki

PRD1 shows homology to NCLDV viruses

- capsid protein folds
- ATPase for capsid filling

A three lineage consortium explains many aspects of eukaryotic design, including the nucleus, mRNA capping, the endomembrane system and mitochondria

If the VE hypothesis is valid the eukaryotic cell cycle derived from three different sources

1) Nuclear replication cycle was originally a viral replication cycle

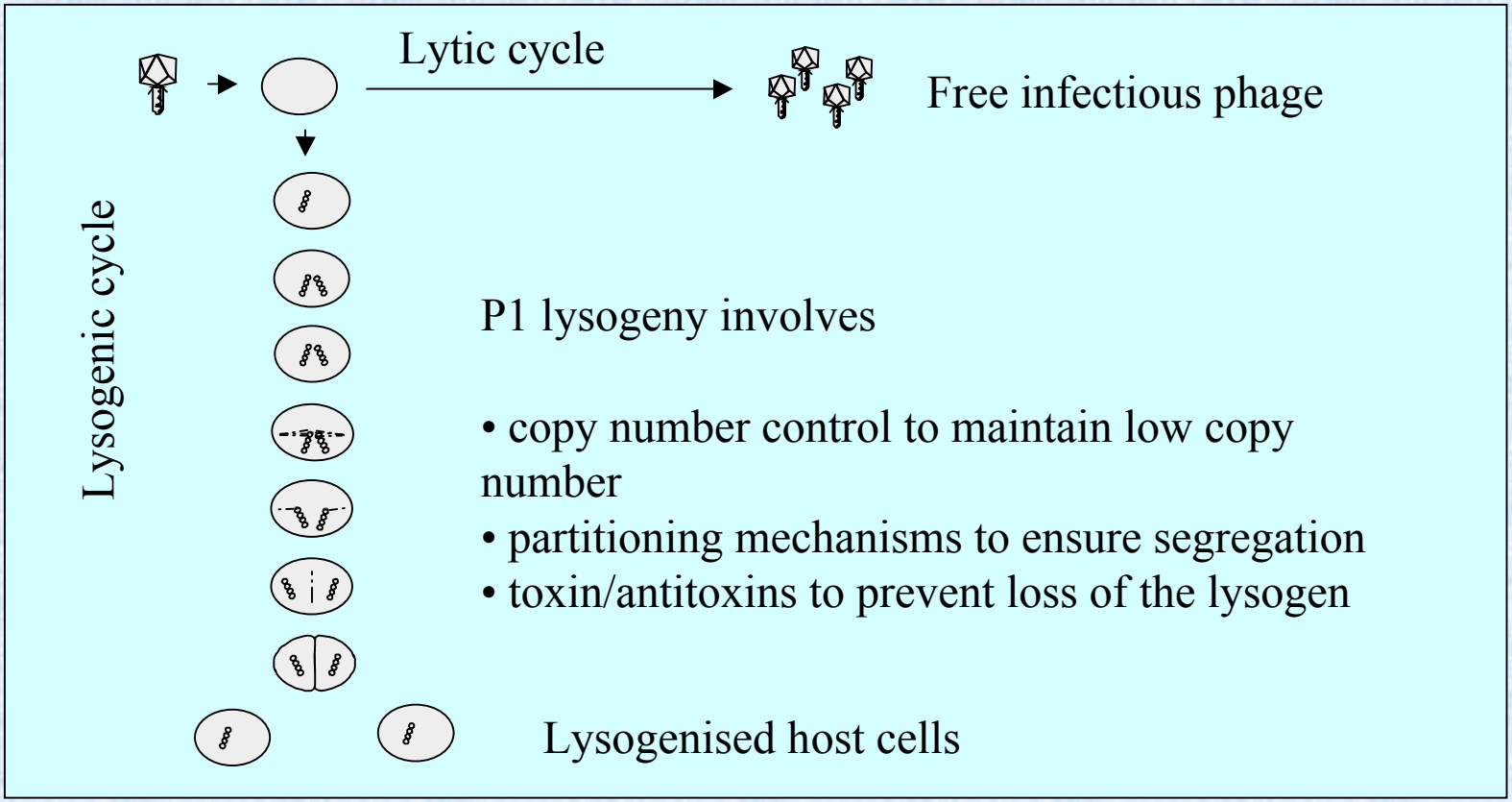
2) Cytoplasmic replication was originally an archaeal replication cycle

3) Mitochondrial replication was originally a bacterial replication cycle

How did the consortium members evolve a co-ordinated replication cycle?

