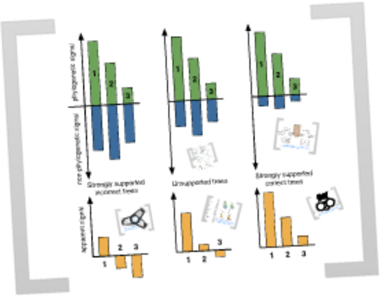
With challenging phylogenetic problems, the difficulty is to make the needle bigger without enlarging the haystack!

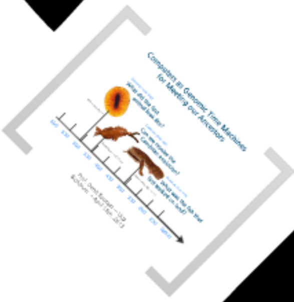
more sequences
are not enough



phylogenetic vs.
non-phylogenetic signal

The coelacanth is a large marine fish with fleshy fins that resembles the limbs of terrestrial vertebrates.

phylogenomics can resolve
the evolution of animals

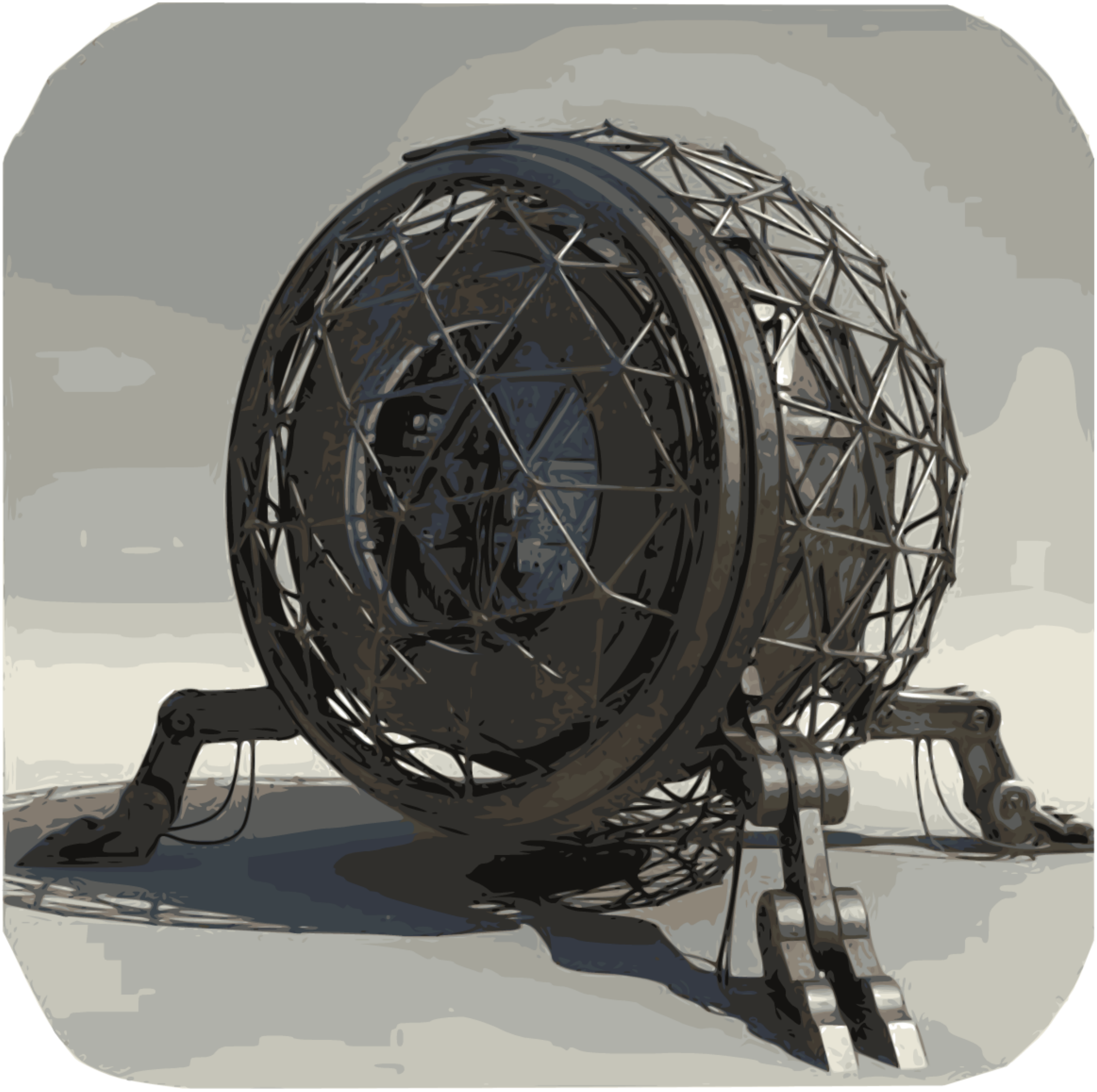


Computers as Genome Time Machines for Meeting our Ancestors

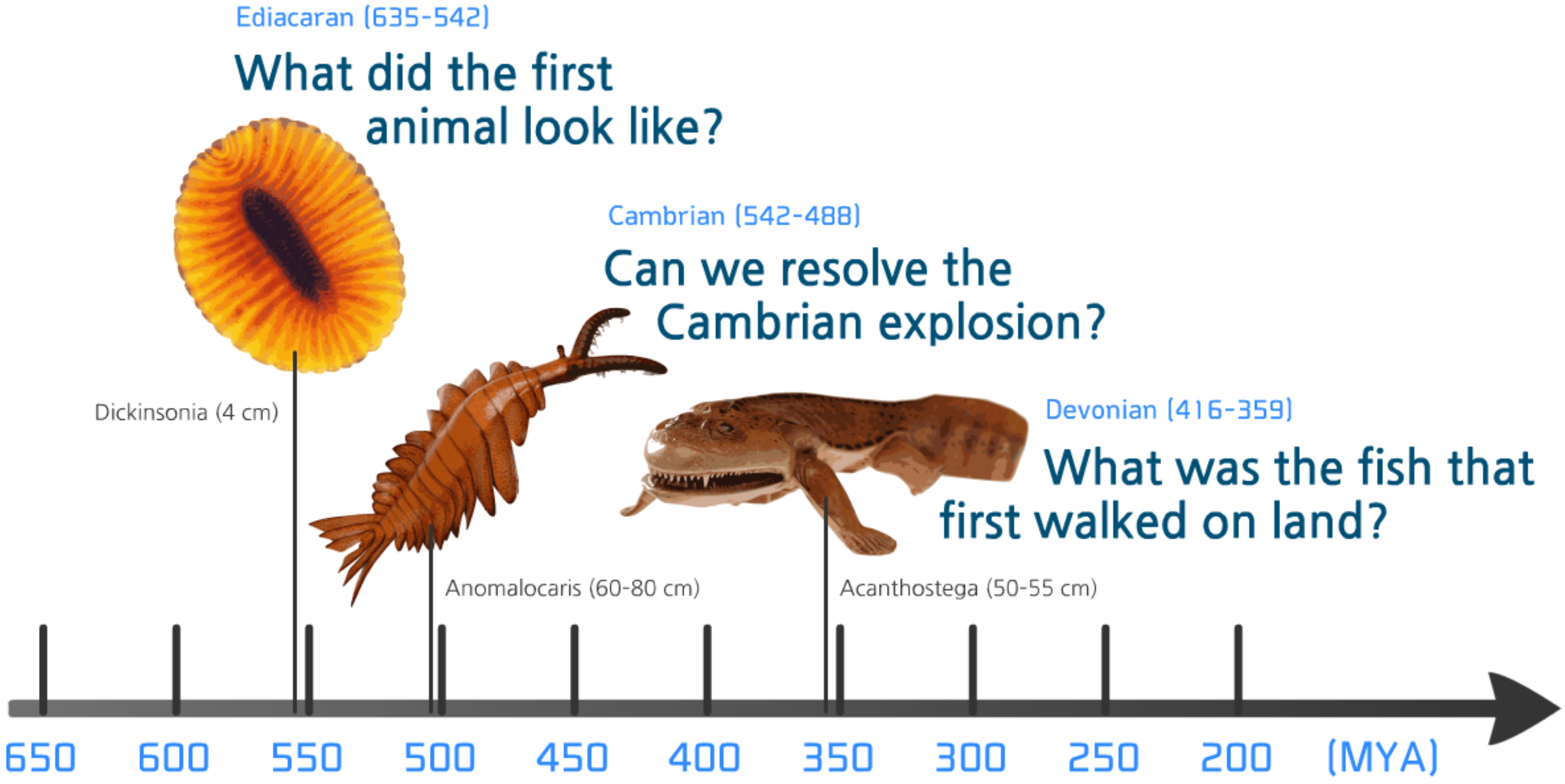
What did the first animal look like?

Can we resolve the Cambrian explosion?

What was the fish that first walked on land?



Computers as Genomic Time Machines for Meeting our Ancestors



Ediacaran (635-542)

What did the first animal look like?

Cambrian (542-488)

Can we resolve the Cambrian explosion?

Devonian (416-359)

What was the fish that first walked on land?

Prof. Denis Baurain — ULg
Bioforum — April 18th, 2013

ve the explosion?

Devonian (416-359)

What was the fish that first walked on land?

nthostega (50-55 cm)

Latimeria chalumnae

Until its rediscovery, the coelacanth was thought to have been extinct since the Late Cretaceous period (70 MYA).

Protopterus annectens

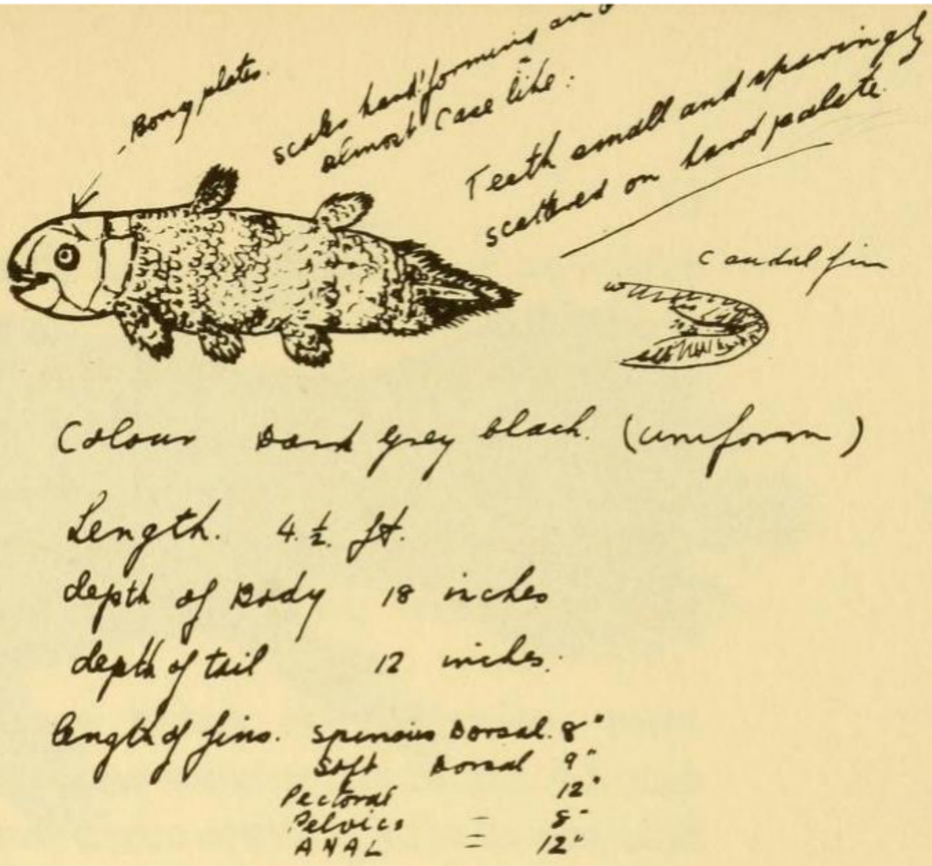
The lungfish has lungs and can breath air. It is also a lobe-finned fish, yet living in freshwater.

Phylogenomics is useful but can suffer from artifacts!

Let's look at their causes!

The coelacanth is a large marine fish with fleshy fins that resembles the limbs of terrestrial vertebrates.

Miss M. Courtenay-Latimer



Colour dark grey black. (uniform)

Length. 4.½. ft.
depth of body 18 inches
depth of tail 12 inches.
length of fins. spinous dorsal. 8"
Soft dorsal 9"
Pectoral 12"
Pelvics = 8"
ANAL = 12"

Miss Latimer's sketch and notes

It was named after its discoverer, who was the curator of a small museum in South Africa.

However, the coelacanth is not a living fossil, as this concept is fundamentally incorrect!

Until its rediscovery, the coelacanth was thought to have been extinct since the Late Cretaceous period (70 MYA).

wrong

right

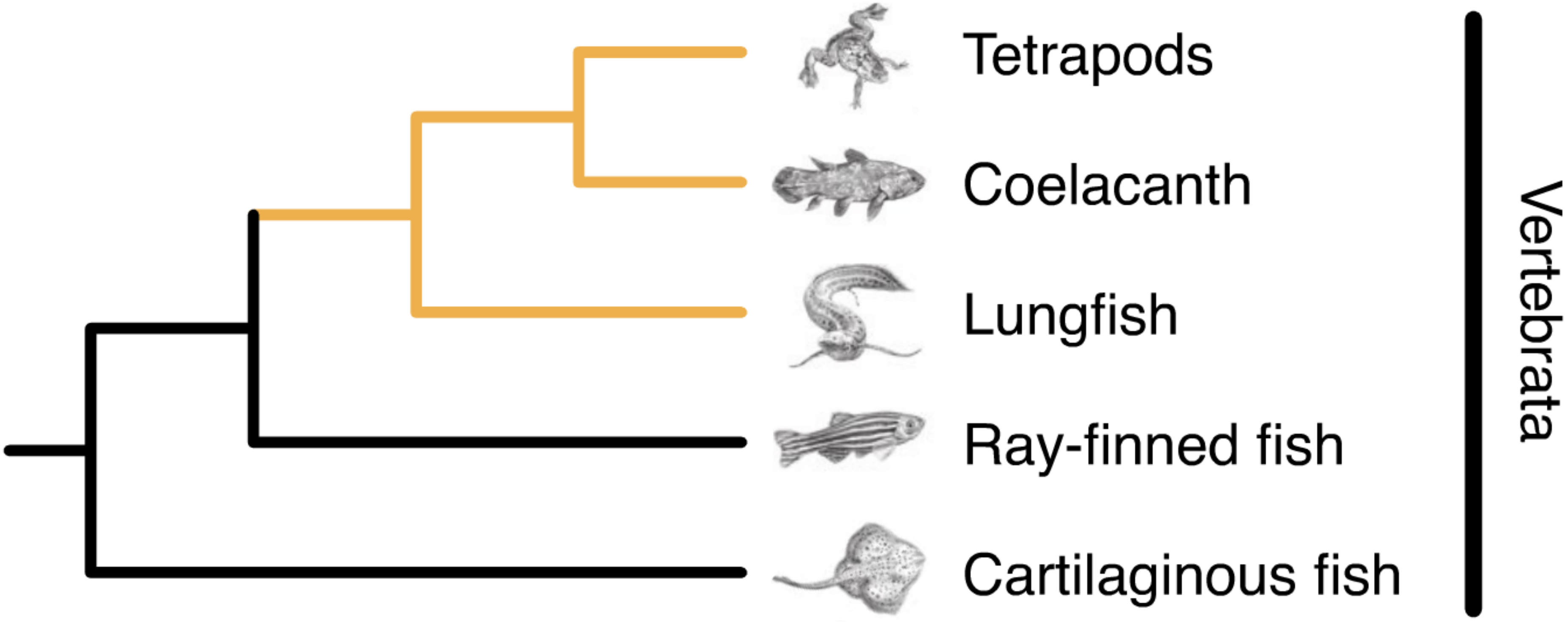
However, the coelacanth is not a living fossil, as this concept is fundamentally incorrect!

Protopterus annectens

The lungfish has lungs and can breath air. It is also a lobe-finned fish, yet living in freshwater.

