THE TIMING OF EUKARYOTIC EVOLUTION:
FROM ROCKS TO RELAXED MOLECULAR CLOCKS, AND RECIPROCALLY

Emmanuel JP DOUZERY [1]
Frédéric DELSUC [1,2]
& Hervé PHILIPPE [2]

[1] Paléontologie, Phylogénie et Paléobiologie,
Institut des Sciences de l'Evolution,
Université Montpellier II, France

[2] CIAR, Département de Biochimie,
Centre Robert-Cedergren,
Université de Montréal, Québec, Canada

1. WHEN DID ANIMALS
AND OTHER EUKARYOTES APPEAR ?

2. VARIETY OF THE MOLECULAR ANSWERS

3. MOLECULAR DATING:
FROM GLOBAL TO RELAXED CLOCKS

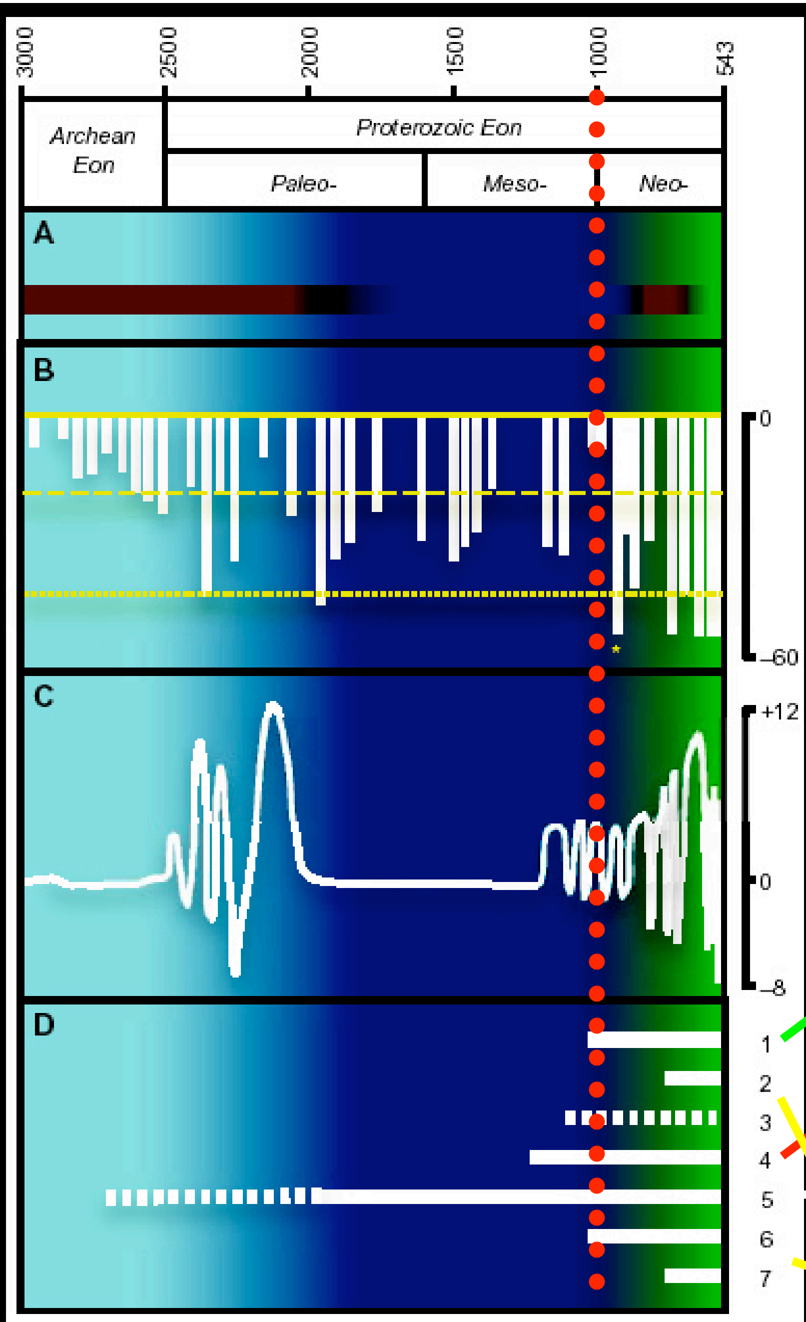
4. LARGE-SCALE PROTEIN CLOCKS
FOR THE CHRONOLOGY
OF EUKARYOTES DIVERSIFICATION

EVOLUTION OF PROTEROZOIC OCEANS

Fe^{2+} deposits

$\Delta^{34}S$ ($\delta^{34}S$ S^{2-}/SO_4^{2-})

$\delta^{13}C$ (carbonates)



Eukaryotes diversified ~1 000 Mya

Green algae

Red algae

Non identified Eukar.

Protists

anoxic

anoxic & sulfidic (S^{2-})

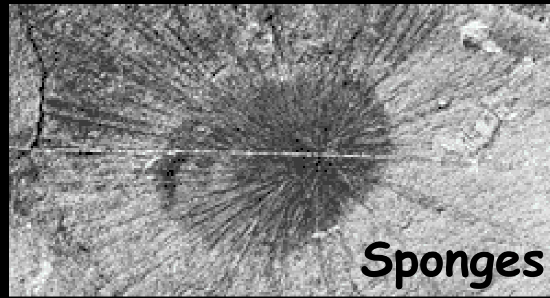
oxic

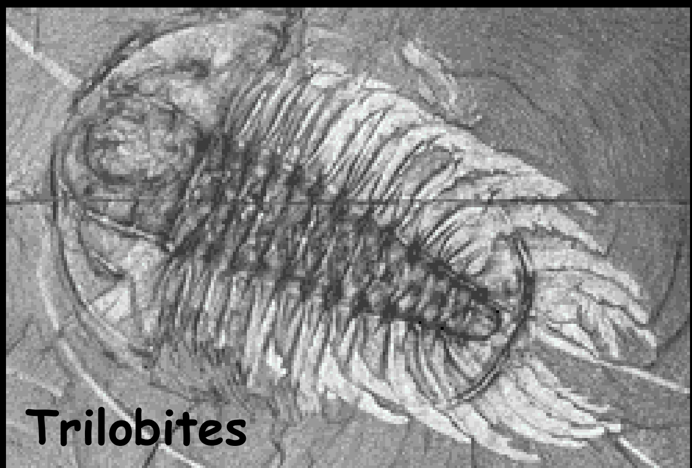
[Anbar & Knoll 2002]