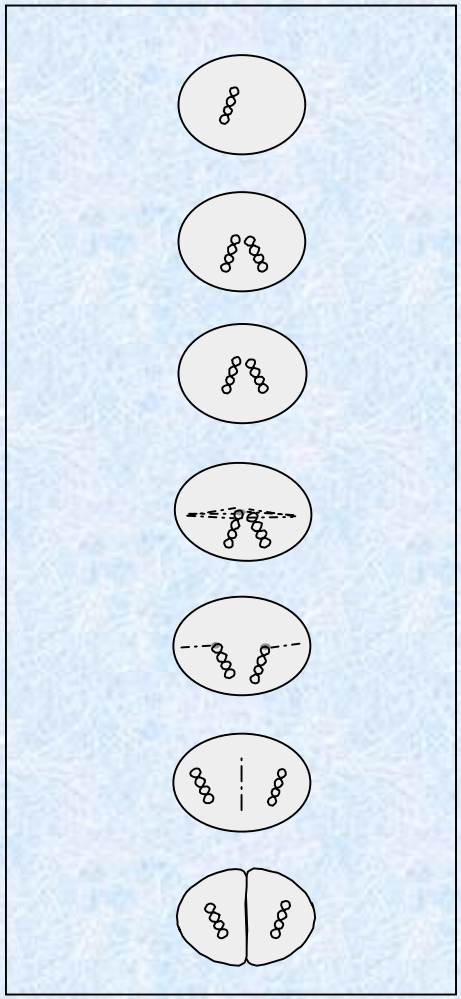
Modern viruses co-ordinate stable replication with their host using two strategies

- Lysogenise host by integrating into host genome (eg lambda, retroviruses)
- Lysogenise host as plasmid like element (eg P1 phage, Bam35, N15, Herpes)



Low copy number prokaryotic lysogens (plasmids and viruses) have convergently evolved mechanisms with common themes

R1 or F plasmid, P1, N15 phage etc



Common themes

After replication, daughter copies handcuffed at centromere

Centromere prevents further replication of plasmid

Segregation proteins bind to centromeric region

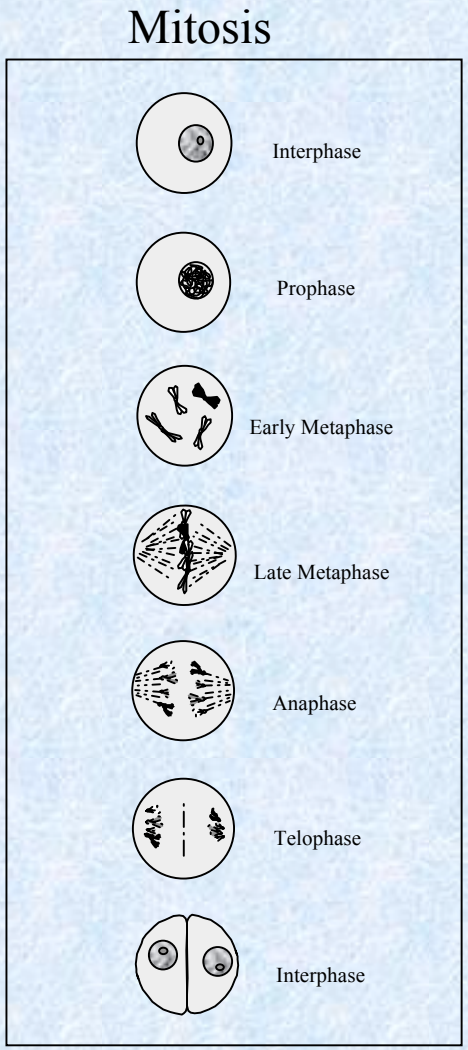
Chromosomes segregated using filament polymerisation (tubulin, actin and other protein!)