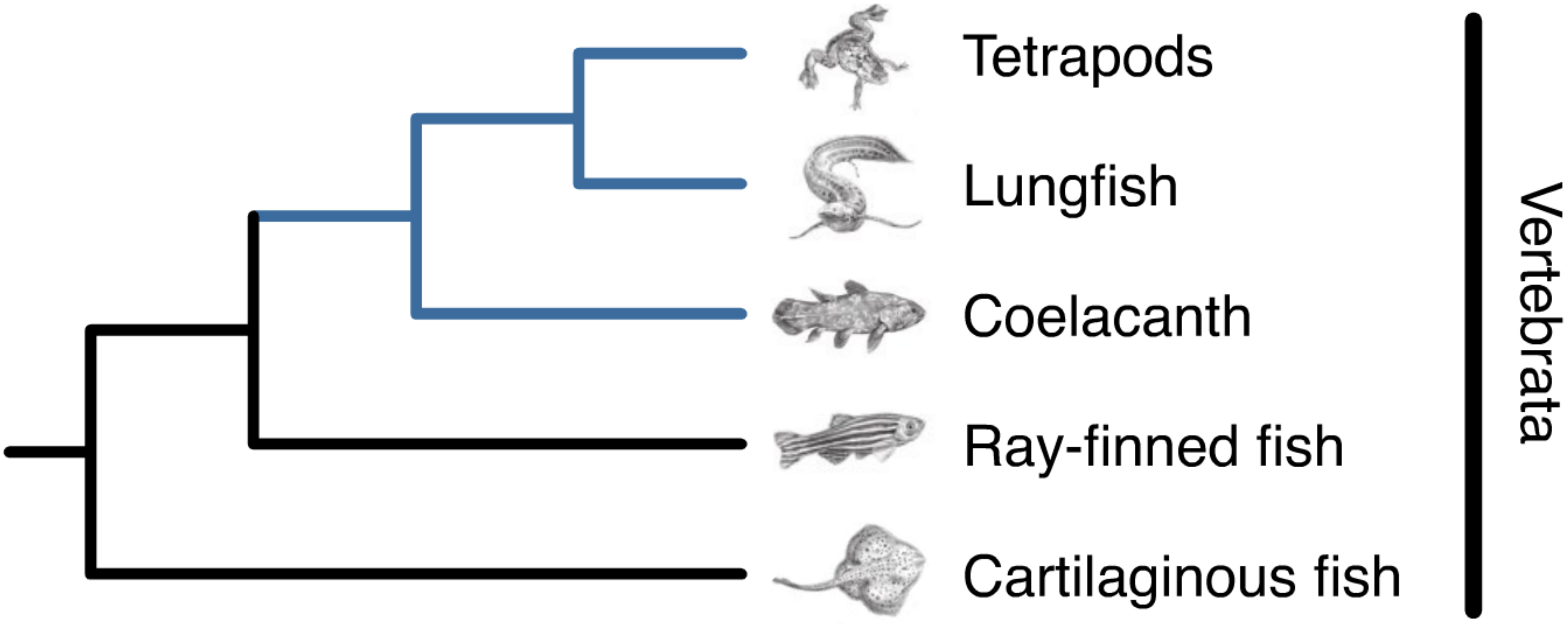
The coelacanth as the closest relative of tetrapods

Gorr et al. (1991) – haemoglobins a & b (nucl.)



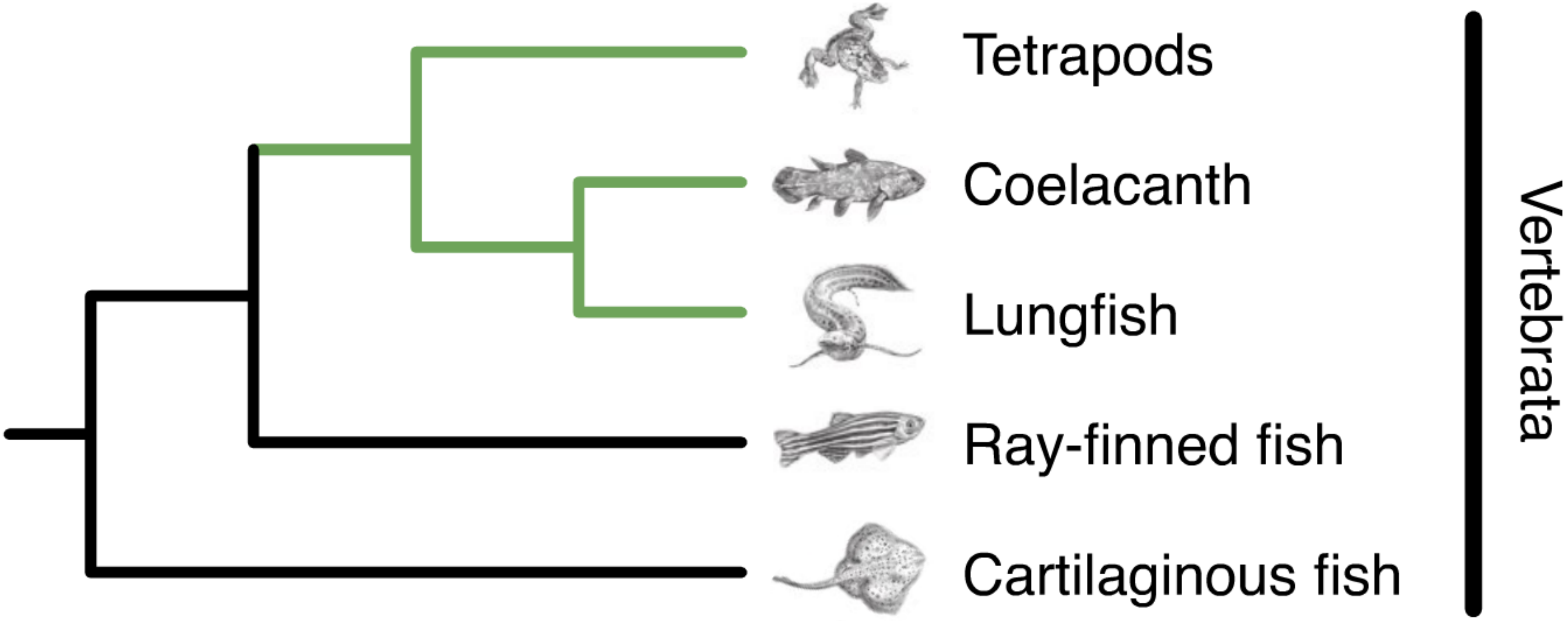
The lungfish as the closest relative of tetrapods

e.g., Meyer & Dolven (1992) – 12S rRNA & cob (mitochondr.)
Brinkmann et al. (2004) – RAS1 & 2 (nucl.)



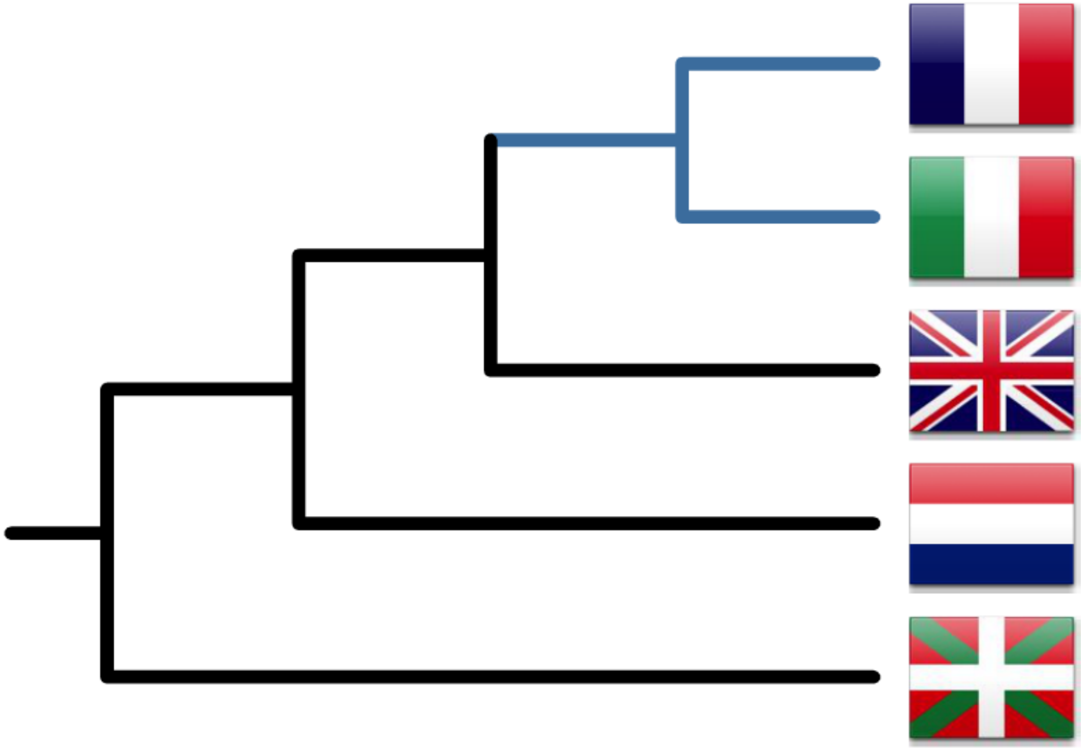
Both the coelacanth and the lungfish!

e.g., Yokobori et al. (1994) – cox1 (mitochondr.)
Zardoya & Meyer (1996) – 28S rRNA (nucl.)



Potential relationships among European languages

based on the naive analysis of 1 word

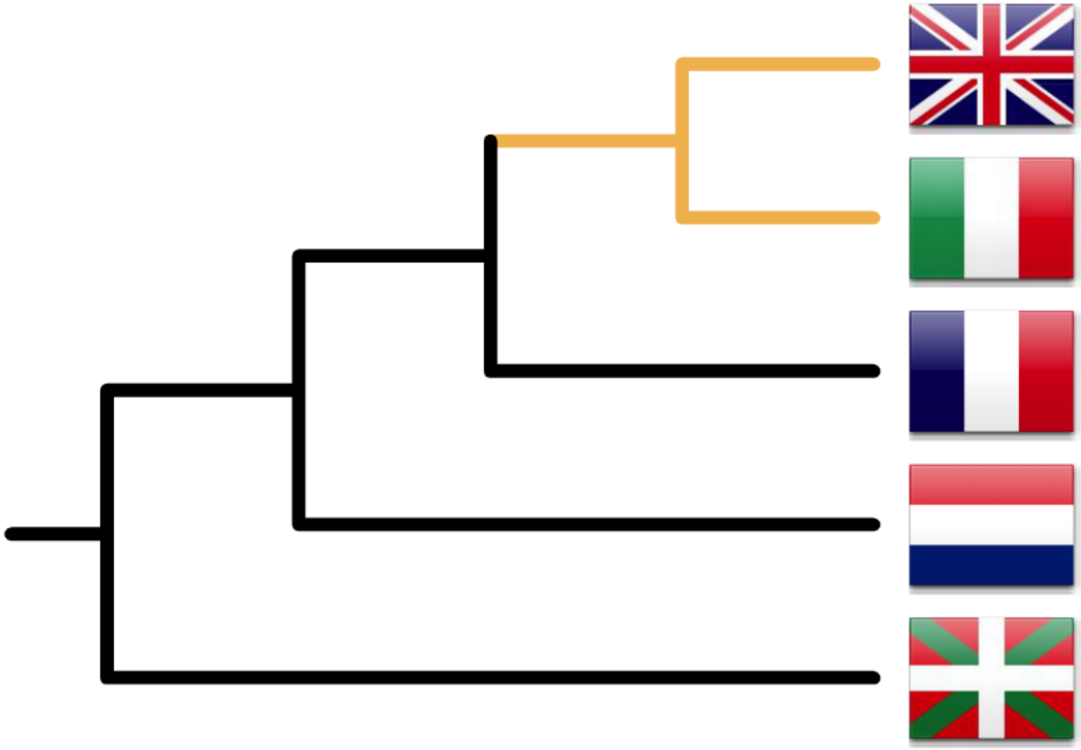


Français	TU
Italiano	TU
English	YOU
Nederlands	JE
Euskara	DUZU

The Basque language is our outgroup here.

Potential relationships among European languages

based on the naive analysis of 1 word



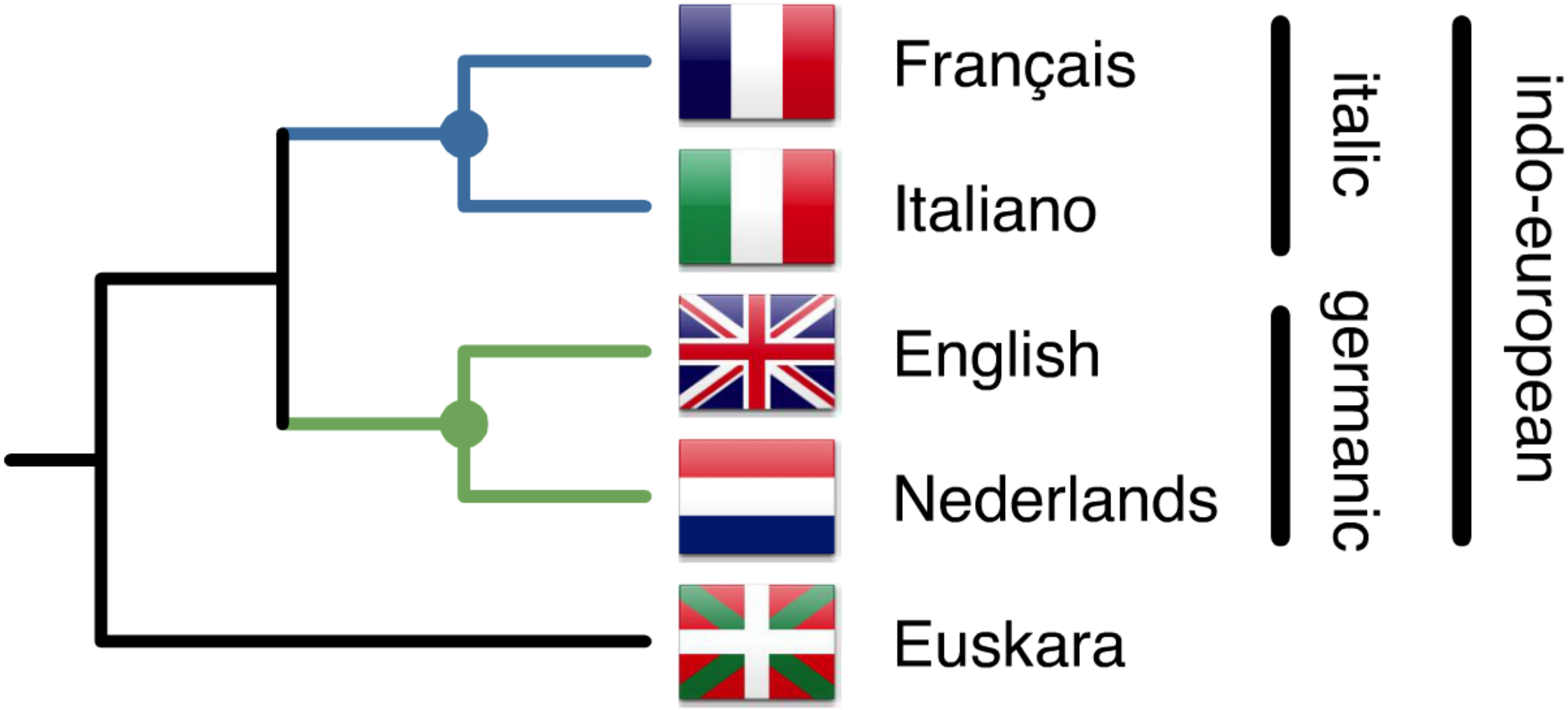
English	NO
Italiano	NO
Français	NON
Nederlands	NEE
Euskara	EZ

Examining only one or two words exposes us to the stochastic error because of the lack of information.

	français	italiano	english	nederlands	euskara
1	un	uno	one	een	bat
2	deux	due	two	twee	bi
3	trois	tre	three	drie	hiru
4	je	io	I	ik	I
5	tu	tu	you	je	duzu
6	qui ?	chi?	who?	wie?	nor?
7	oui	si	yes	ja	bai
8	non	no	no	nee	ez
9	mère	madre	mother	moeder	ama
10	père	padre	father	vader	aita
11	dent	dente	tooth	tand	hortz
12	coeur	cuore	heart	hart	bihotza
13	pied	piede	foot	voet	oinez
14	souris	topolino	mouse	muis	saguaren

Known relationships among European languages

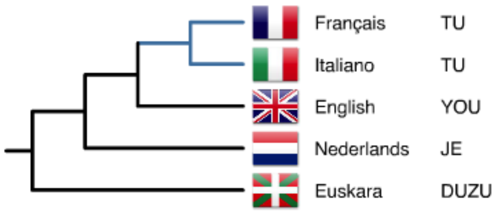
strongly supported by the naive analysis of 14 words



We now have robust statistical support
and phylogenetic resolution!

Potential relationships among European languages

based on the naive analysis of 1 word



The Basque language is our outgroup here.

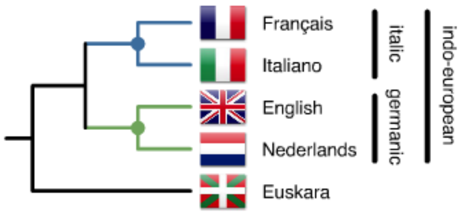
Potential relationships among European languages

based on the naive analysis of 1 word

Examining only one or two words exposes us to the stochastic error because of the lack of information.

Known relationships among European languages

strongly supported by the naive analysis of 14 words



We now have robust statistical support and phylogenetic resolution!

	français	italiano	english	nederlands	euskara
1	un	uno	one	een	bat
2	deux	due	two	twee	bi
3	trois	tre	three	drie	hiru
4	je	io	I	ik	I
5	tu	tu	you	je	duzu
6	qui ?	chi?	who?	wie?	nor?
7	oui	si	yes	ja	bai
8	non	no	no	nee	ez
9	mère	madre	mother	moeder	ama
10	père	padre	father	vader	aita
11	dent	dente	tooth	tand	hortz
12	coeur	cuore	heart	hart	bihotza
13	pied	piede	foot	voet	oinez
14	souris	topolino	mouse	muis	saguaren

Similarly, phylogenomics is phylogenetics applied to many genes at once.