540 Ma
(Burgess, CA)

CAMBRIAN EXPLOSION OF ANIMALS

Hallucigenia





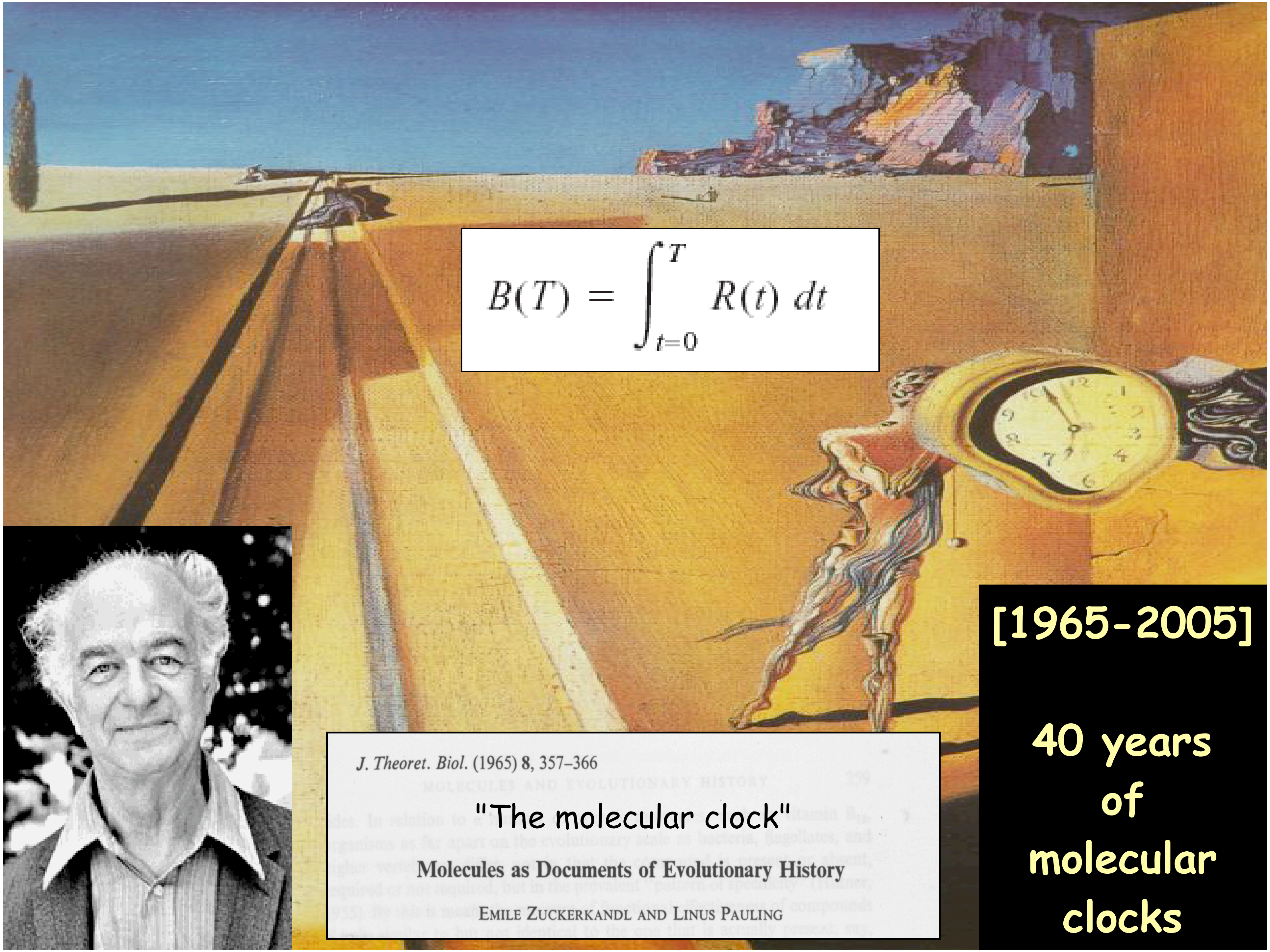
Their explosive diversification initiates the **Primary Era.**