Plasmid segregation is completed before cell division

Upon completion, two daughter cells are produced both containing a copy of the plasmid

Several of these ‘themes’ are also found in the mitotic cycle

Mitosis is consistent with a viral origin of the eukaryotic nucleus



After replication, daughter copies joined at centromere

Nuclear membrane breaks down and chromosomes condense

Segregation proteins bind to centromeric region (kinetochore)

Chromosomes segregated using filament polymerisation (tubulin->microtubules)

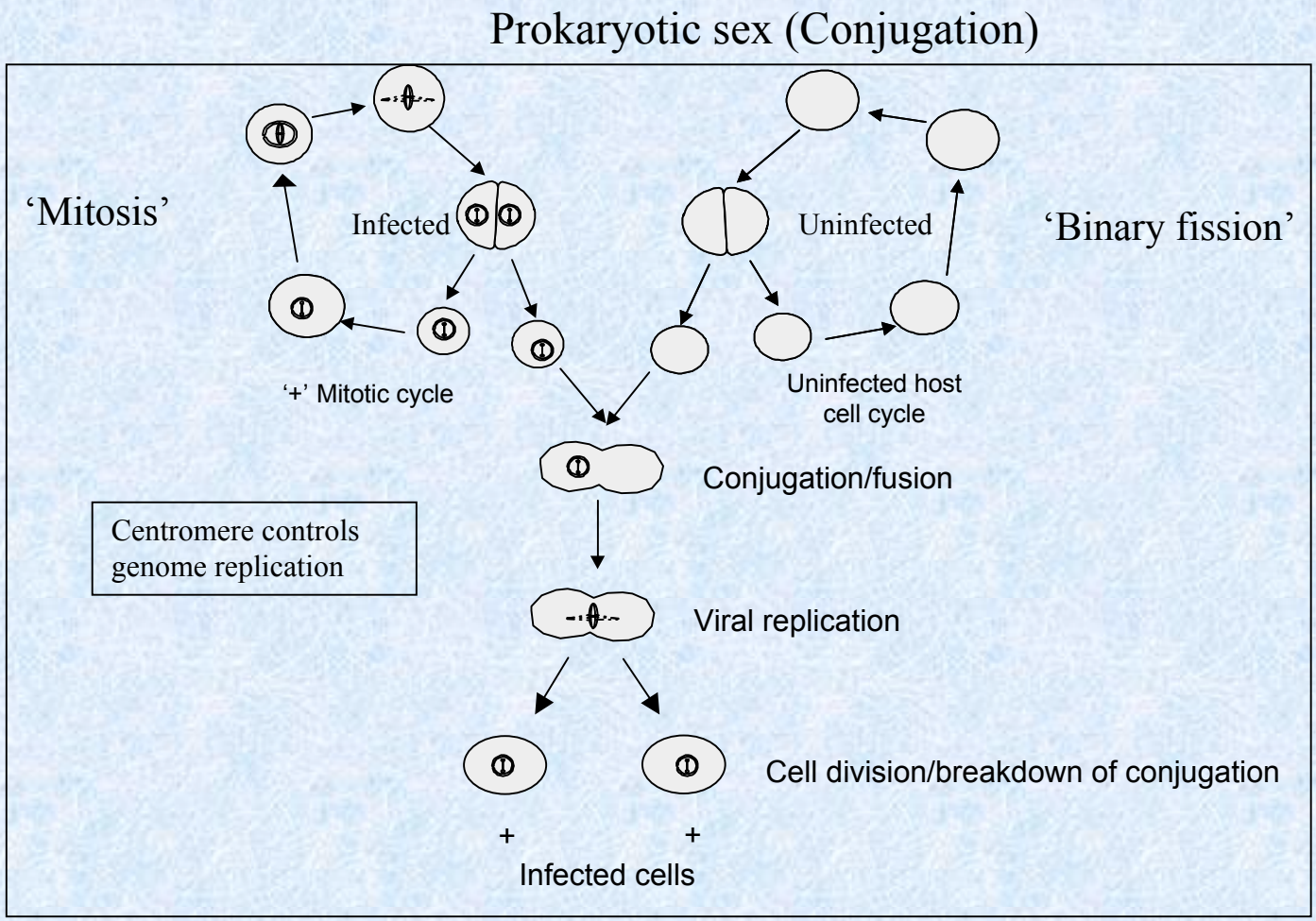
Chromosome segregation is completed before cell division

Upon completion, two daughter cells are produced both containing a copy of each linear chromosome

VE hypothesis also provides a simple model for the origin of sex and meiosis

In the VE hypothesis, the lysogenic virus evolved conjugation to enable horizontal transfer to new hosts