African coelacanth

- blood DNA + muscle RNA library
- assembly with ALLPATHS-LG and Trinity
- annotation with the Ensembl pipeline

genome size: 2.86 Gbp (2.18 Gbp)

- 3 RNA libraries (brain, gonad/kidney, gut/liver)
- assembly with Trinity
- no annotation

West African lungfish

genome size: est. 50-100 Gbp

NGS (Illumina) sequencing was used to sequence the two lobe-finned fishes.

collection of proteomes

- 13 tetrapods (Ensembl)
- 2 ray-finned fishes (Ensembl)
- 2 lobe-finned fishes
  - coelacanth (Broad Institute)
  - lungfish (RNA-seq contigs)

319,459 proteins

after filtration

- tetrapods (Ensembl)
- 2 ray-finned fishes (Ensembl
- 2 lobe-finned fishes
  - coelacanth (Broad Institu
  - lungfish (RNA-seq contig

319,459 proteins
after filtration

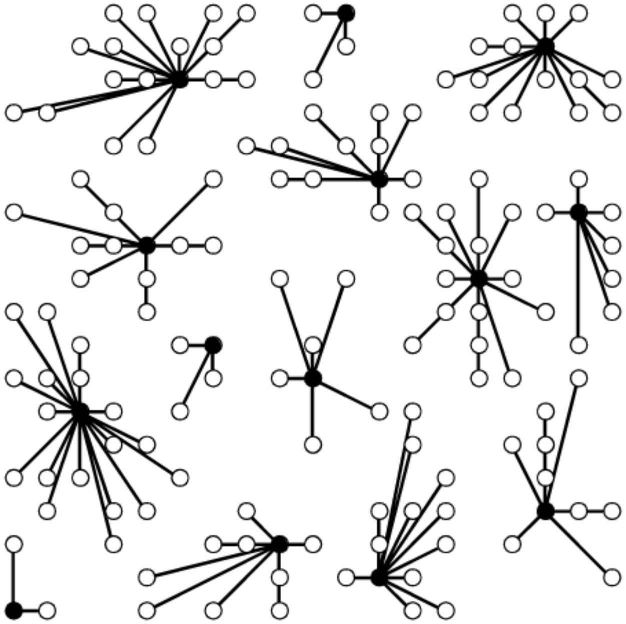
all-vs-all comparison

51,026,866,611 pairwise similarities

overnight on a desktop workstation

100x faster than BLAST using USEARCH

clustering



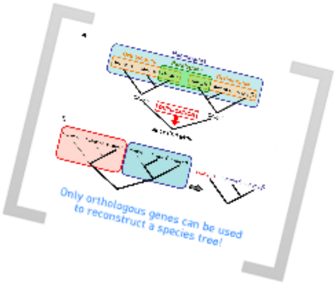
OrthoMCL
(1 hour)

7,764 groups

with both the lungfish
and the coelacanth

addition of transcriptomes

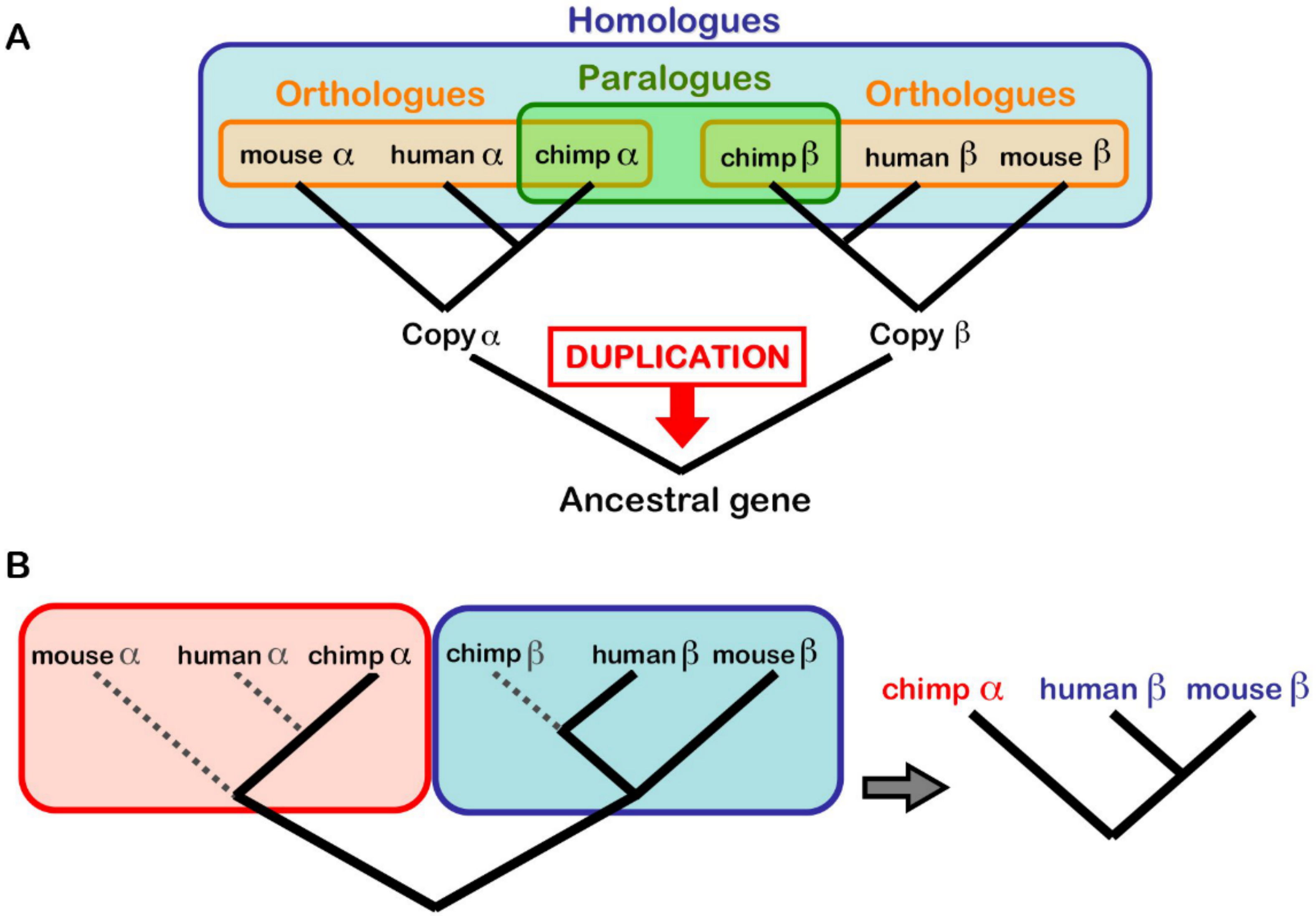
- 1 tetrapod
- 1 ray-finned fish
- 3 cartilaginous fishes



Only orthologous genes can be used to reconstruct a species tree!

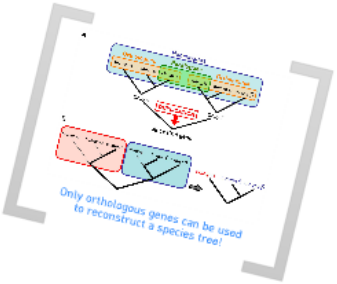
from RNA-seq contigs
using HaMStR (1 week)

373 single-copy alignments



Only orthologous genes can be used to reconstruct a species tree!

addition of transcriptomes



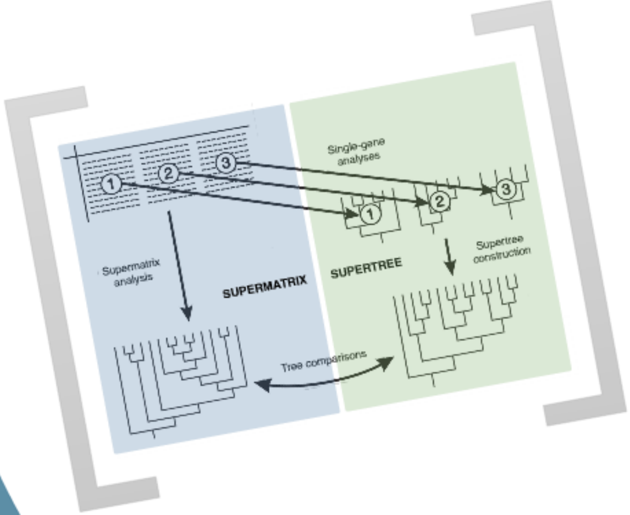
Only orthologous genes can be used to reconstruct a species tree!

- 1 tetrapod
- 1 ray-finned fish
- 3 cartilaginous fishes

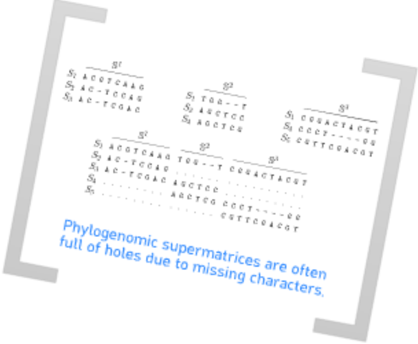
from RNA-seq contigs
using HaMStR (1 week)

373 single-copy alignments

dataset assembly

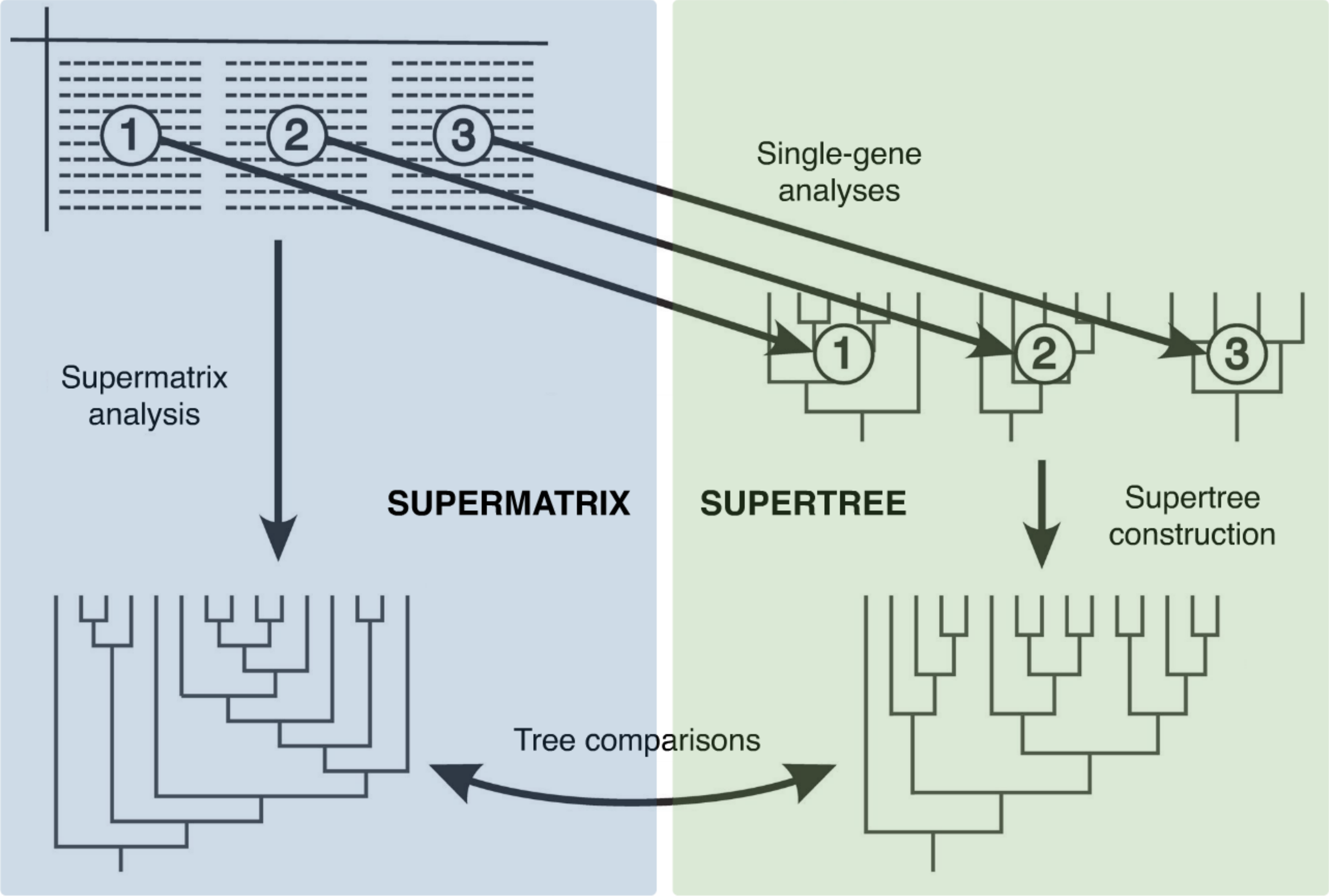


supermatrix made with SCaFoS (1 hour)



Phylogenomic supermatrices are often full of holes due to missing characters.

22 species x 251 genes (100,583 aligned AA)


1
2
3
Single-gene
analyses
Supermatrix
analysis
1
2
3
SUPERMATRIX
SUPERTREE
Supertree
construction
Tree comparisons

	$\mathbb{S}^1$
S_1	A C G T C A A G
S_2	A C - T C C A G
S_3	A C - T C G A C

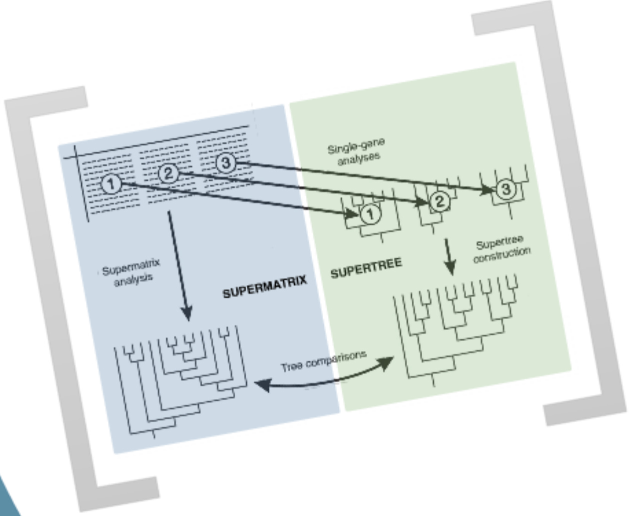
	$\mathbb{S}^2$
S_1	T G G - - T
S_3	A G C T C C
S_4	A G C T C G

	$\mathbb{S}^3$
S_1	C G G A C T A C G T
S_4	C C C T - - - - G G
S_5	C G T T C G A C G T

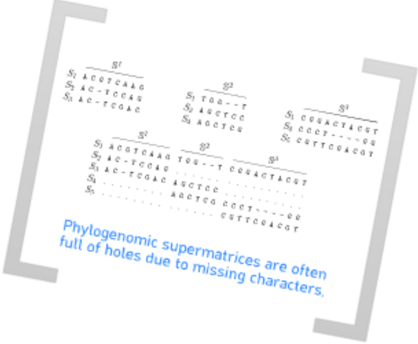
	$\mathbb{S}^1$	$\mathbb{S}^2$	$\mathbb{S}^3$
S_1	A C G T C A A G	T G G - - T	C G G A C T A C G T
S_2	A C - T C C A G	. . . . . .	. . . .
S_3	A C - T C G A C	A G C T C C	. . . .
S_4	. . . .	A G C T C G	C C C T - - - - G G
S_5	. . . .	. . . . . .	C G T T C G A C G T

Phylogenomic supermatrices are often full of holes due to missing characters.

dataset assembly



supermatrix made
with SCaFoS (1 hour)


Phylogenomic supermatrices are often full of holes due to missing characters.

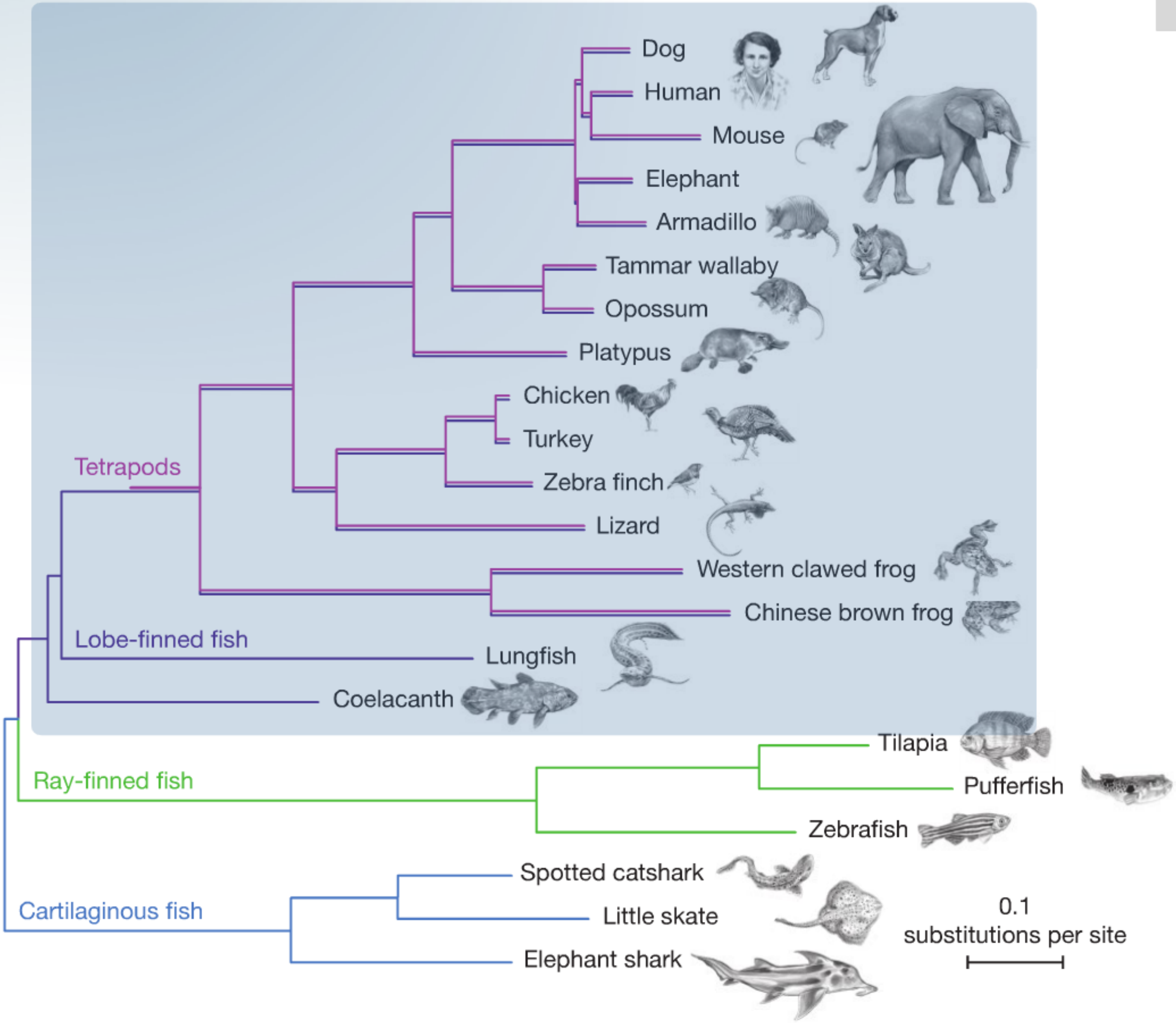
22 species x 251 genes
(100,583 aligned AA)

!

using PhyloBayes
(CATGTR model)

3 months on a
grid computer

**phylogenetic
analysis**



Tetrapods are more closely related to the lungfish than to the coelacanth. Easy, eh?