1. WHEN DID ANIMALS AND OTHER EUKARYOTES APPEAR ?

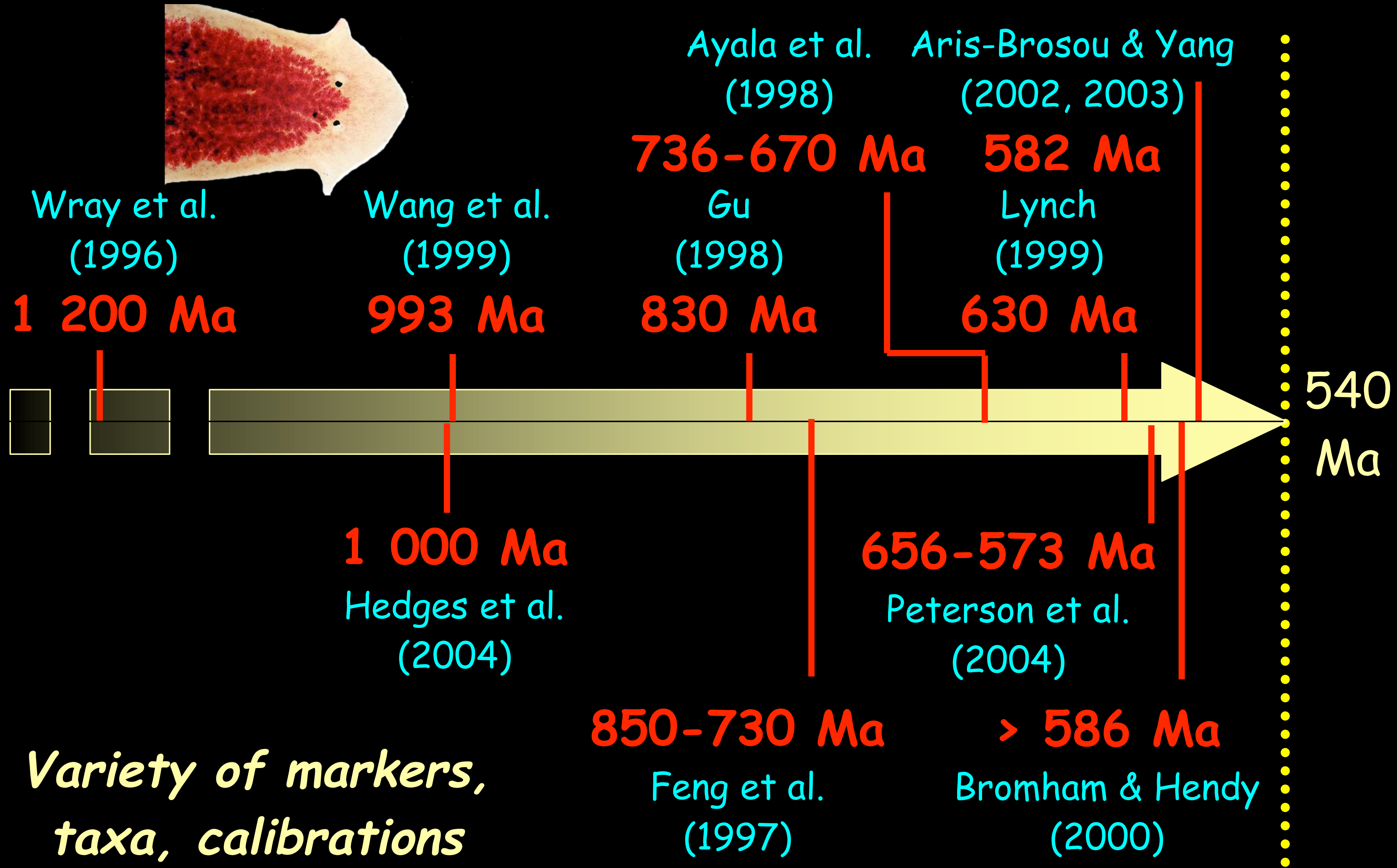
2. VARIETY OF THE MOLECULAR ANSWERS

3. MOLECULAR DATING: FROM GLOBAL TO RELAXED CLOCKS

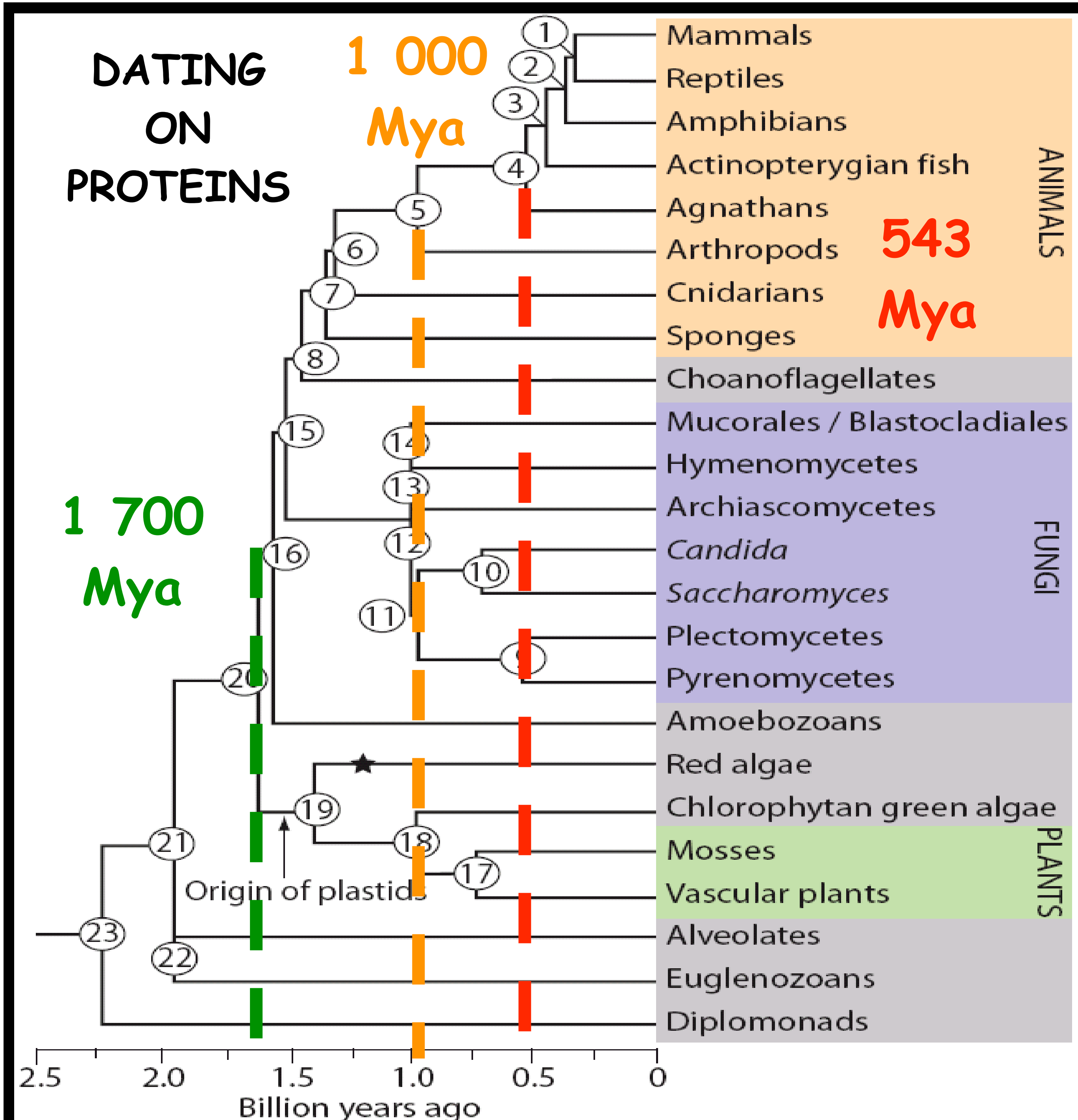
4. LARGE-SCALE PROTEIN CLOCKS FOR THE CHRONOLOGY OF EUKARYOTES DIVERSIFICATION


$$B(T) = \int_{t=0}^{T} R(t)\, dt$$
J. Theoret. Biol. (1965) 8, 357–366
"The molecular clock"
Molecules as Documents of Evolutionary History
EMILE ZUCKERKANDL AND LINUS PAULING
[1965-2005]
40 years
of
molecular
clocks

WHEN DID BILATERIAN ANIMALS DIVERSIFY ?



Variety of markers, taxa, calibrations and methods !



1. Discrepancy between molecules and fossils ;

2. No explicit uncertainty

[Hedges et al. 2004]

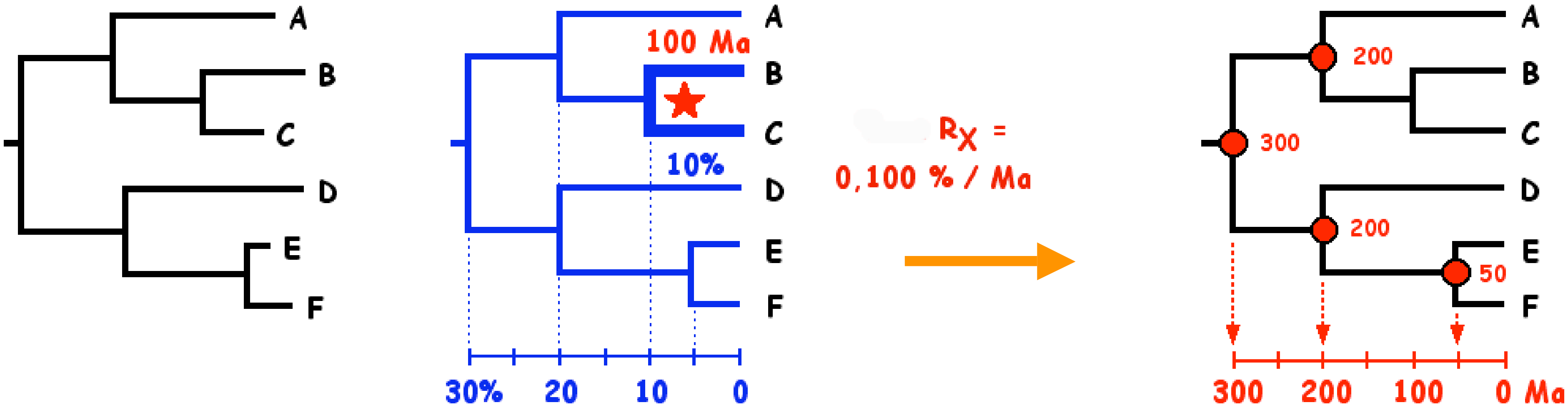
1. WHEN DID ANIMALS
AND OTHER EUKARYOTES APPEAR ?