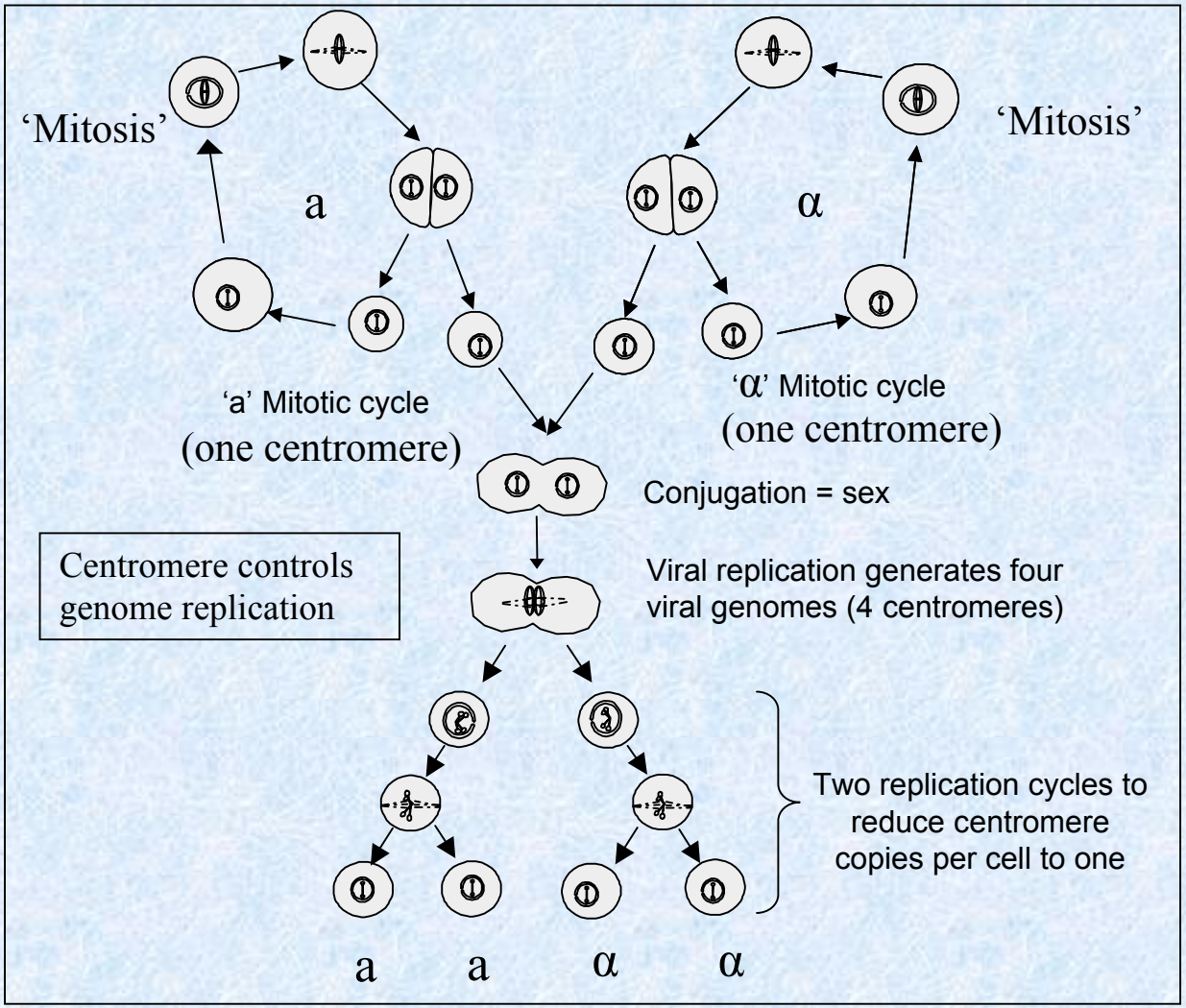
•Conjugation has evolved several times to spread low copy number lysogens such as R1, F to uninfected new lineages



Has some elements of sexual cycle

- Transfer of genetic information (plasmid and lateral gene transfer)
- Cell to cell conjugation

Eukaryotic sex emerges when the viral ancestor diverged into two closely related forms