Our genomic time machine!

collection of proteomes

all-vs-all comparison

51,026,866,611 pairwise similarities

overnight on a desktop workstation

100x faster than BLAST using USEARCH

clustering

OrthoMCL (1 hour)

7,764 groups

with both the lungfish and the coelacanth

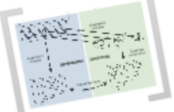
addition of transcriptomes

- 1 tetrapod
- 1 ray-finned fish
- 3 cartilaginous fishes

from RNA-seq contigs using HaMStR (1 week)

373 single-copy alignments

dataset assembly



supermatrix made with SCaFoS (1 hour)

22 species x 251 genes (100,583 aligned AA)

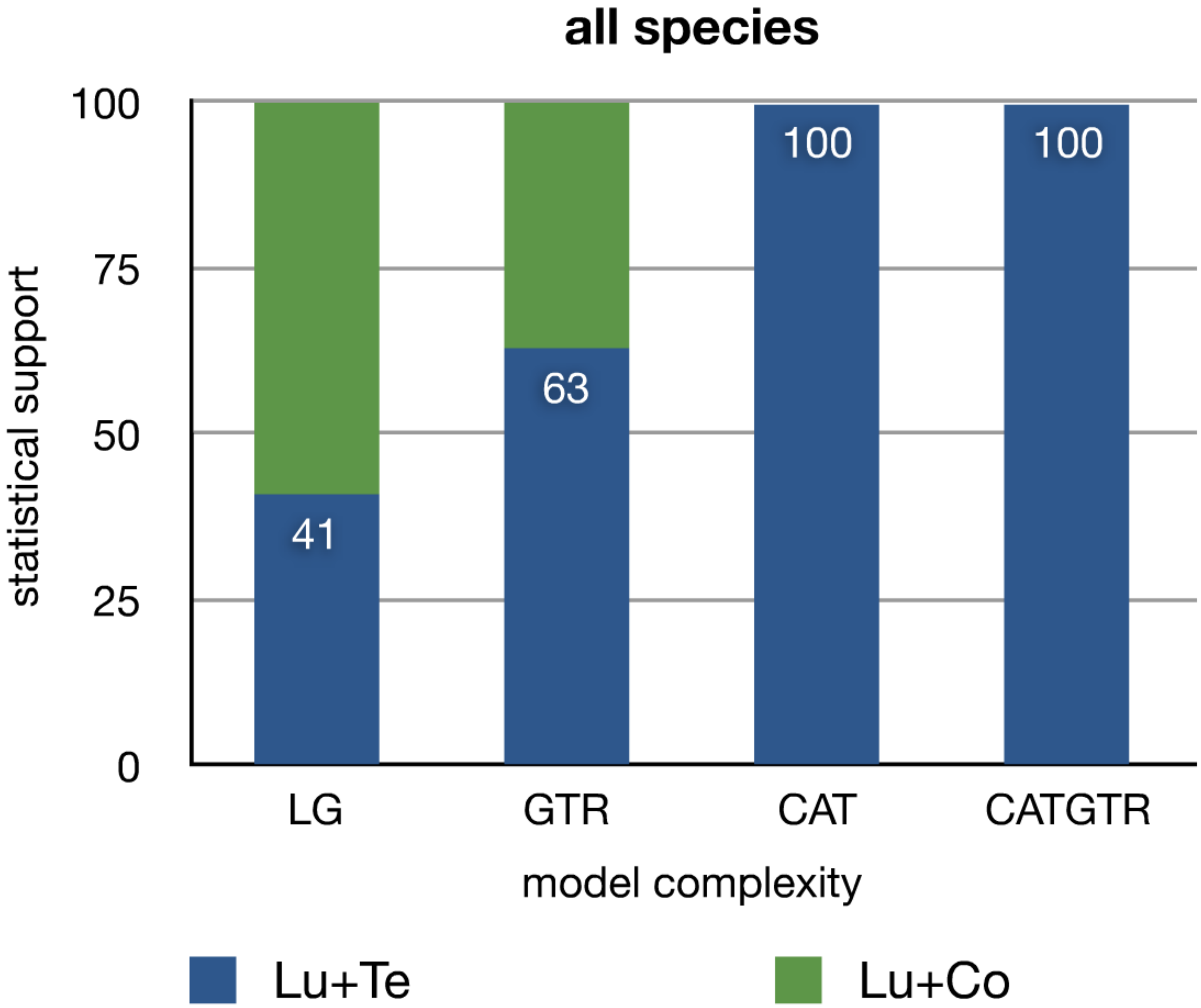
using PhyloBayes (CATGTR model)

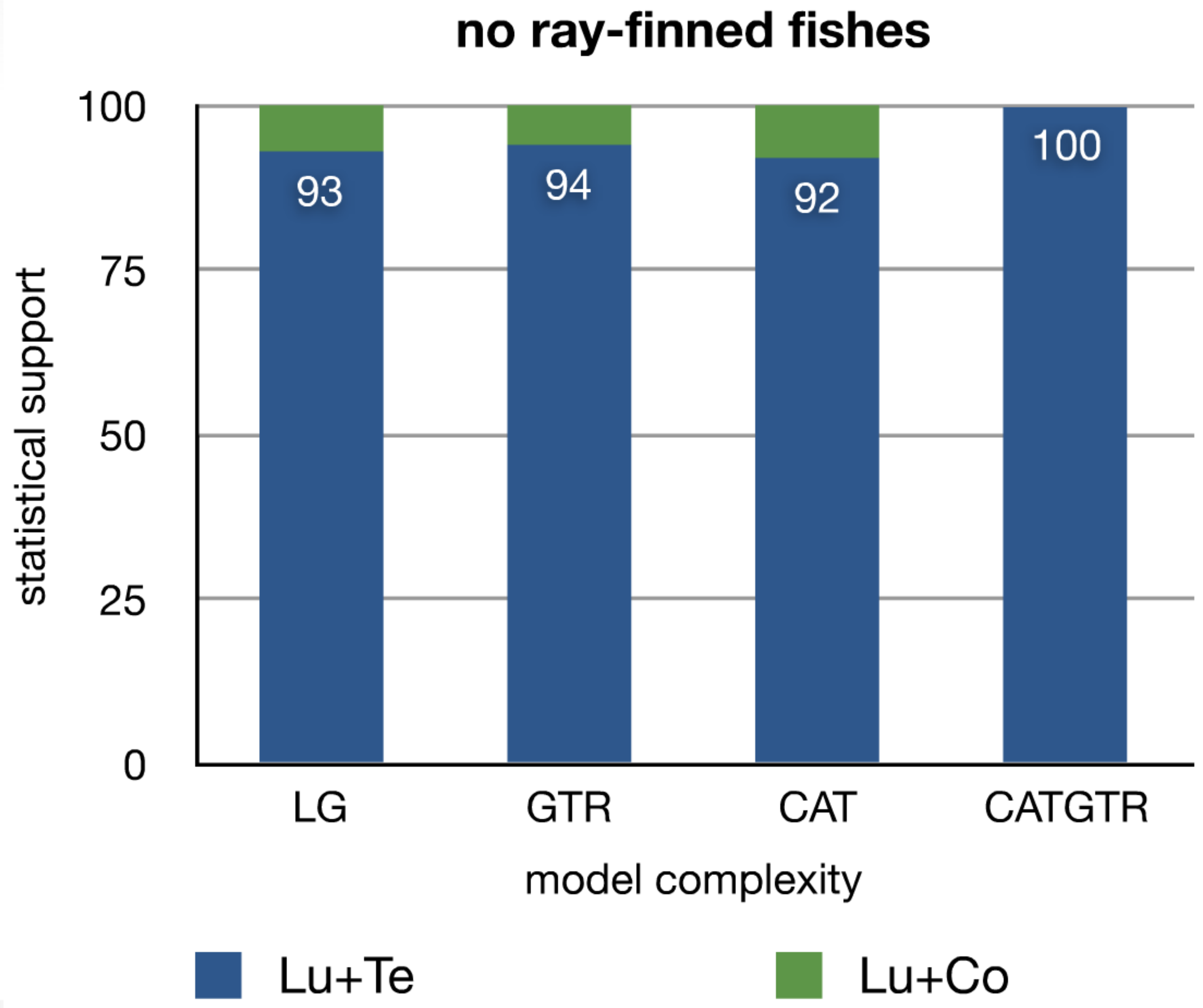
3 months on a grid computer

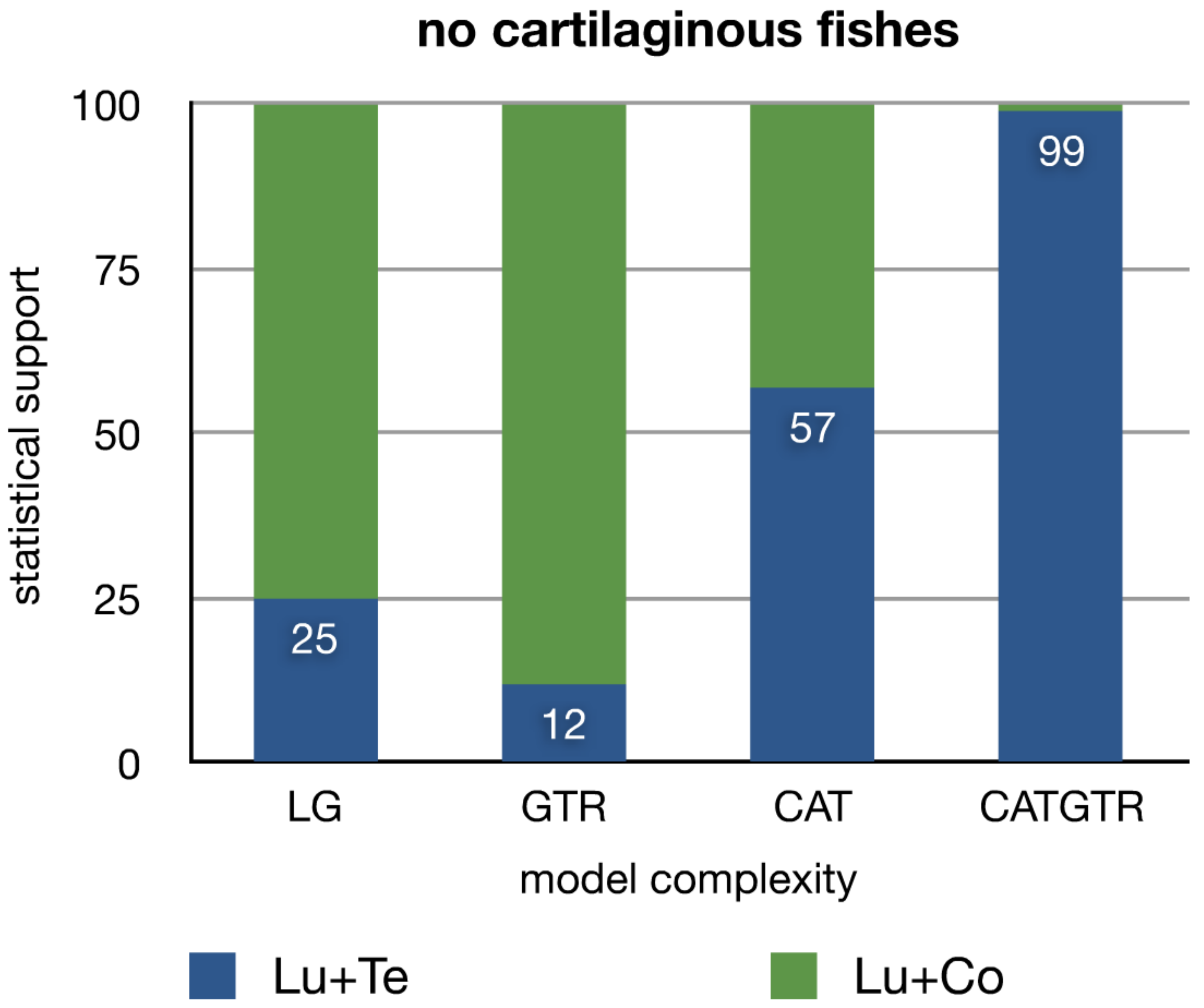
phylogenetic analysis

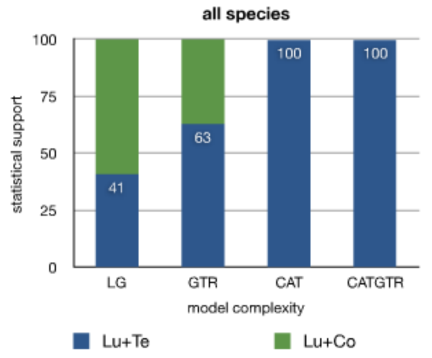
Tetrapods are more closely related to the lungfish than to the coelacanth. Easy, eh?

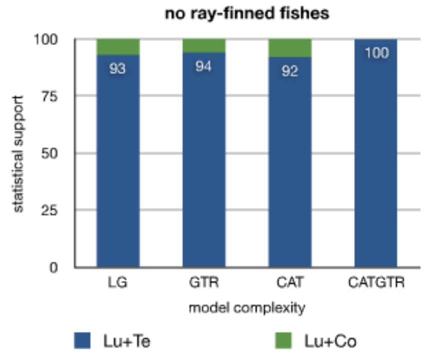
phylogenomic protocol

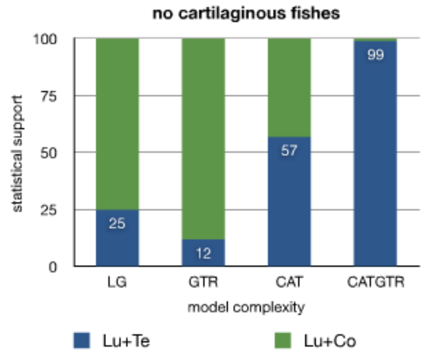

all species
statistical support
100
75
50
25
0
41
63
100
100
LG
GTR
CAT
CATGTR
model complexity
Lu+Te
Lu+Co

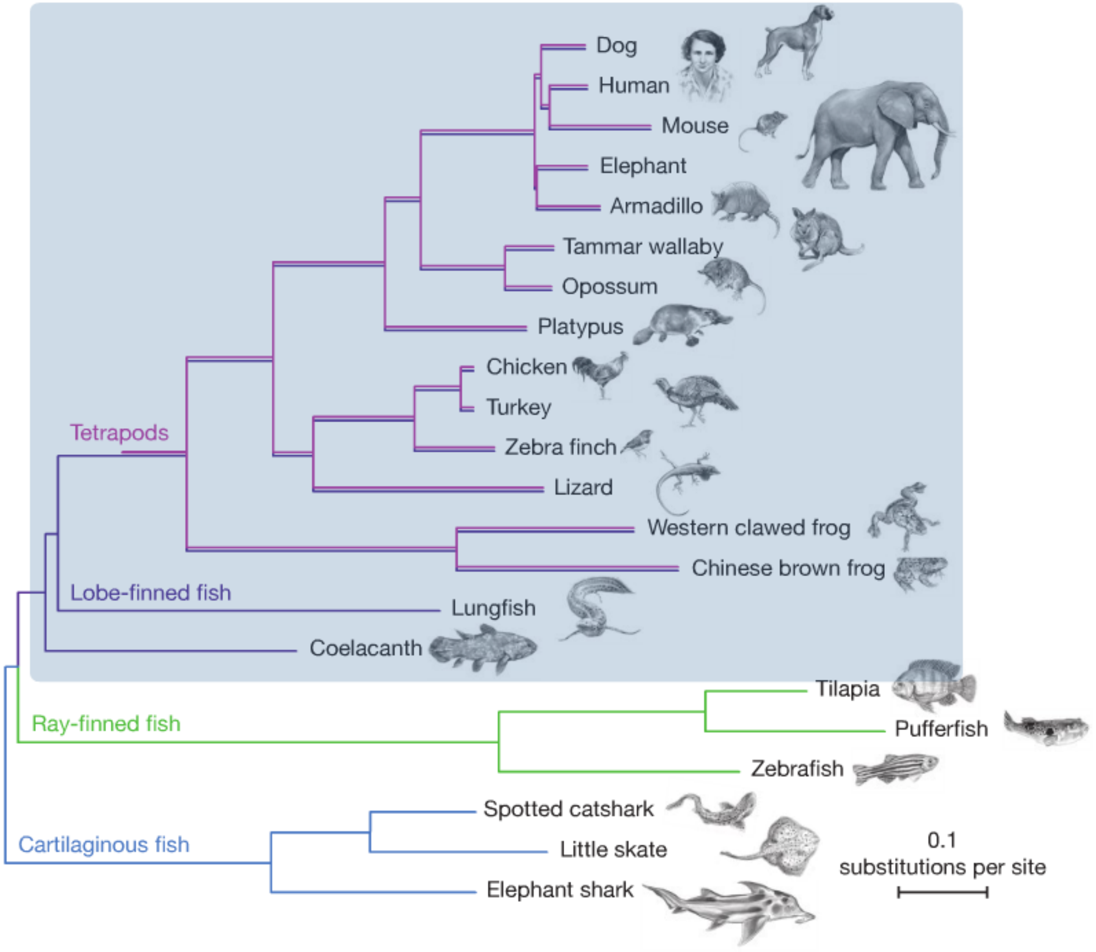

no ray-finned fishes
statistical support
100
75
50
25
0
93
94
92
100
LG
GTR
CAT
CATGTR
model complexity
Lu+Te
Lu+Co


no cartilaginous fishes
statistical support
100
75
50
25
0
25
12
57
99
LG
GTR
CAT
CATGTR
model complexity
🟦 Lu+Te
🟩 Lu+Co





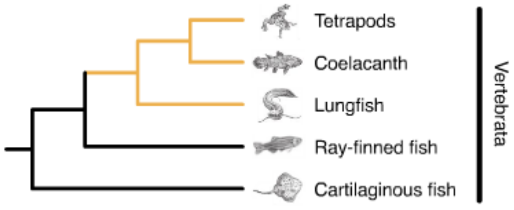




Why do we have such an unstable phylogeny?