2. VARIETY OF THE MOLECULAR ANSWERS

**3. MOLECULAR DATING:
FROM GLOBAL TO RELAXED CLOCKS**

4. LARGE-SCALE PROTEIN CLOCKS
FOR THE CHRONOLOGY
OF EUKARYOTES DIVERSIFICATION

Gene X



ML tree

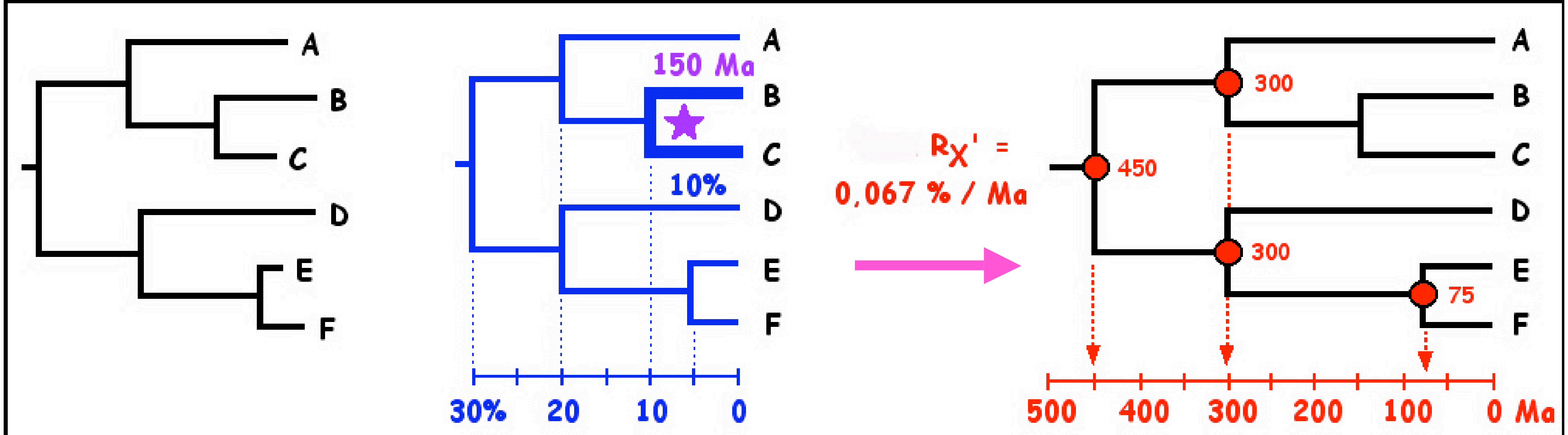
ML clock-like tree

Fossil calibration

Time estimates

DATING WITH MOLECULAR CLOCKS

Gene X



ML tree

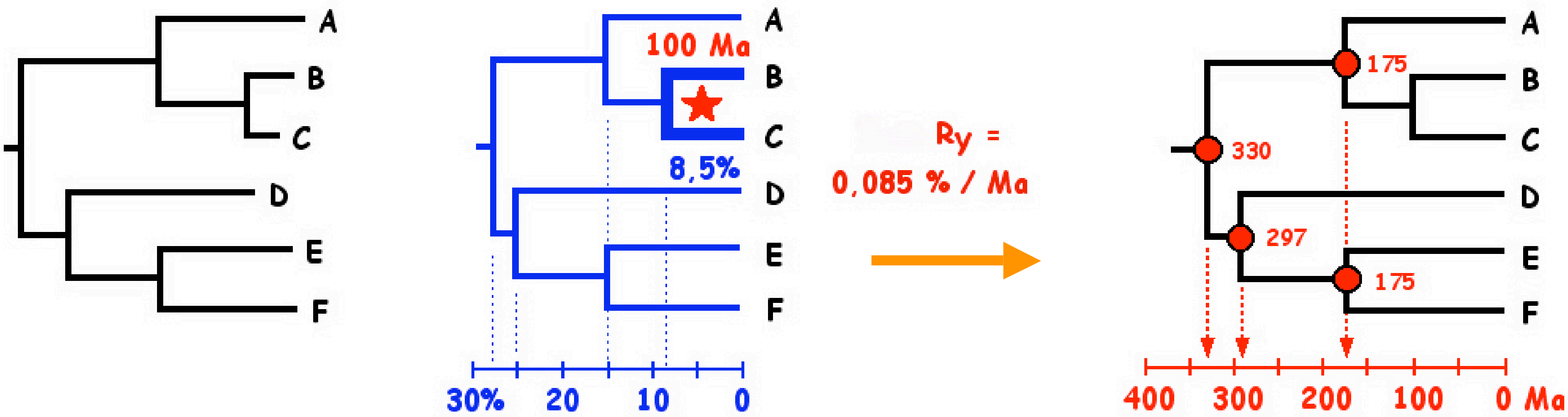
ML clock-like tree

A different fossil calibration

Slower rate, deeper time estimates

THE PALEONTOLOGICAL ERROR

Gene Y



Different ML branch lengths

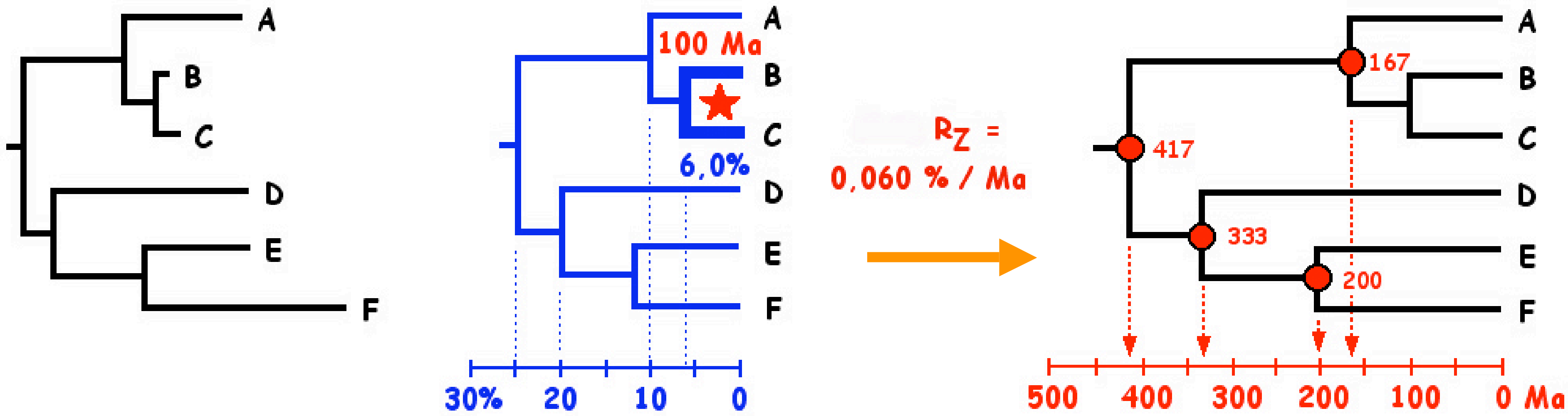
ML clock-like tree

Initial fossil calibration

Different rate, different time estimates