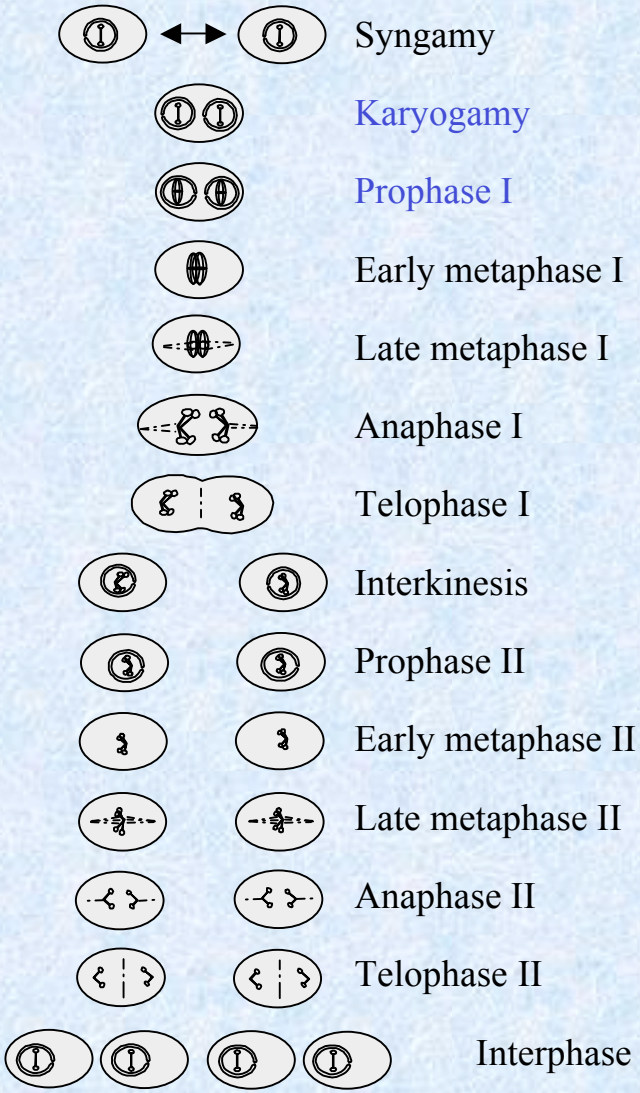


Closely related viruses are

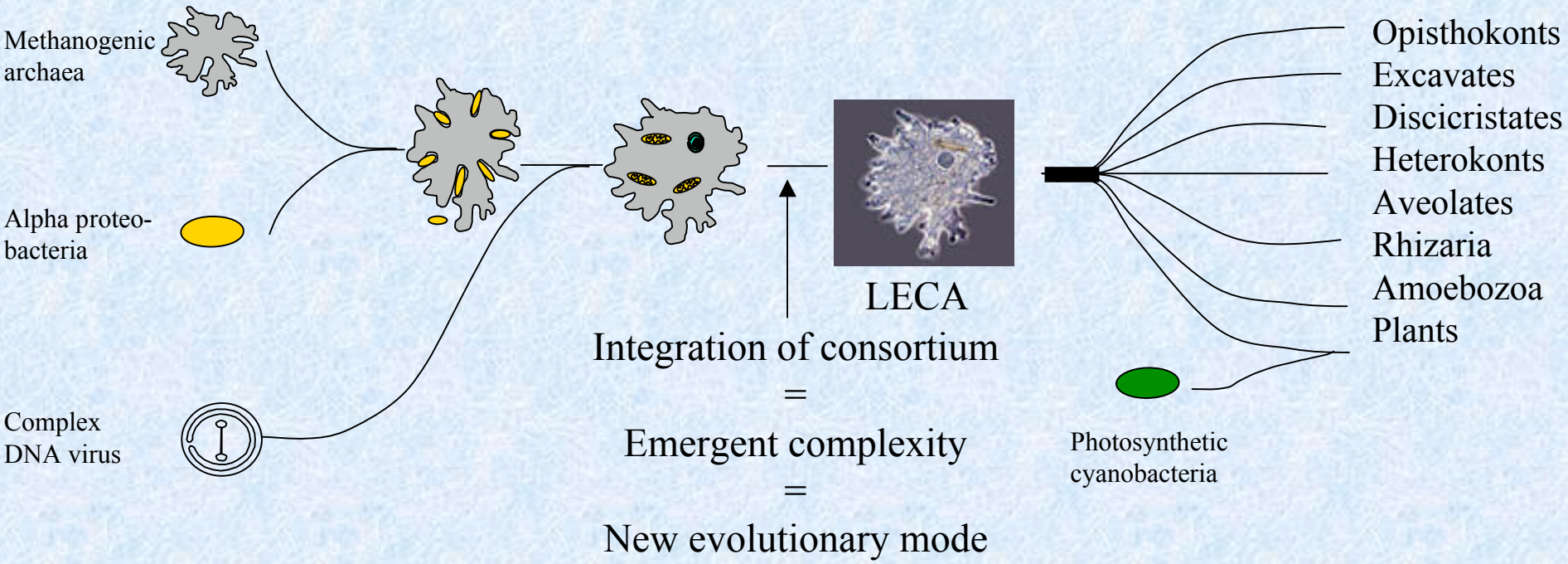
- highly homologous
- use homologous centromeric sequences for copy number control
- fail to recognize other lysogen as 'self'
- The two lysogens would be 'incompatible' due to use of homologous centromeric control regions