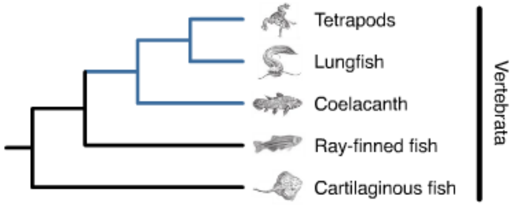
The coelacanth as the closest relative of tetrapods

Gorr et al. (1991) – haemoglobins a & b (nucl.)



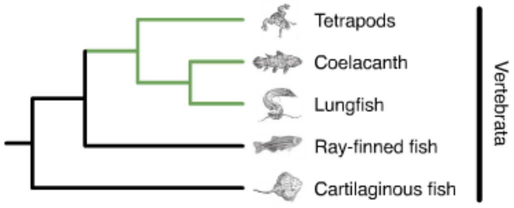
The lungfish as the closest relative of tetrapods

e.g., Meyer & Dolven (1992) – 12S rRNA & cob (mitochondr.)
Brinkmann et al. (2004) – RAS1 & 2 (nucl.)

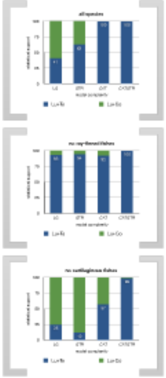


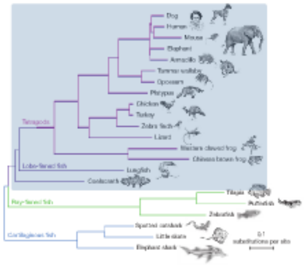
Both the coelacanth and the lungfish!

e.g., Yokobori et al. (1994) – cox1 (mitochondr.)
Zardoya & Meyer (1996) – 28S rRNA (nucl.)



Similarly, phylogenomics is phylogenetics applied to many genes at once.





Why do we have such an unstable phylogeny?

Phylogenomics is useful but can suffer from artifacts!

Let's look at their causes!

African coelacanth

- blood DNA + muscle RNA library
- assembly with ALLPATHS-LG and Trinity
- annotation with the Ensembl pipeline

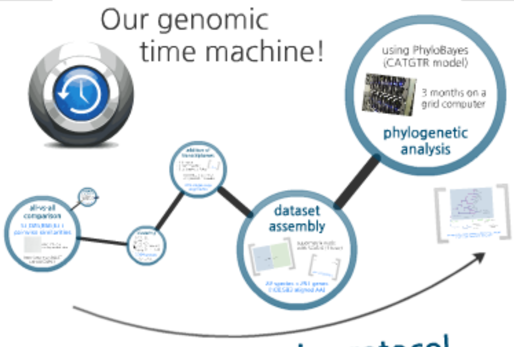
genome size: 2.86 Gbp (2.18 Gbp)

- 3 RNA libraries (brain, gonad/kidney, gut/liver)
- assembly with Trinity
- no annotation

West African lungfish

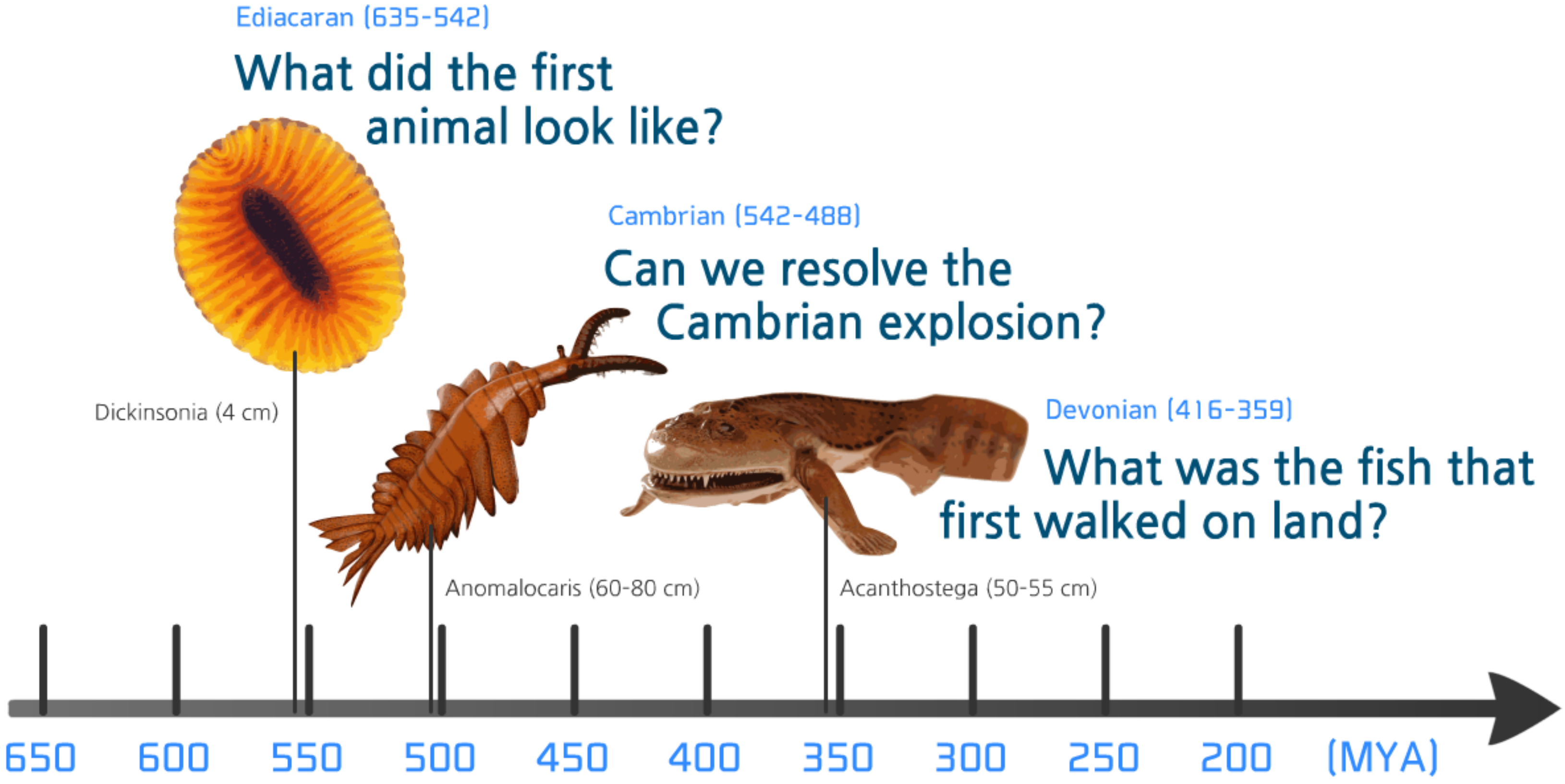
genome size: est. 50-100 Gbp

NGS (Illumina) sequencing was used to sequence the two lobe-finned fishes.



phylogenomic protocol

Computers as Genomic Time Machines for Meeting our Ancestors



Ediacaran (635-542)

What did the first animal look like?

Cambrian (542-488)

Can we resolve the Cambrian explosion?

Devonian (416-359)

What was the fish that first walked on land?

Prof. Denis Baurain — ULg
Bioforum — April 18th, 2013

irst
ook like?

Cambrian (542-488)

Can we resolve the Cambrian explosion?

Devonian

What
first walke

Ediacaran (635-542 MYA)

Only odd soft animals live the sea.

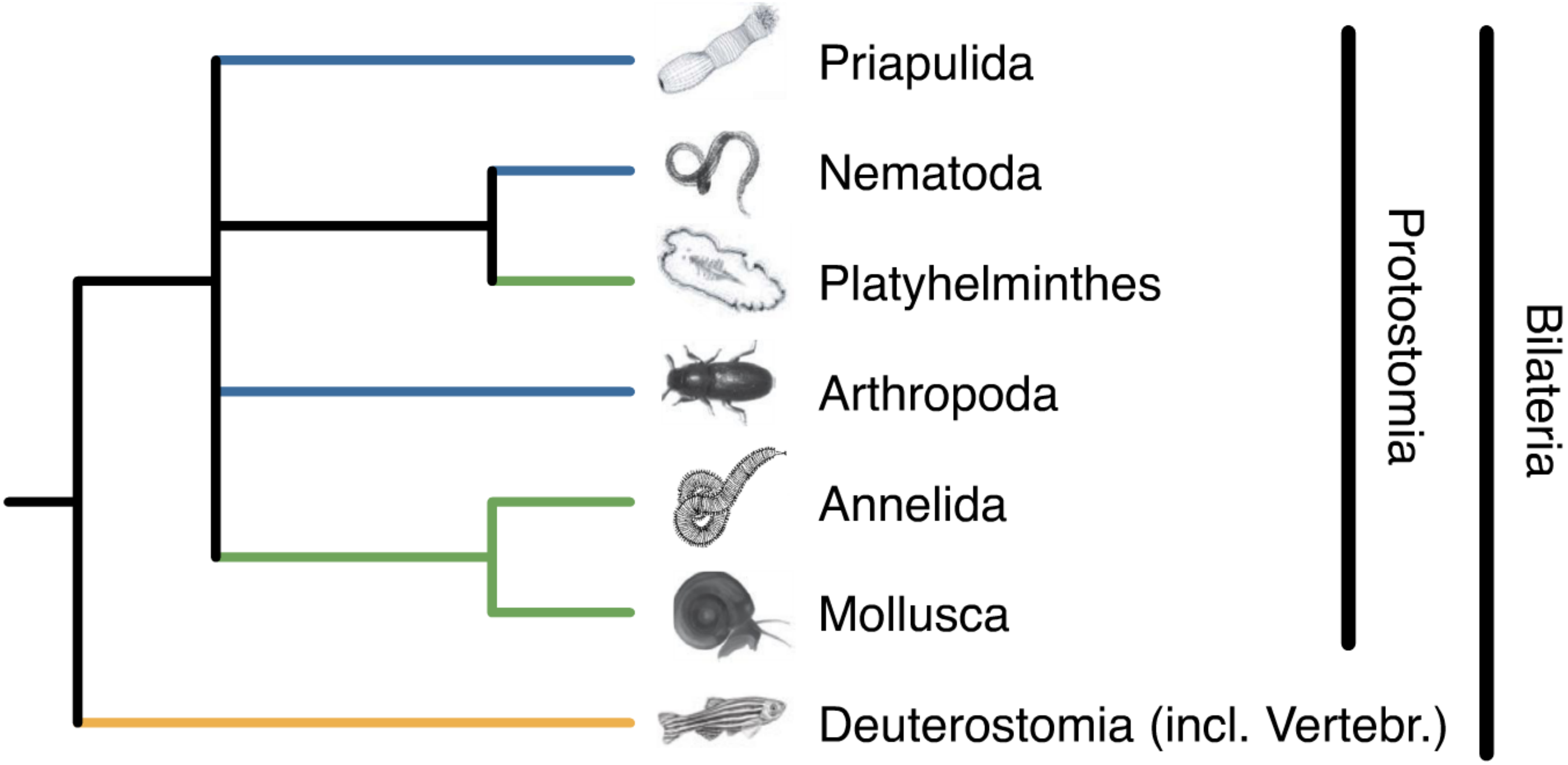
Cambrian (542-488 MYA)

Many animals now thrive in the sea!

Looking back at the past, it seems like
all bilaterian lineages have appeared at once!

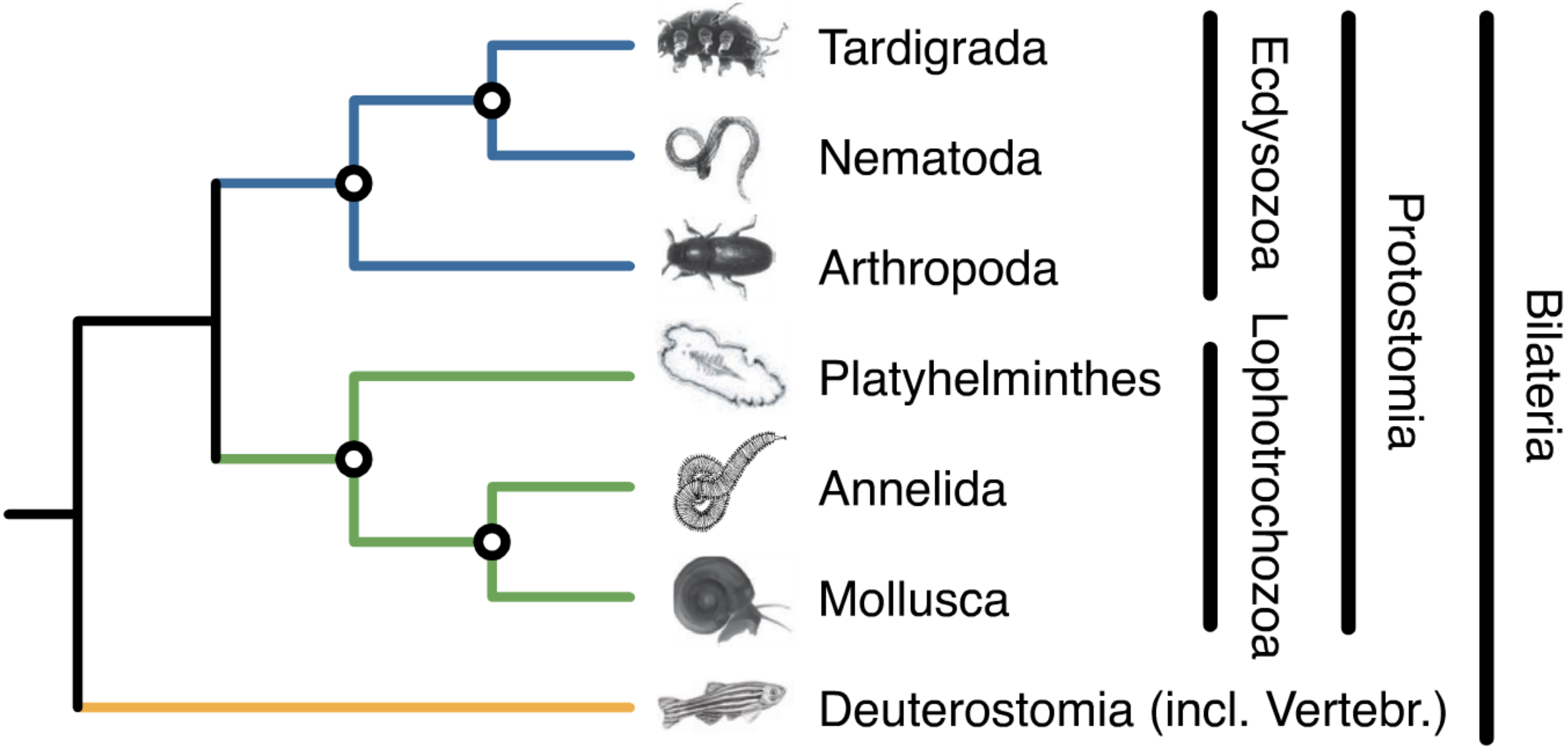
The « Cambrian explosion » — nothing to see!