THE STOCHASTIC ERROR

Gene Z



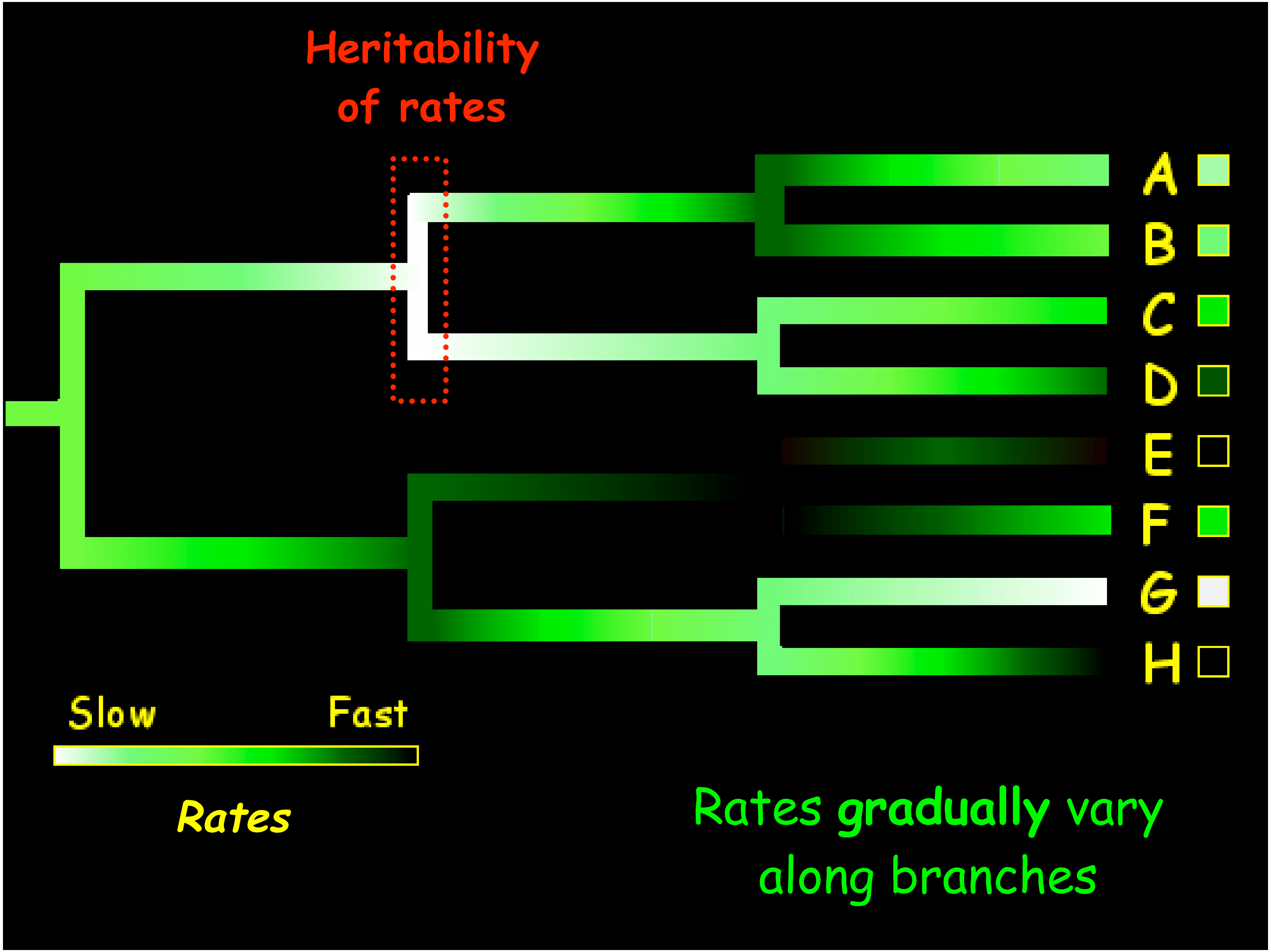
Different ML branch lengths

Distorted ML clock-like tree

Initial fossil calibration

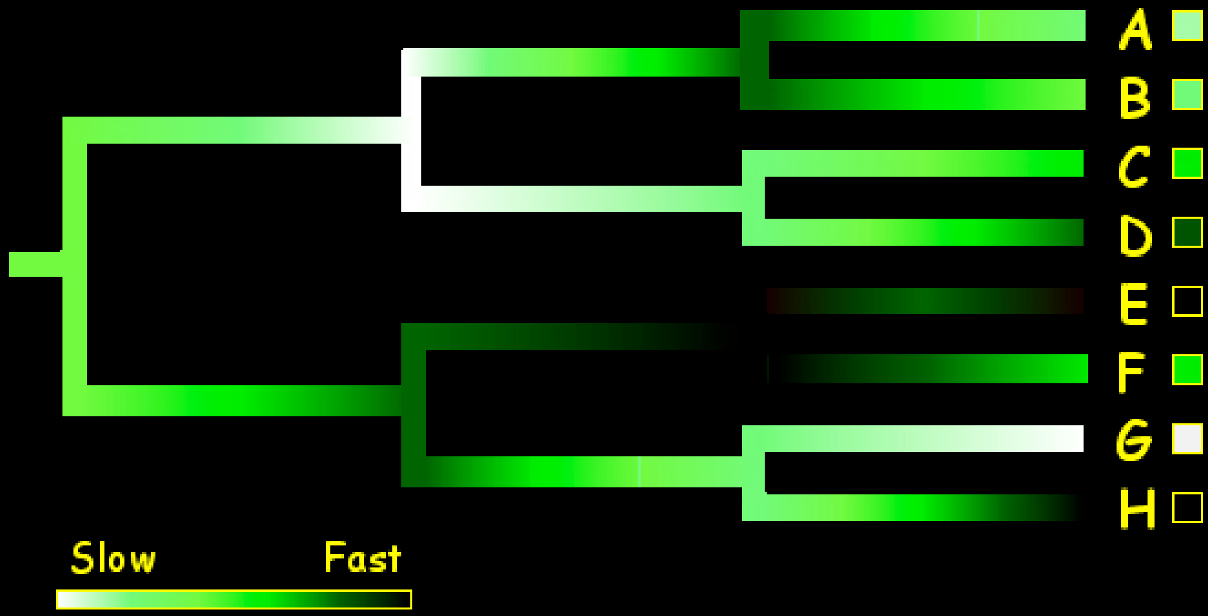
Different rate, different time estimates

THE LACK OF CLOCK


Heritability of rates
A
B
C
D
E
F
G
H
Slow
Fast
Rates
Rates gradually vary along branches

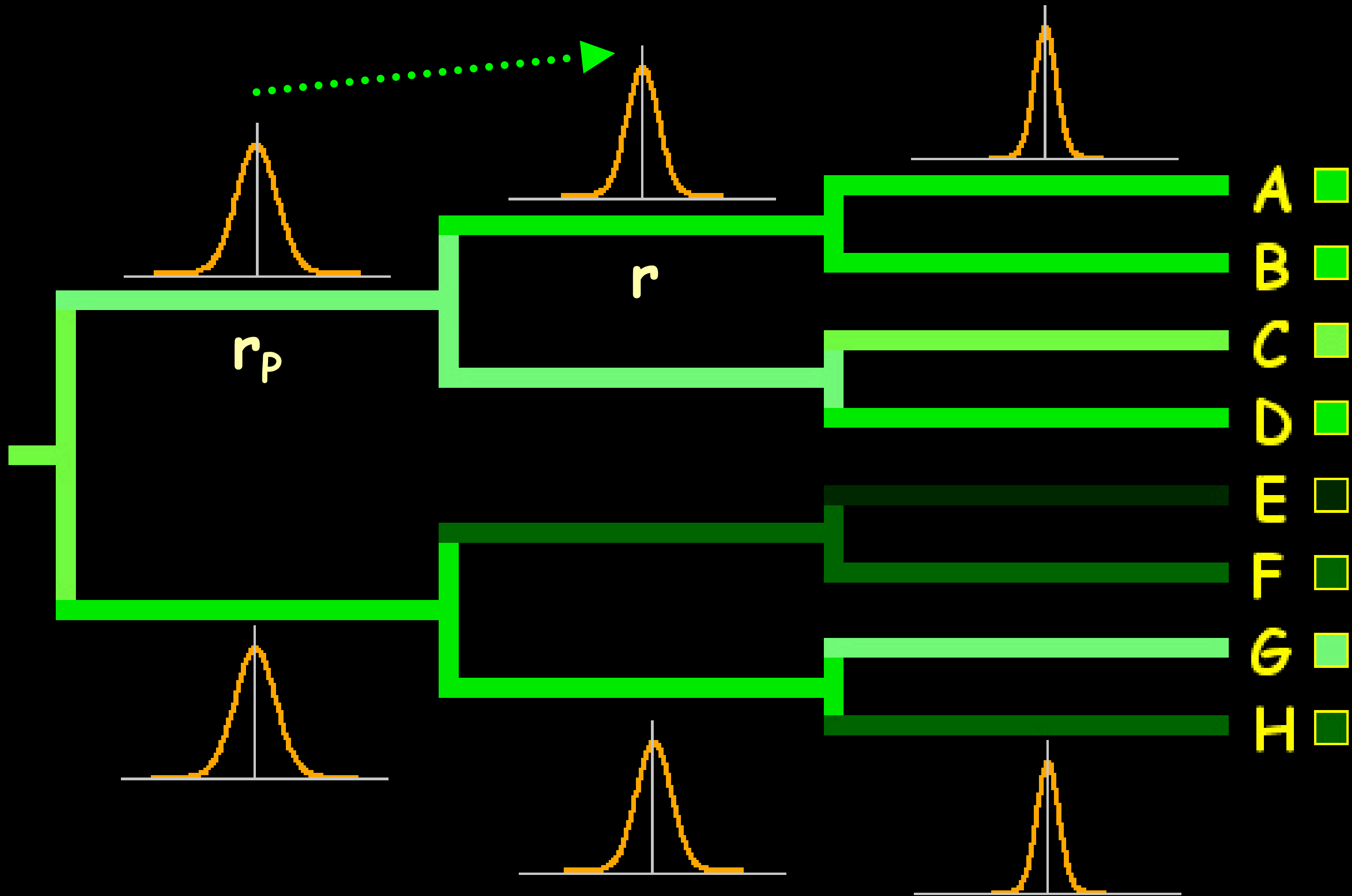
"Closely related evolutionary lineages are likely to evolve at similar rates. A consequence of this **rate autocorrelation** is that branches that are nearby in a tree will have correlated lengths"

[Thorne, Kishino & Painter 1998]