Conjugation between two related lineages provides rationale for origin of sex, homologous chromosomes, and two mating

The reductive division cycle due to presence four centromeres generates a rational for the origin of meiosis

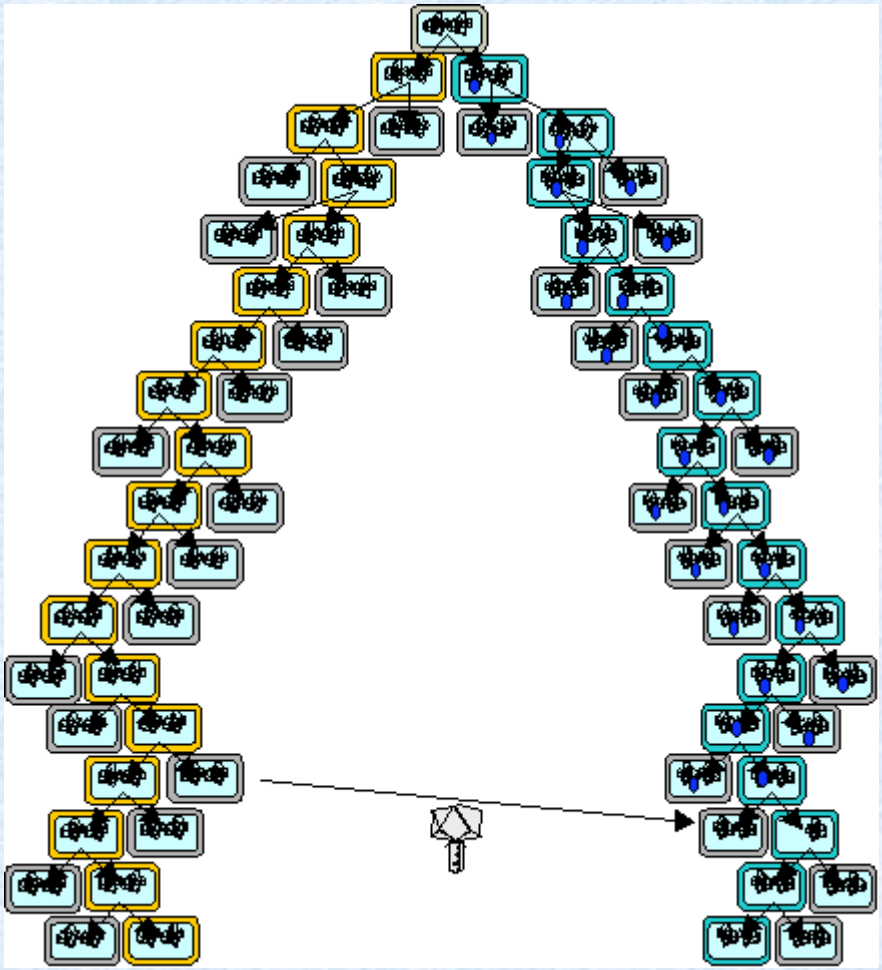


The evolution of the eukaryotes according to the VE hypothesis



- Novel ‘cellular’ design arises from tripartite consortium
- Novel replication cycles due to co-ordination of replication cycles of three lineages
- Fusion of viral and cellular replication strategies produces sex, meiosis and mitosis
- Phagocytosis (predation) propels eukaryotes in new evolutionary directions
- Phylogenetic ‘big bang’ occurs because eukaryogenesis changed genome evolution