Rokas et al. (2005) – 50 nuclear genes



The « new animal phylogeny »

Aguinaldo et al. (1997) – 18S rRNA
Philippe et al. (2005) – 71 nuclear genes

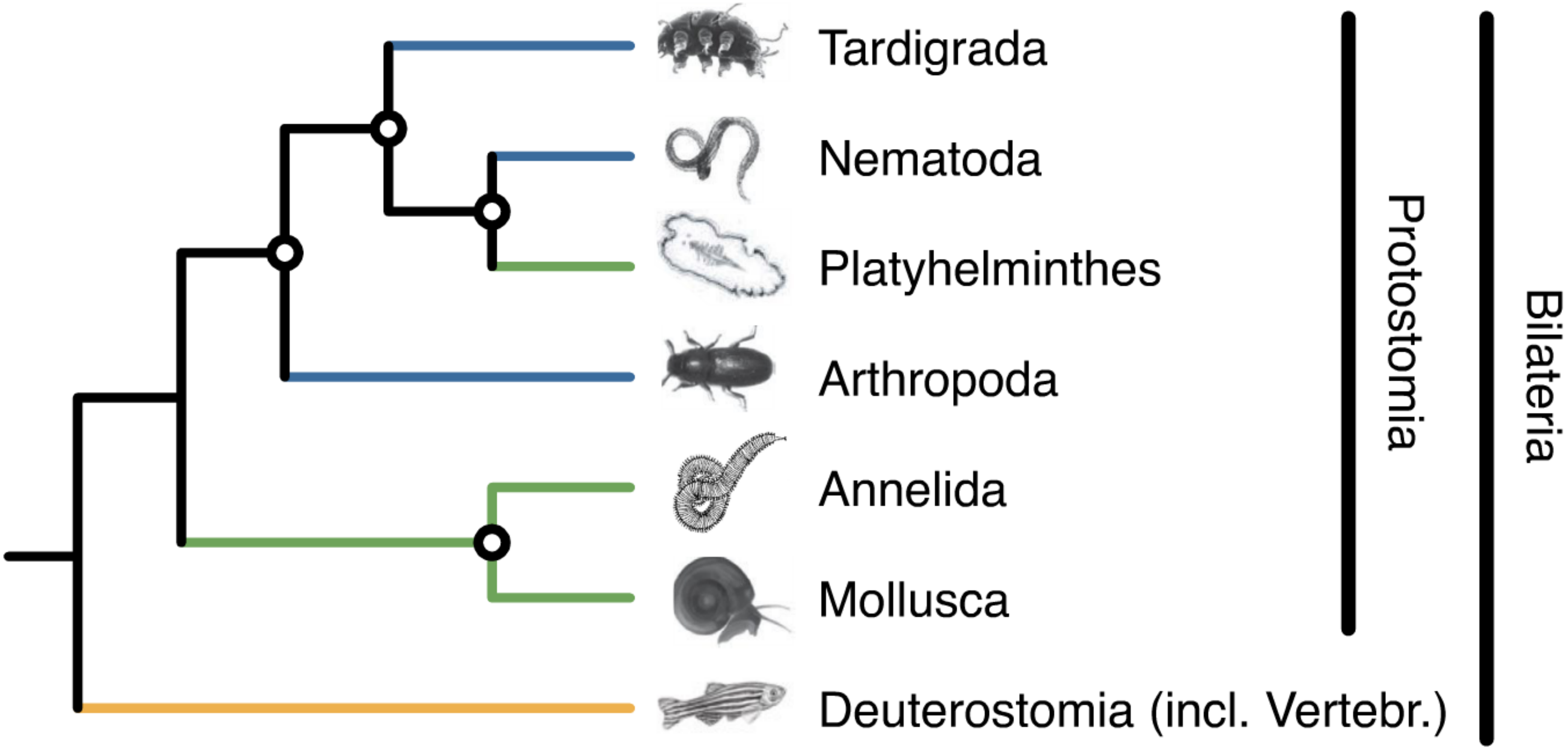


Tardigrada
Nematoda
Arthropoda
Platyhelminthes
Annelida
Mollusca
Deuterostomia (incl. Vertebr.)

Ecdysozoa
Lophotrochozoa
Protostomia
Bilateria

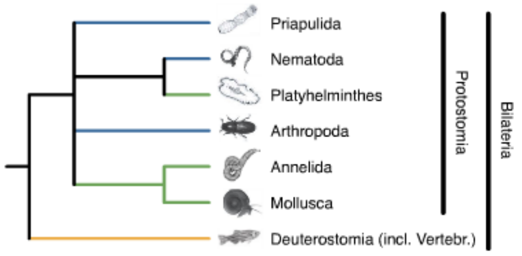
Example of alternative phylogeny

Philippe et al. (2005) – 146 nuclear genes



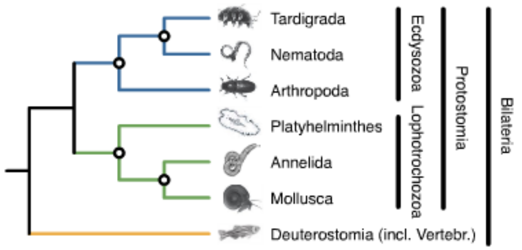
The « Cambrian explosion » — nothing to see!

Rokas et al. (2005) – 50 nuclear genes



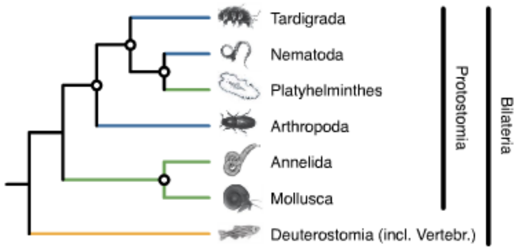
The « new animal phylogeny »

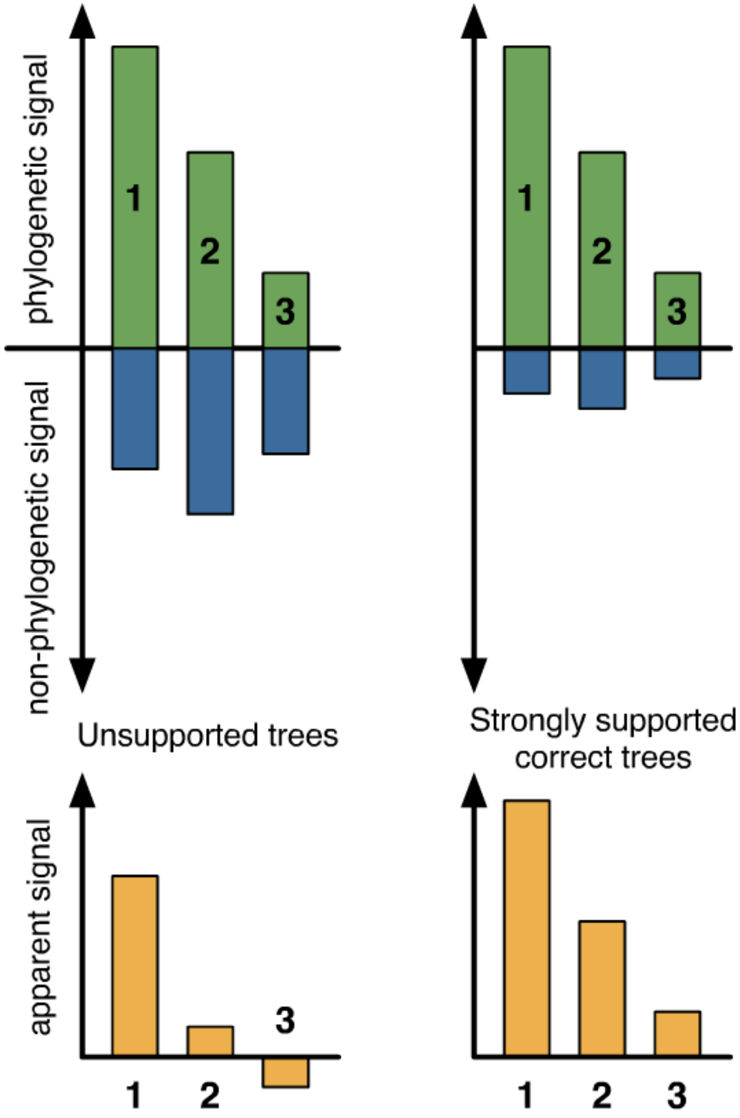
Aguinaldo et al. (1997) – 18S rRNA
Philippe et al. (2005) – 71 nuclear genes



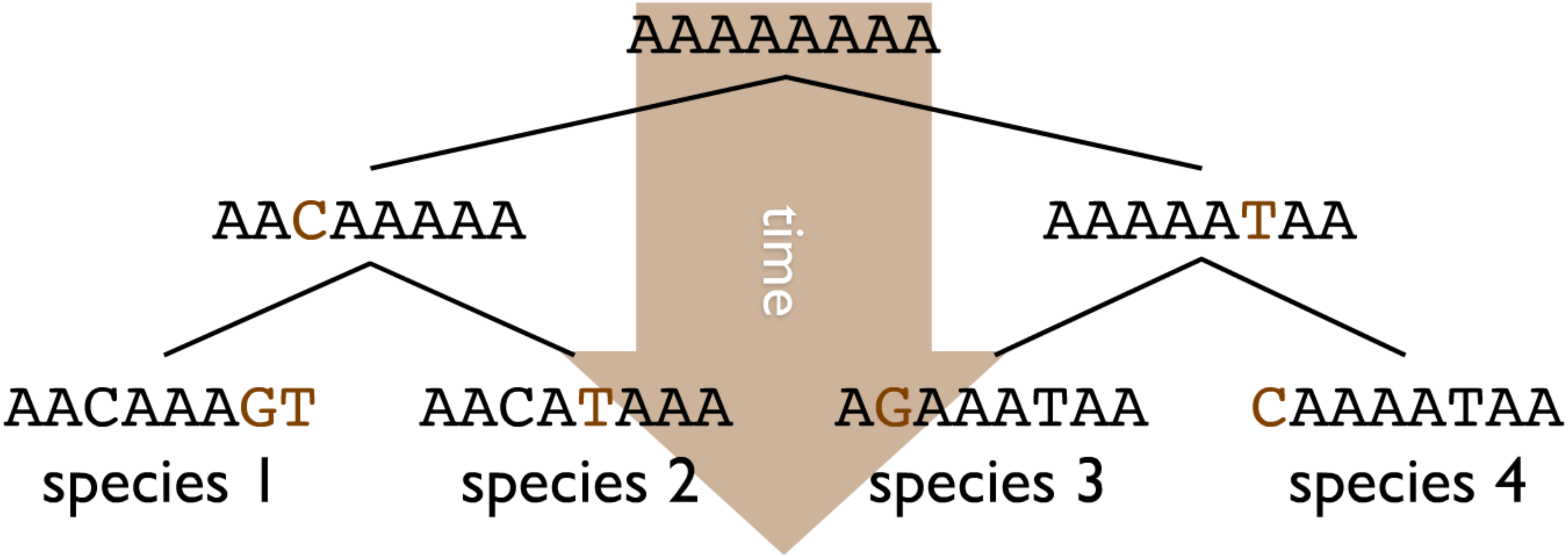
Example of alternative phylogeny

Philippe et al. (2005) – 146 nuclear genes





Working hypothesis: the lack of resolution stems from an excess of non-phylogenetic signal.



AAAAAAAA

AACAAAAA
AAAAATAA

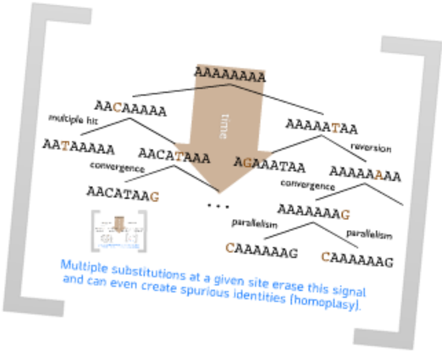
AACAAAGT
species 1

AACATAAA
species 2

AGAAATAA
species 3

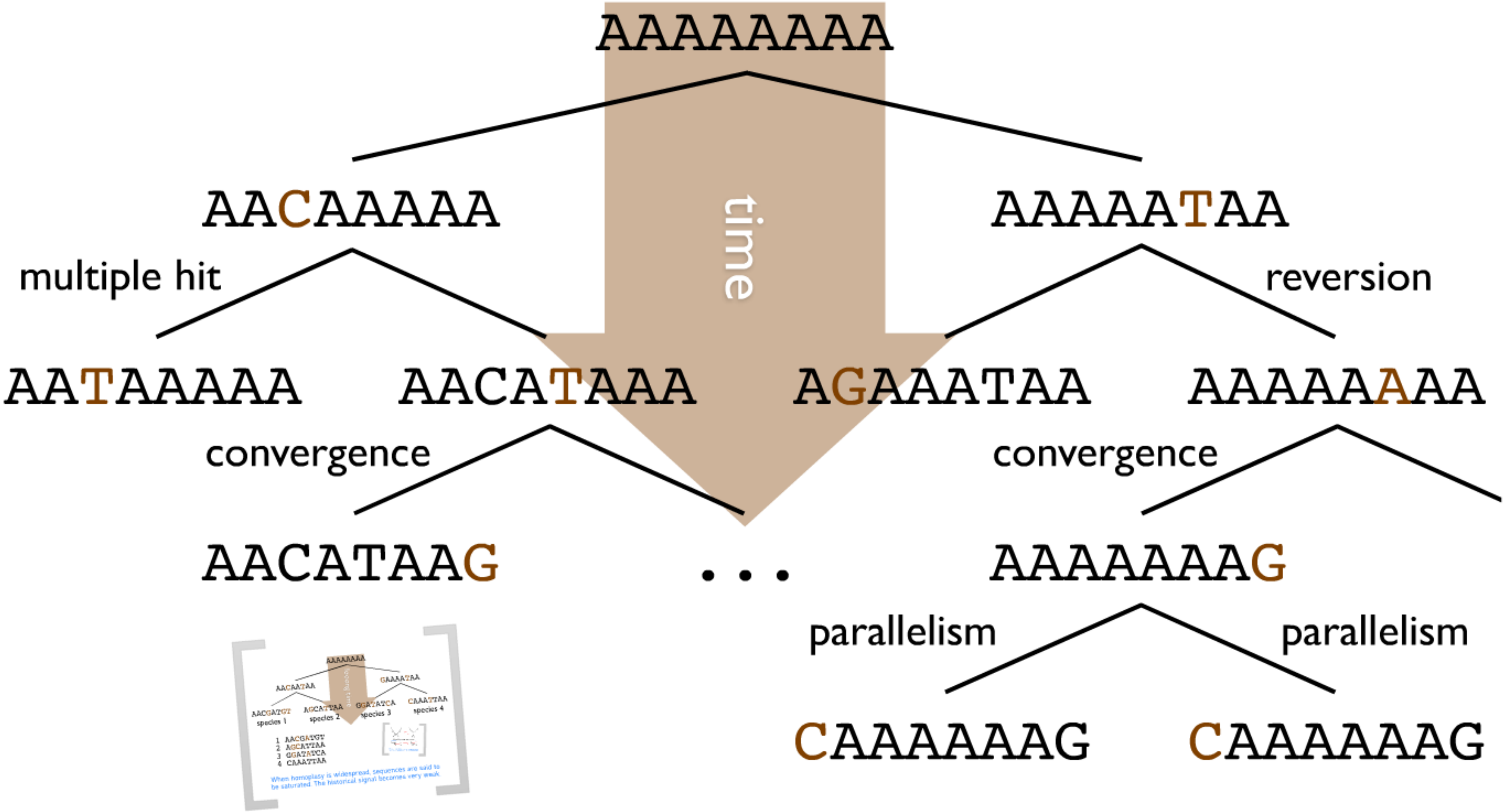
CAAAATAA
species 4

1 AACAAAGT
2 AACATAAA
3 AGAAATAA
4 CAAAATAA



Multiple substitutions at a given site erase this signal and can even create spurious identities (homoplasy).

The phylogenetic signal lies in the substitutions inherited from the common ancestors of the sequences.



AAAAAAAA

time

AACAAAAA

AAAAATAA

multiple hit

reversion

AATAAAAA

AACATAAA

AGAAATAA

AAAAAAAA

convergence

convergence

AACATAAG

. . .