Stationary log-normal distribution of rates

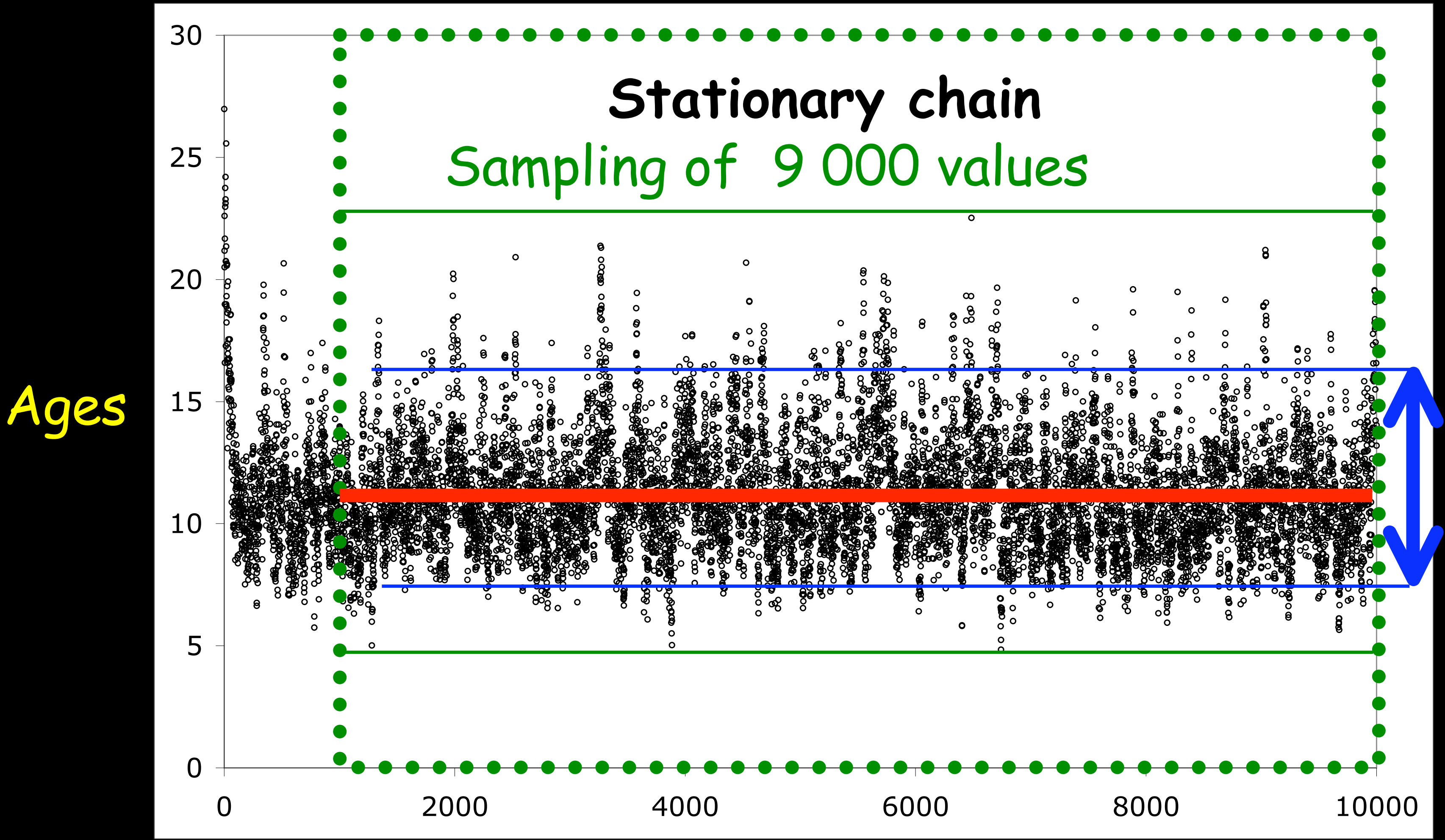
$$\frac{1}{2\beta t}\left[\log(r/r_p)+\beta t/2\right]^2+\log\left(r\sqrt{2\beta t}\right)$$



$$VAR(\ln r_P) = \nu \cdot t_P$$

[Kishino, Thorne & Bruno 2001]

BAYESIAN APPROACH
FOR THE RELAXED MOLECULAR CLOCK



μ = 11,0 Ma ± 2,1

95% credibility interval = 7,5 - 15,9 Ma

1. WHEN DID ANIMALS
AND OTHER EUKARYOTES APPEAR ?

2. VARIETY OF THE MOLECULAR ANSWERS

3. MOLECULAR DATING:
FROM GLOBAL TO RELAXED CLOCKS

**4. LARGE-SCALE PROTEIN CLOCKS
FOR THE CHRONOLOGY
OF EUKARYOTES DIVERSIFICATION**

HOW TO IMPROVE THE MOLECULAR DATING OF THE EUKARYOTIC EVOLUTION ?