Prokaryotic evolution proceeds primarily via vertical inheritance but plasmids and viruses are key evolutionary forces



Prokaryotes only have one ancestor alive through all of history ("serial evolution")

• A mutation of a gene would be trapped in one vertical lineage indefinitely

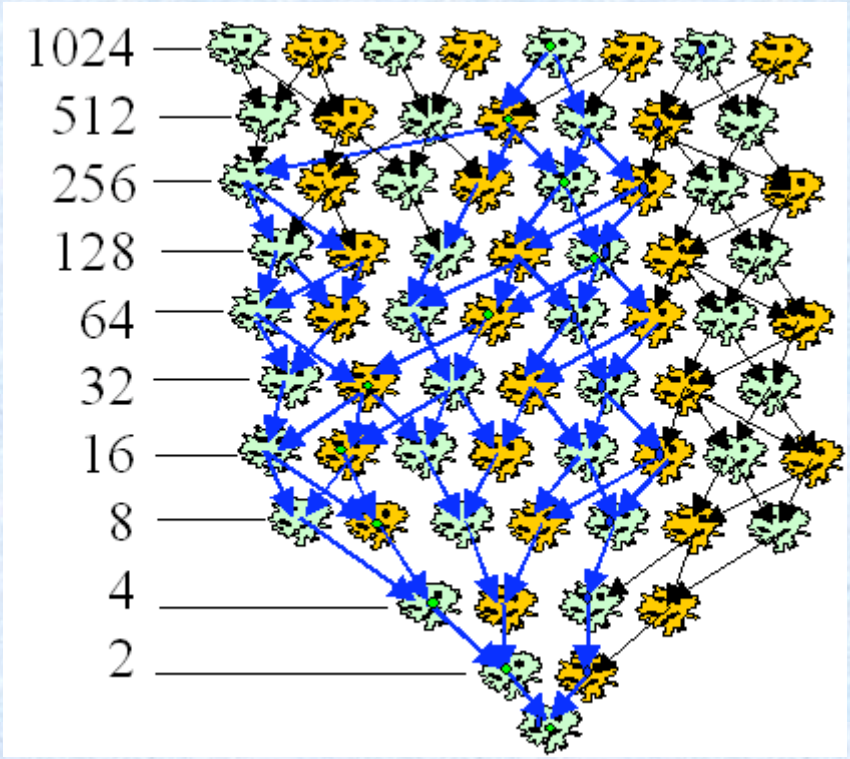
• **Müller's ratchet should lead to genome degeneration**

• Lateral gene transfer necessary to move genes to a different genetic background

• Lateral gene transfer common due to viruses and plasmids

The eukaryotic mode of replication is fundamentally different

Sexually reproducing eukaryotes evolve as populations



$10g = 10^3$
$20g = 10^6$
$30g = 10^9$
$40g = 10^{12}$
$50g = 10^{15}$

(50 human generations @ 20y/gen = 1000 years!)
(1000 AD world pop = 400 x 10^6)

Over geological time eukaryotes have an entire population as its ancestor ("parallel evolution")

Mutations in any successful lineage are ultimately available for the entire species population

The 'Neodarwinian' synthesis only applies to eukaryotic evolution

- Species concepts
- Hardy-Weinberg equilibrium
- Dominant and recessive alleles
- Sexual selection etc

Eukaryotes, prokaryotes (and viruses) evolve according to fundamentally different principles

By changing the nature of replication, the eukaryotic cell changed the nature of evolution

Serial evolution (asexual replication with no gene transfer)

In 1 million years, a bacterial lineage replicating on average at 1 hour per generation and accumulating one beneficial mutation per 100 generations,

Total new beneficial mutations:

((24 x 365 x 1,000,000) x 1) /100 = 8.8×10^7

Parallel evolution (obligate sexual evolution as a population)

In 1 million years, a human population replicating once every 20 years, with a population size of 100, 000 and accumulating one beneficial mutation per generation per 100 individuals

Total new beneficial mutations:

((1,000,000/20) x 100,000)/100 = 5×10^7

Sex allows complex slow replicating eukaryotes to assemble new genetic information as quickly as simple fast replicating asexual prokaryotes !