AAAAAAAG

parallelism

parallelism

CAAAAAAG

CAAAAAAG

AAAAAAAA

AACAATAA

GAAAATAA

AACGATGT

species 1

AGCATTAA

species 2

GGATATCA

species 3

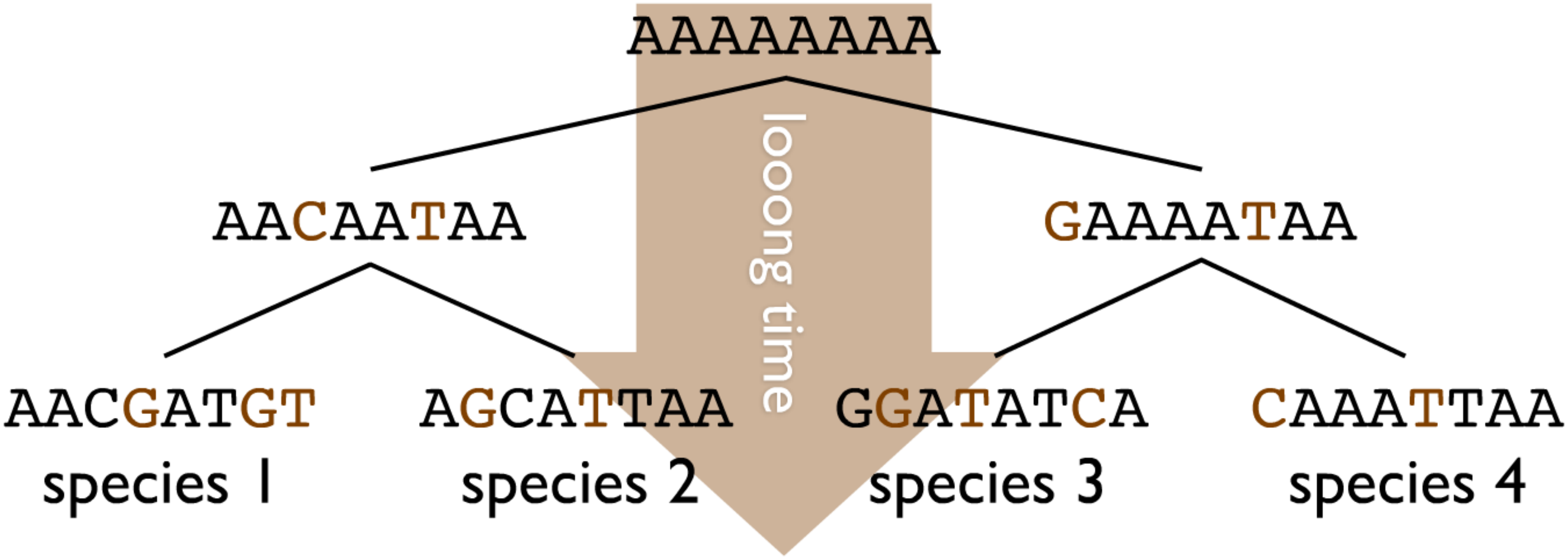
CAAATTAA

species 4

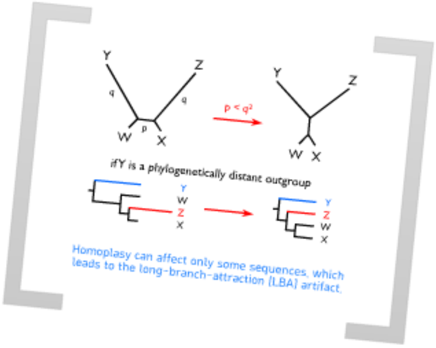
1 AACGATGT
2 AGCATTAA
3 GGATATCA
4 CAAATTAA

When homoplasy is widespread, sequences are said to be saturated. The historical signal becomes very weak.

Multiple substitutions at a given site erase this signal and can even create spurious identities (homoplasy).



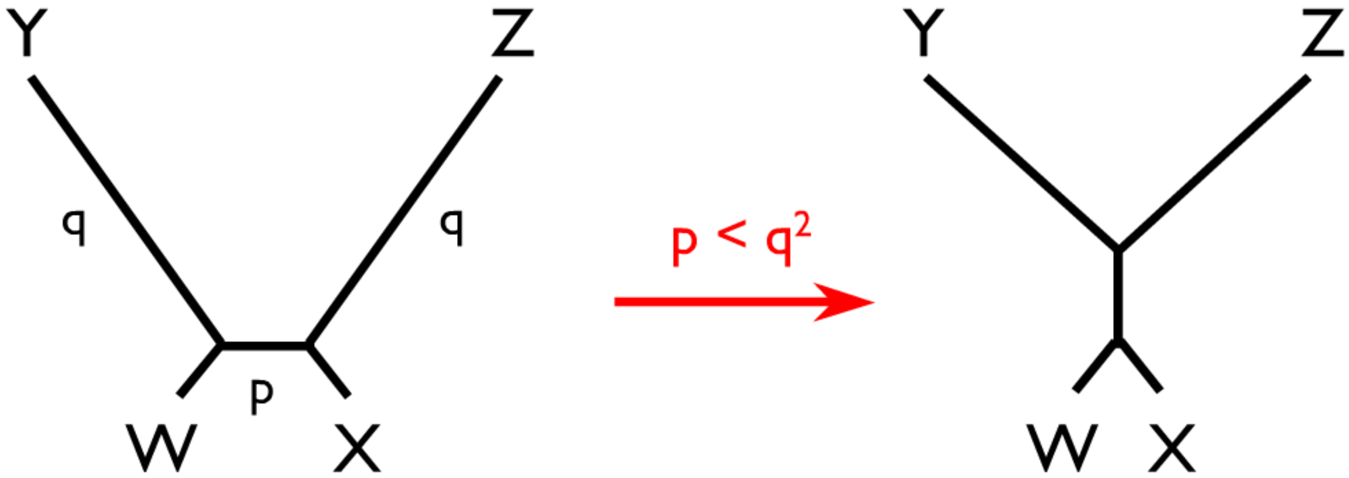
1 AACGATGT
2 AGCATTAA
3 GGATATCA
4 CAAATTAA



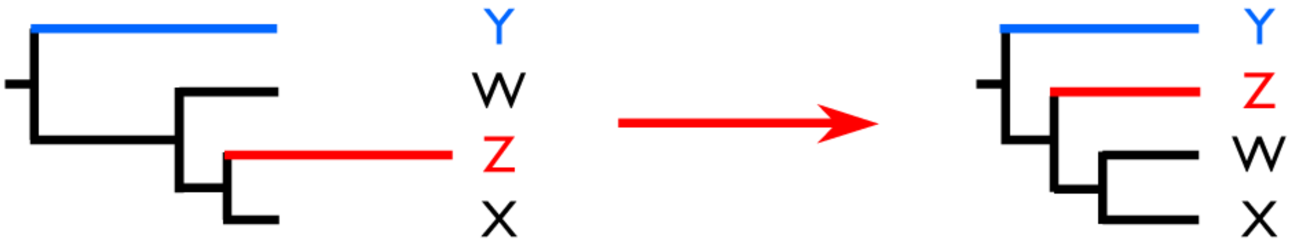
ifY is a phylogenetically distant outgroup

Homoplasy can affect only some sequences, which leads to the long-branch-attraction (LBA) artifact.

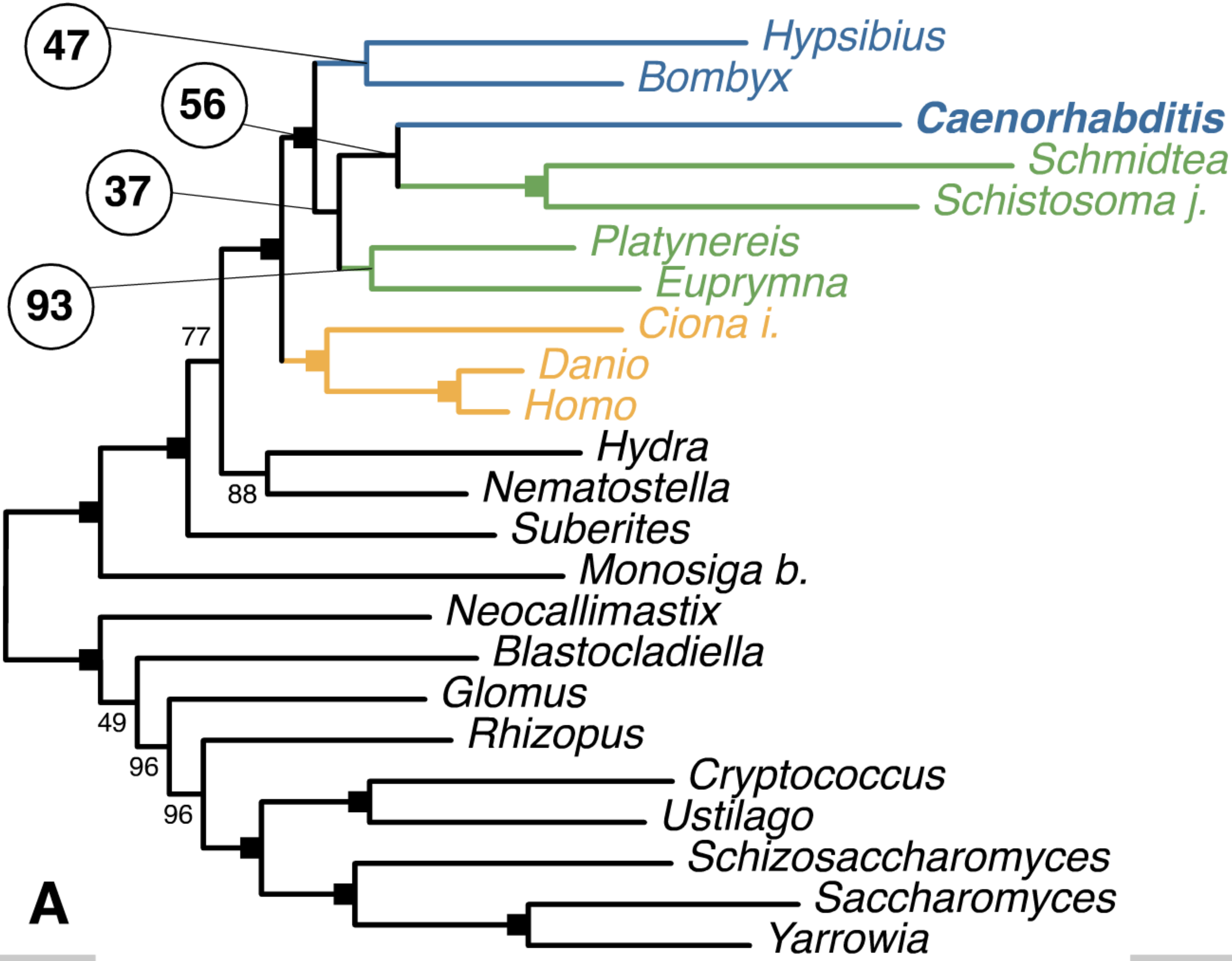
When homoplasy is widespread, sequences are said to be saturated. The historical signal becomes very weak.

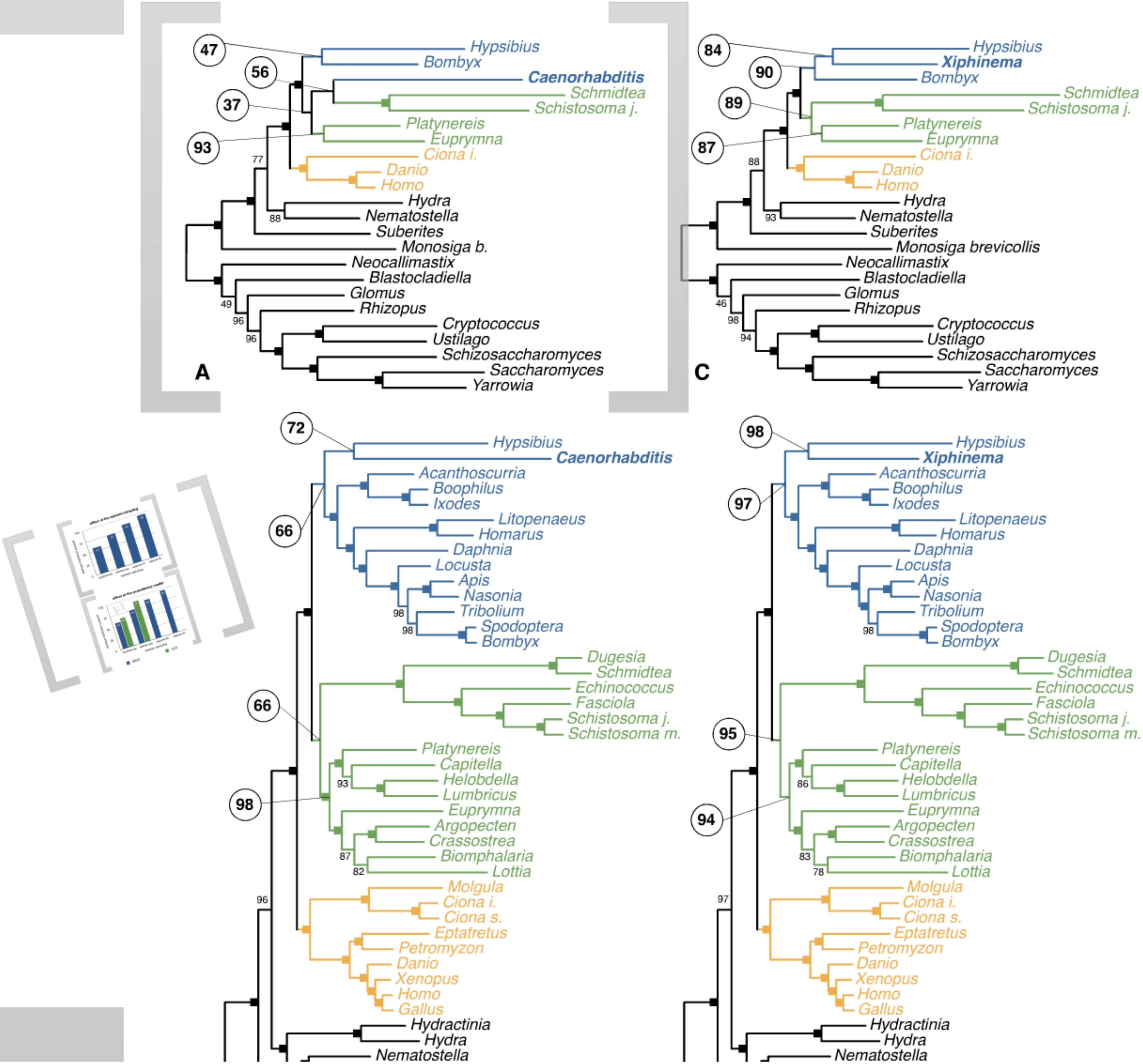


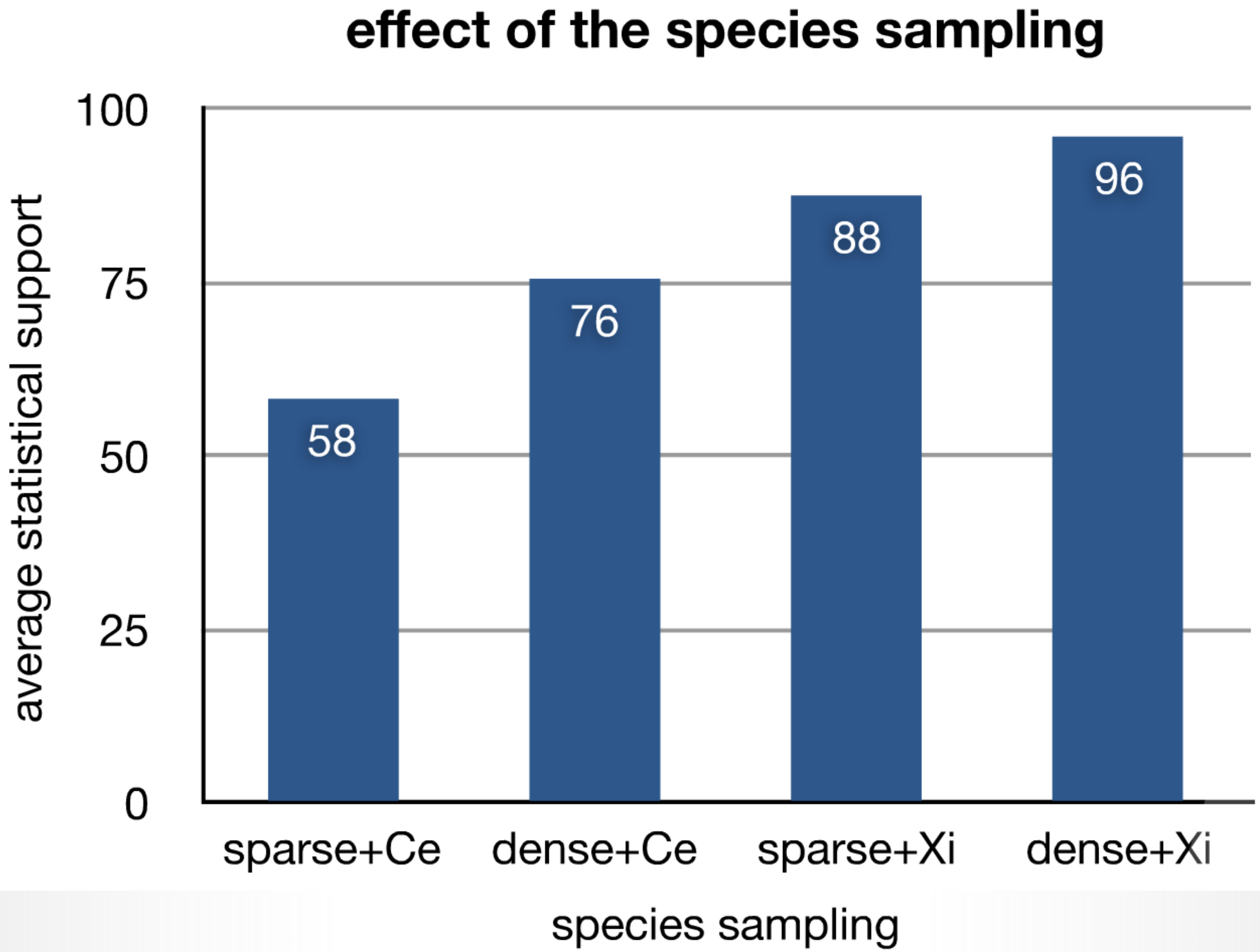
if Y is a phylogenetically distant outgroup