1. REDUCE THE STOCHASTIC ERRORS !

→ Use of a greater number of proteins

2. REDUCE THE PHYLOGENETIC ERRORS !

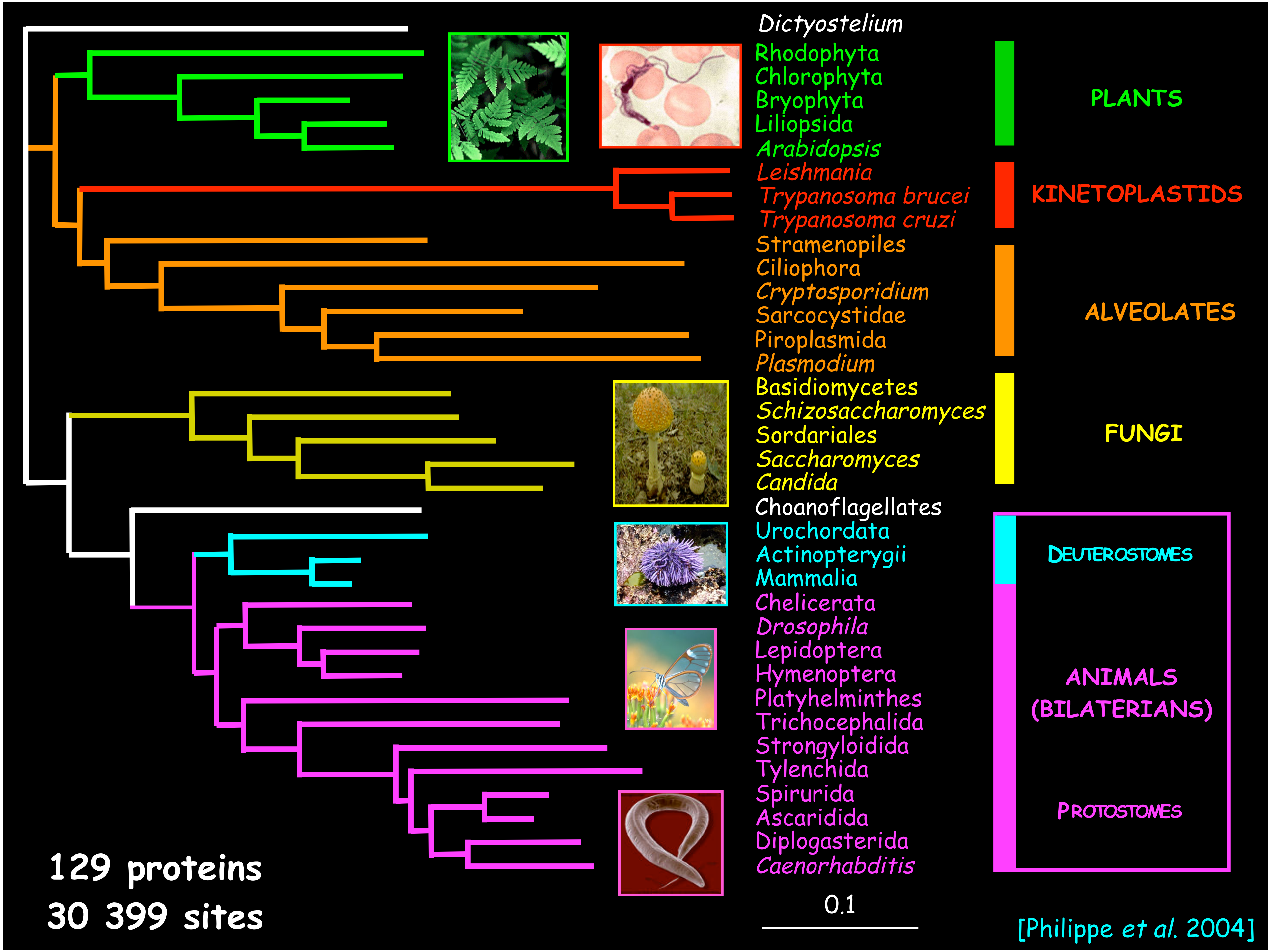
→ Use of a greater number of species

3. REDUCE THE PALEONTOLOGICAL ERRORS !

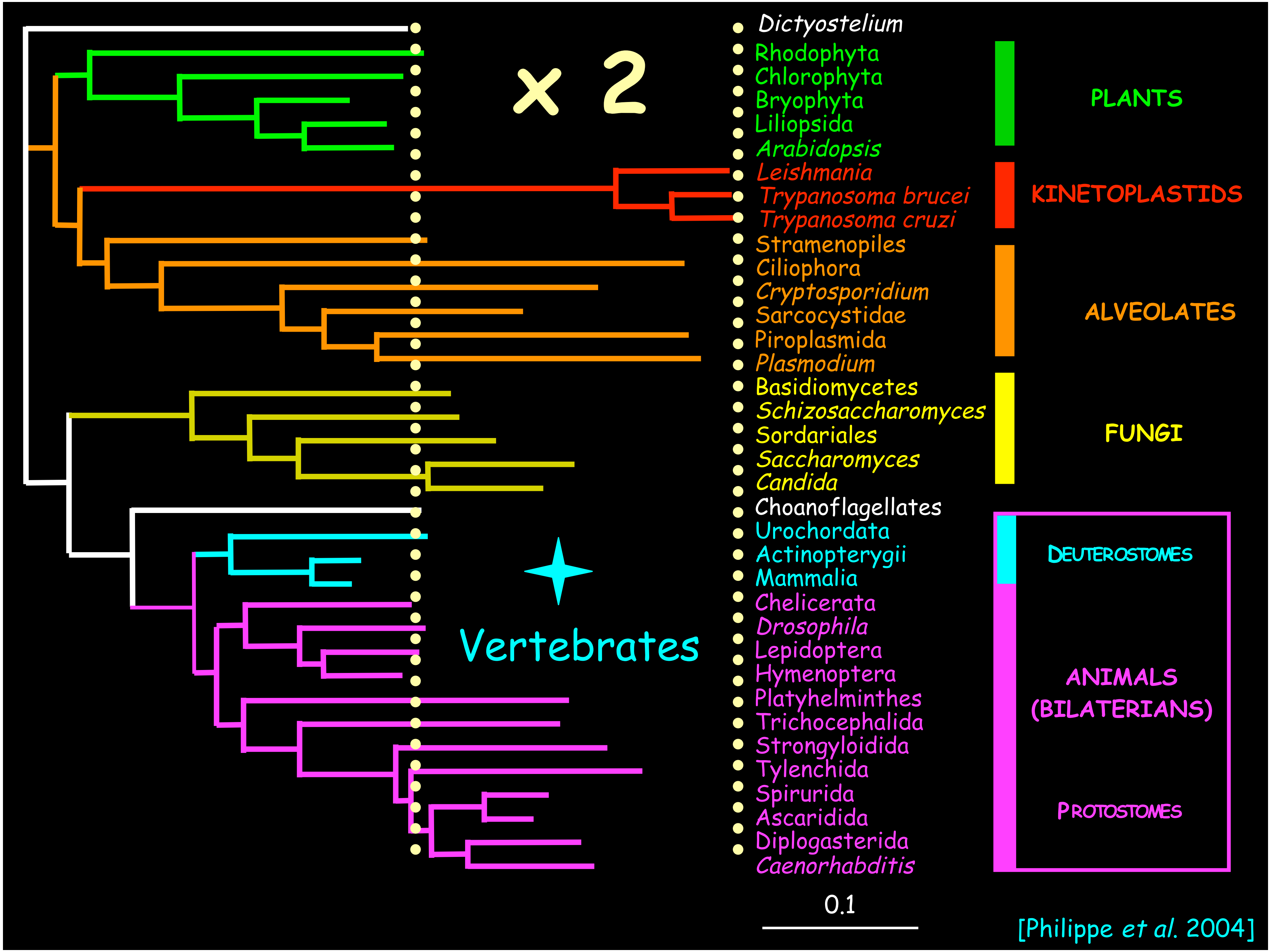
→ Use of multiple calibration intervals.

4. REDUCE THE DATING ERRORS !

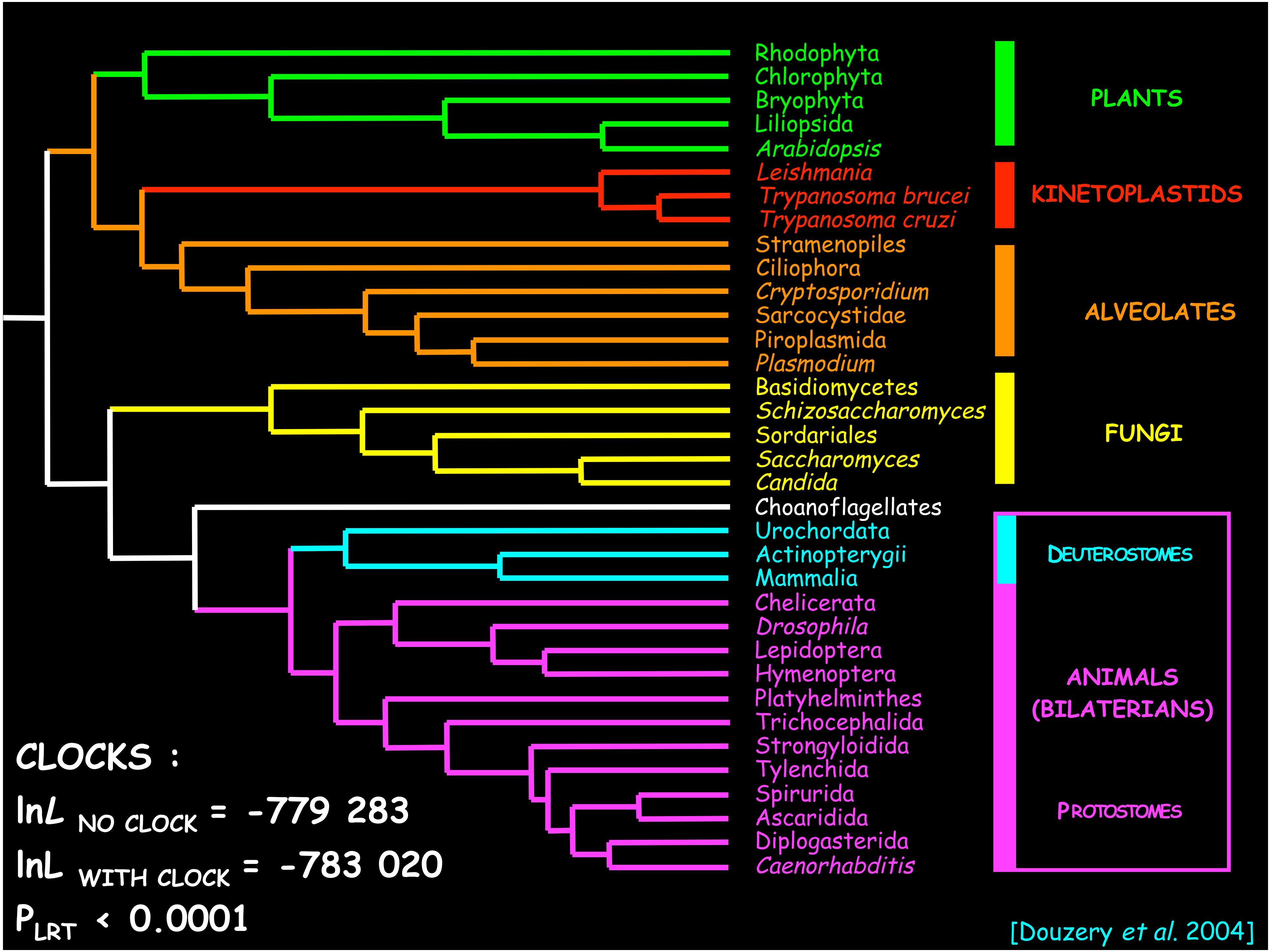
→ Use of relaxed molecular clocks.

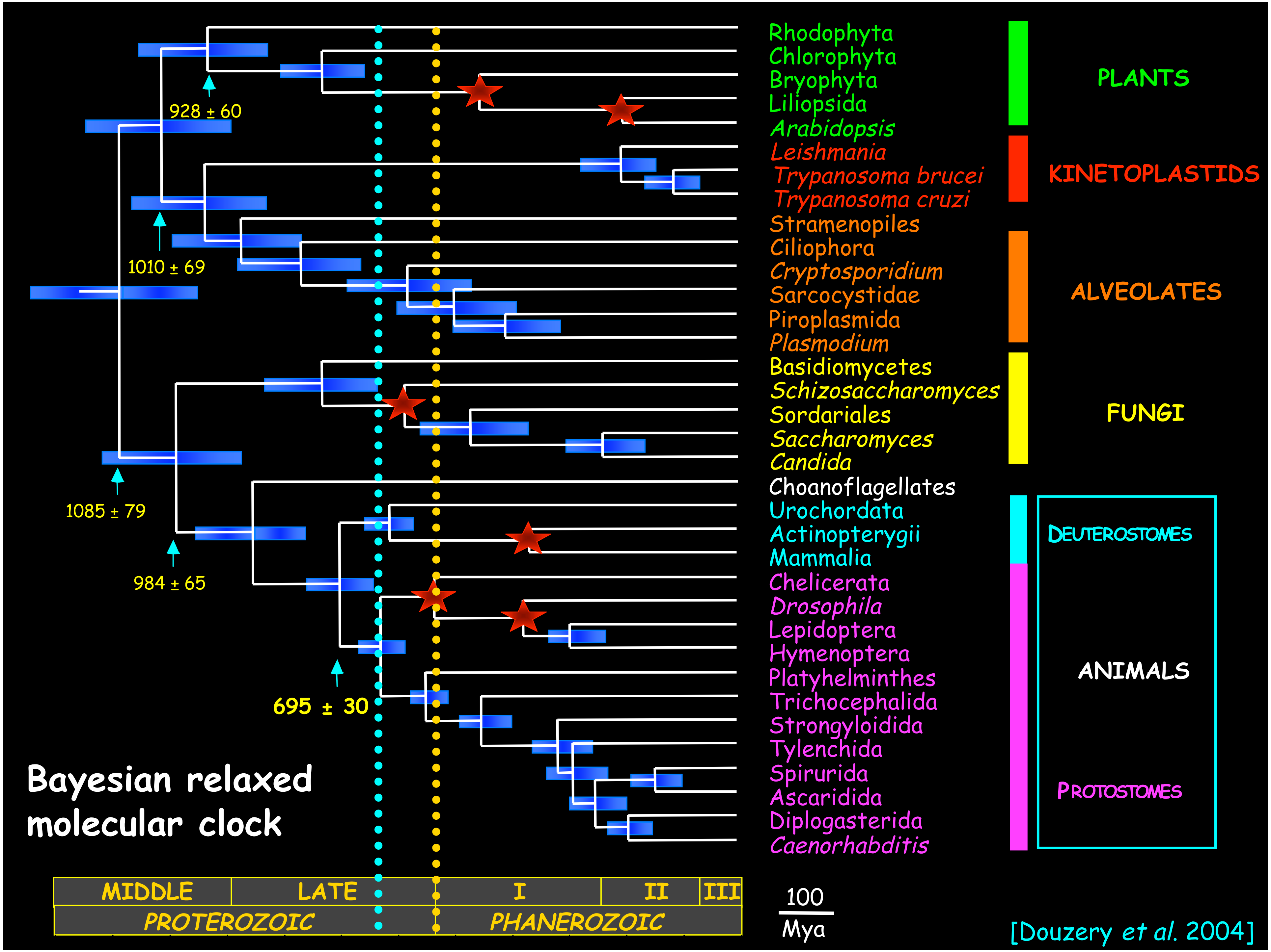


Dictyostelium
Rhodophyta
Chlorophyta
Bryophyta
Liliopsida
Arabidopsis
Leishmania
Trypanosoma brucei
Trypanosoma cruzi
Stramenopiles
Ciliophora
Cryptosporidium
Sarcocystidae
Piroplasmida
Plasmodium
Basidiomycetes
Schizosaccharomyces
Sordariales
Saccharomyces
Candida
Choanoflagellates
Urochordata
Actinopterygii
Mammalia
Chelicerata
Drosophila
Lepidoptera
Hymenoptera
Platyhelminthes
Trichocephalida
Strongyloidida
Tylenchida
Spirurida
Ascaridida
Diplogasterida
Caenorhabditis
PLANTS
KINETOPLASTIDS
ALVEOLATES
FUNGI
DEUTEROSTOMES
ANIMALS
(BILATERIANS)
PROTOSTOMES
129 proteins
30 399 sites
0.1
[Philippe et al. 2004]



x 2
Vertebrates
Dictyostelium
Rhodophyta
Chlorophyta
Bryophyta
Liliopsida
Arabidopsis
Leishmania
Trypanosoma brucei
Trypanosoma cruzi
Stramenopiles
Ciliophora
Cryptosporidium
Sarcocystidae
Piroplasmida
Plasmodium
Basidiomycetes
Schizosaccharomyces
Sordariales
Saccharomyces
Candida
Choanoflagellates
Urochordata
Actinopterygii
Mammalia
Chelicerata
Drosophila
Lepidoptera
Hymenoptera
Platyhelminthes
Trichocephalida
Strongyloidida
Tylenchida
Spirurida
Ascaridida
Diplogasterida
Caenorhabditis
PLANTS
KINETOPLASTIDS
ALVEOLATES
FUNGI
DEUTEROSTOMES
ANIMALS
(BILATERIANS)
PROTOSTOMES
0.1
[Philippe et al. 2004]


Rhodophyta
Chlorophyta
Bryophyta
Liliopsida
Arabidopsis
Leishmania
Trypanosoma brucei
Trypanosoma cruzi
Stramenopiles
Ciliophora
Cryptosporidium
Sarcocystidae
Piroplasmida
Plasmodium
Basidiomycetes
Schizosaccharomyces
Sordariales
Saccharomyces
Candida
Choanoflagellates
Urochordata
Actinopterygii
Mammalia
Chelicerata
Drosophila
Lepidoptera
Hymenoptera
Platyhelminthes
Trichocephalida
Strongyloidida
Tylenchida
Spirurida
Ascaridida
Diplogasterida
Caenorhabditis
PLANTS
KINETOPLASTIDS
ALVEOLATES
FUNGI
DEUTEROSTOMES
ANIMALS (BILATERIANS)
PROTOSTOMES
CLOCKS :
lnL NO CLOCK = -779 283
lnL WITH CLOCK = -783 020
P LRT < 0.0001
[Douzery et al. 2004]



Bayesian relaxed molecular clock
928 ± 60
1010 ± 69
1085 ± 79
984 ± 65
695 ± 30
Rhodophyta
Chlorophyta
Bryophyta
Liliopsida
Arabidopsis
Leishmania
Trypanosoma brucei
Trypanosoma cruzi
Stramenopiles
Ciliophora
Cryptosporidium
Sarcocystidae
Piroplasmida
Plasmodium
Basidiomycetes
Schizosaccharomyces
Sordariales
Saccharomyces
Candida
Choanoflagellates
Urochordata
Actinopterygii
Mammalia
Chelicerata
Drosophila
Lepidoptera
Hymenoptera
Platyhelminthes
Trichocephalida
Strongyloidida
Tylenchida
Spirurida
Ascaridida
Diplogasterida
Caenorhabditis
PLANTS
KINETOPLASTIDS
ALVEOLATES
FUNGI
DEUTEROSTOMES
ANIMALS
PROTOSTOMES
MIDDLE
LATE
I
II
III
PROTEROZOIC
PHANEROZOIC
100
Mya
[Douzery et al. 2004]

CONCLUSIONS

The Bayesian relaxed molecular clock

on 129 proteins:

→ reduces the discrepancies between **paleontological** and **molecular** estimates of divergence times of major eucaryote groups ;

→ demonstrates the need of incorporating **variations of evolutionary rates** ;

→ emphasizes the importance of dense **taxonomic** and **genetic** samplings.

PERSPECTIVES

→ Incorporation of

phylogenetic uncertainty ;

→ Development and comparison

of **models** of **evolutionary rate variation** ;

→ Recognition of **different partitions**

among genes or proteins.