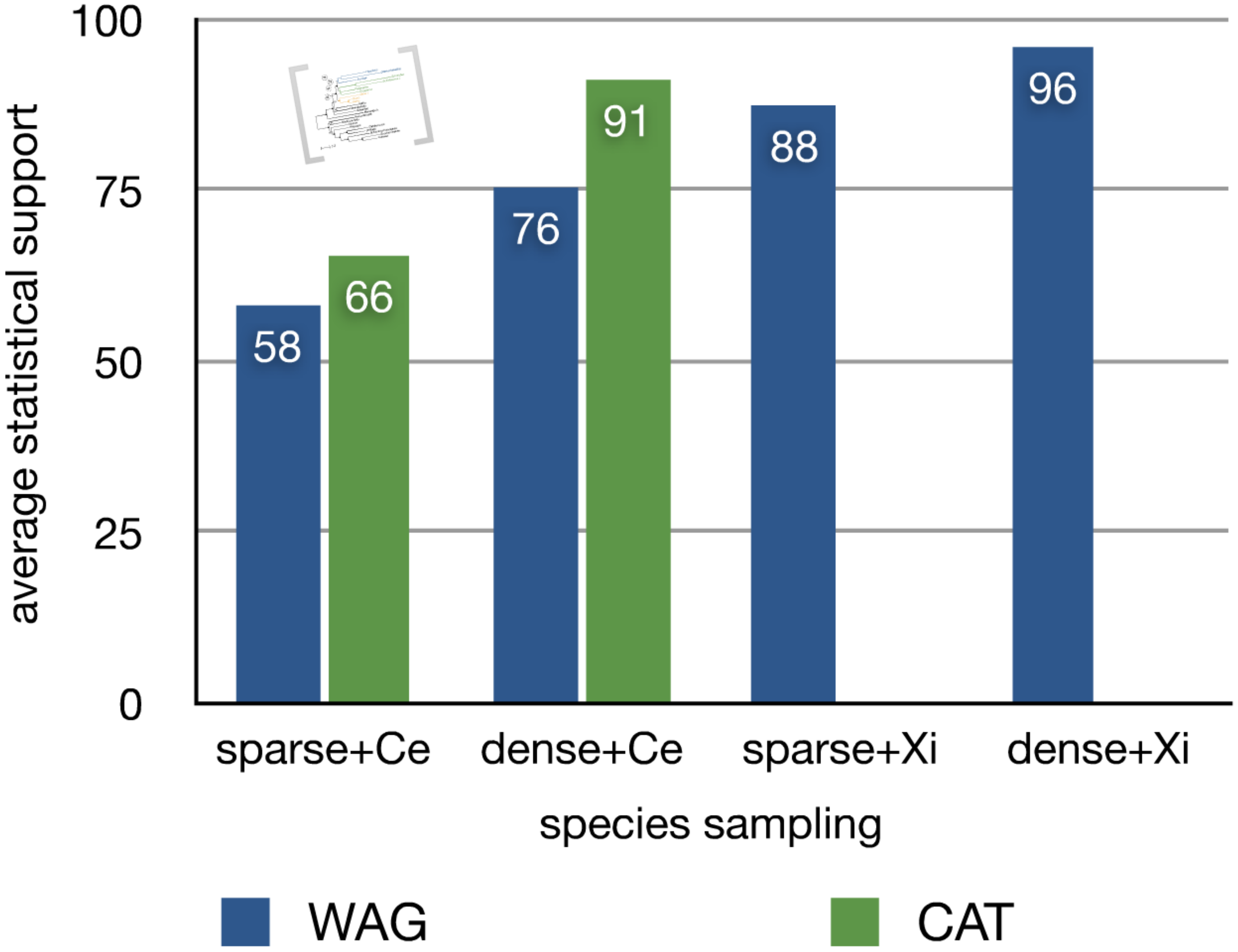
Homoplasy can affect only some sequences, which leads to the long-branch-attraction (LBA) artifact.

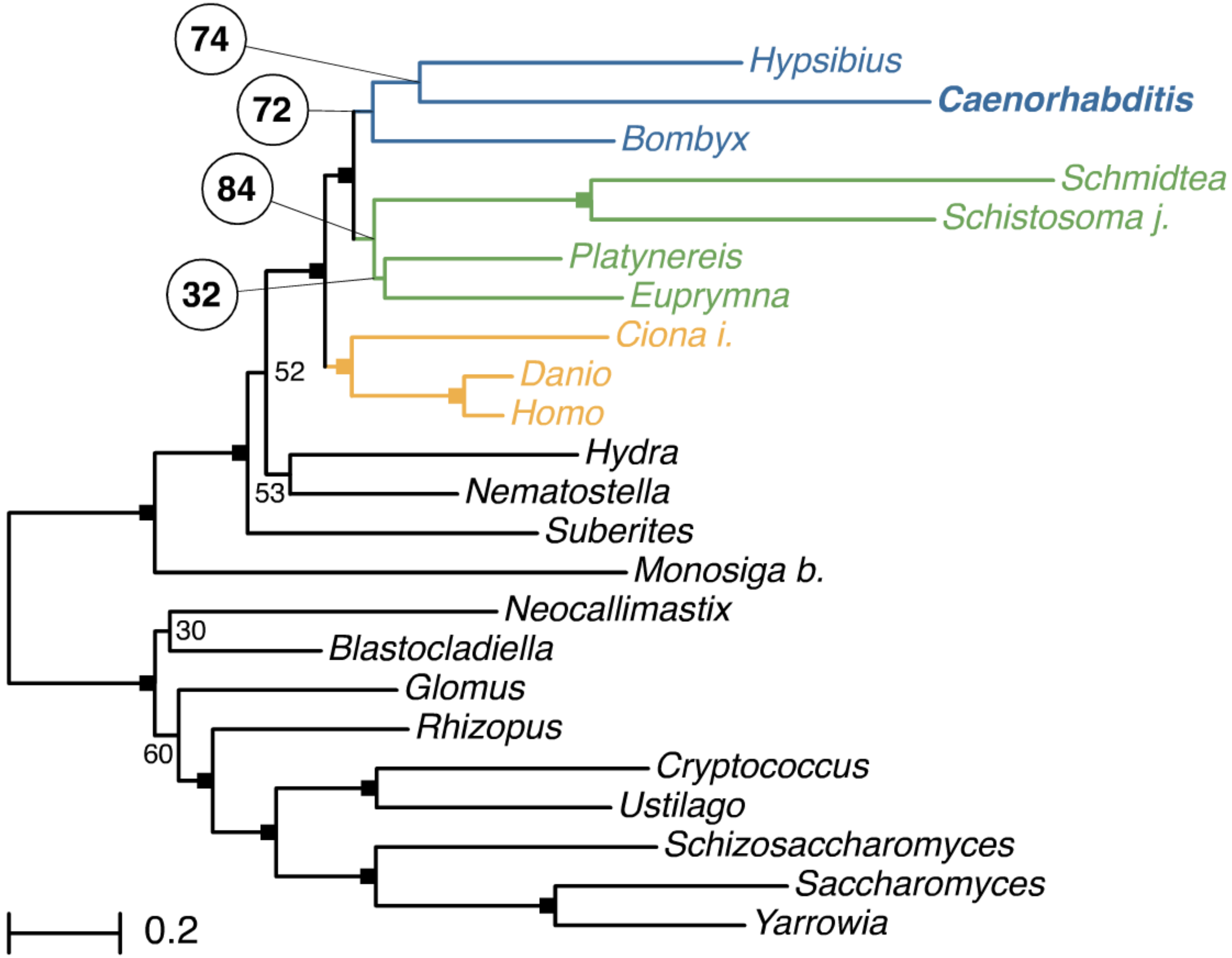

47
56
37
93
77
88
49
96
96
Hypsibius
Bombyx
Caenorhabditis
Schmidtea
Schistosoma j.
Platynereis
Euprymna
Ciona i.
Danio
Homo
Hydra
Nematostella
Suberites
Monosiga b.
Neocallimastix
Blastocladiella
Glomus
Rhizopus
Cryptococcus
Ustilago
Schizosaccharomyces
Saccharomyces
Yarrowia
A


A
47
56
37
93
77
88
49
96
96
Hypsibius
Bombyx
Caenorhabditis
Schmidtea
Schistosoma j.
Platynereis
Euprymna
Ciona i.
Danio
Homo
Hydra
Nematostella
Suberites
Monosiga b.
Neocallimastix
Blastocladiella
Glomus
Rhizopus
Cryptococcus
Ustilago
Schizosaccharomyces
Saccharomyces
Yarrowia
C
84
90
89
87
88
93
46
98
94
Hypsibius
Xiphinema
Bombyx
Schmidtea
Schistosoma j.
Platynereis
Euprymna
Ciona i.
Danio
Homo
Hydra
Nematostella
Suberites
Monosiga brevicollis
Neocallimastix
Blastocladiella
Glomus
Rhizopus
Cryptococcus
Ustilago
Schizosaccharomyces
Saccharomyces
Yarrowia
72
66
66
98
98
98
93
87
82
96
Hypsibius
Caenorhabditis
Acanthoscurria
Boophilus
Ixodes
Litopenaeus
Homarus
Daphnia
Locusta
Apis
Nasonia
Tribolium
Spodoptera
Bombyx
Dugesia
Schmidtea
Echinococcus
Fasciola
Schistosoma j.
Schistosoma m.
Platynereis
Capitella
Helobdella
Lumbricus
Euprymna
Argopecten
Crassostrea
Biomphalaria
Lottia
Molgula
Ciona i.
Ciona s.
Eptatretus
Petromyzon
Danio
Xenopus
Homo
Gallus
Hydractinia
Hydra
Nematostella
98
97
95
94
98
86
83
78
97
Hypsibius
Xiphinema
Acanthoscurria
Boophilus
Ixodes
Litopenaeus
Homarus
Daphnia
Locusta
Apis
Nasonia
Tribolium
Spodoptera
Bombyx
Dugesia
Schmidtea
Echinococcus
Fasciola
Schistosoma j.
Schistosoma m.
Platynereis
Capitella
Helobdella
Lumbricus
Euprymna
Argopecten
Crassostrea
Biomphalaria
Lottia
Molgula
Ciona i.
Ciona s.
Eptatretus
Petromyzon
Danio
Xenopus
Homo
Gallus
Hydractinia
Hydra
Nematostella

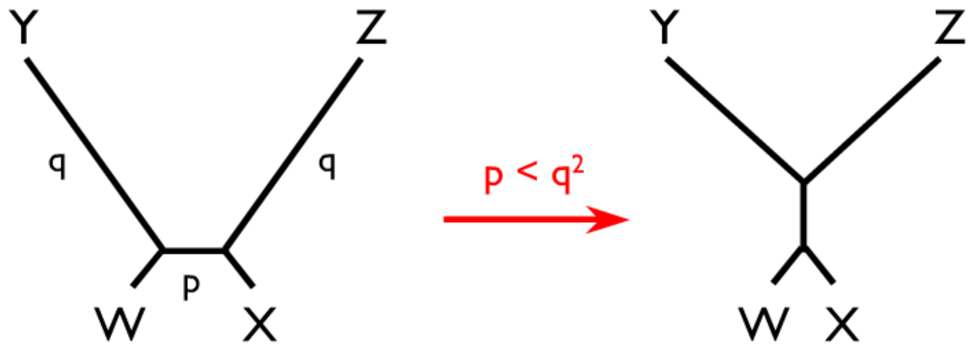

effect of the species sampling
average statistical support
100
75
50
25
0
58
76
88
96
sparse+Ce
dense+Ce
sparse+Xi
dense+Xi
species sampling

effect of the evolutionary model




74
72
84
32
52
53
30
60
Hypsibius
Caenorhabditis
Bombyx
Schmidtea
Schistosoma j.
Platynereis
Euprymna
Ciona i.
Danio
Homo
Hydra
Nematostella
Suberites
Monosiga b.
Neocallimastix
Blastocladiella
Glomus
Rhizopus
Cryptococcus
Ustilago
Schizosaccharomyces
Saccharomyces
Yarrowia
0.2

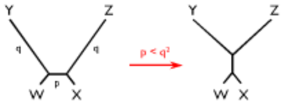
Long branches (old or fast-evolving) contain much more substitutions than reflected in their length.



If undetected, these multiple substitutions generate a non-phylogenetic signal that hinders reconstruction (systematic error).

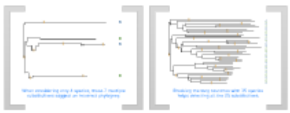
How to reduce systematic error?

Long branches (old or fast-evolving) contain much more substitutions than reflected in their length.



If undetected, these multiple substitutions generate a non-phylogenetic signal that hinders reconstruction (systematic error).

We'd better improve species sampling!



We need better models!

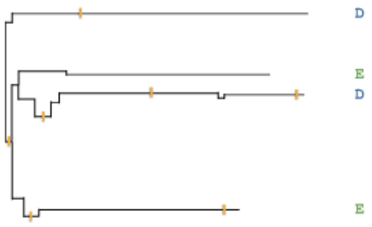
We should use slow-evolving species only!

This is trivial: we simply minimize the number of multiple substitutions.

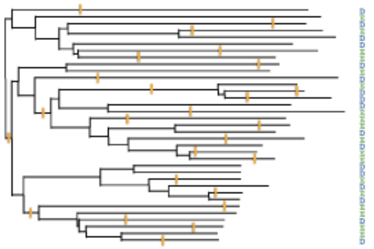
We should use slow-evolving species only!

This is trivial: we simply minimize the number of multiple substitutions.

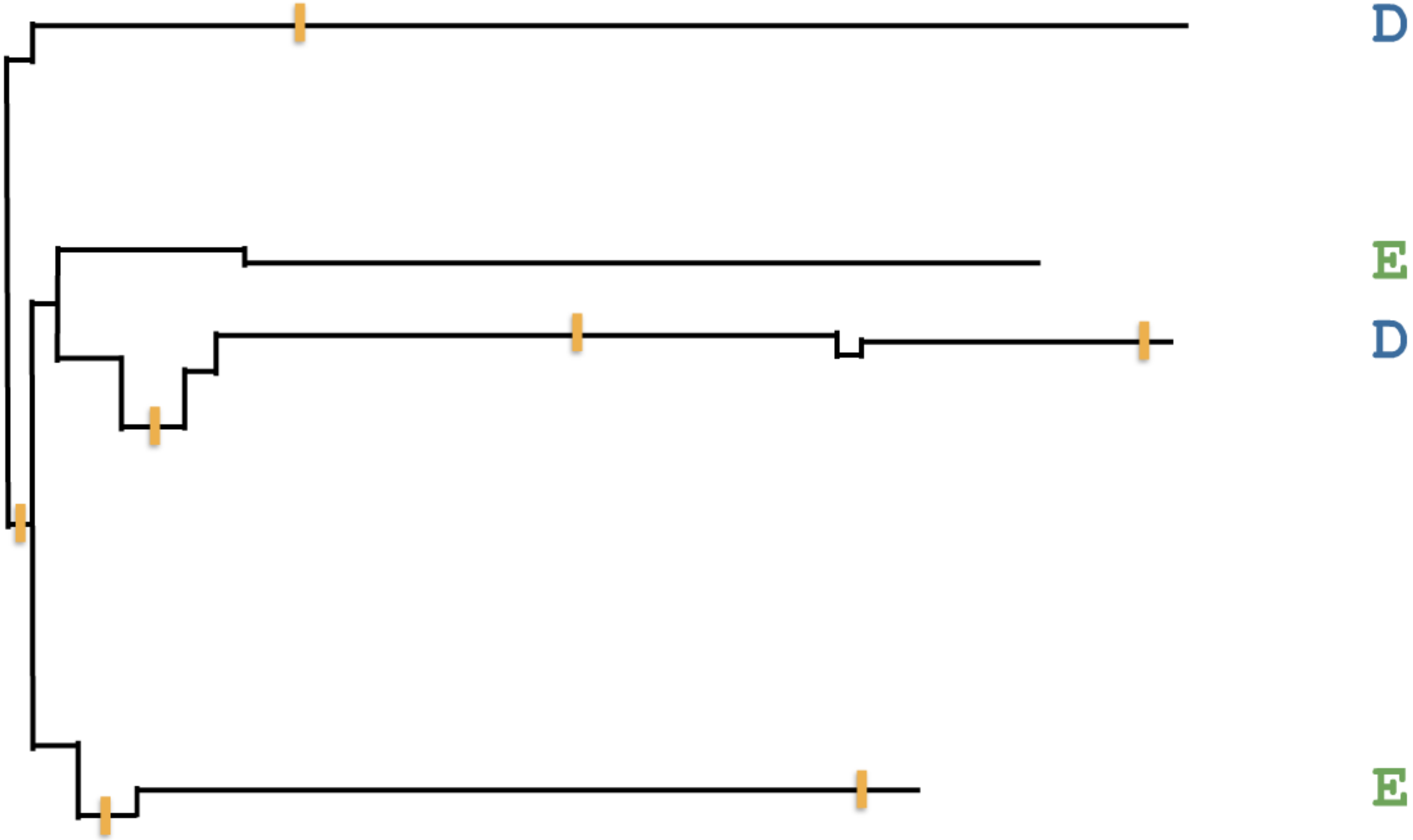
We'd better improve species sampling!



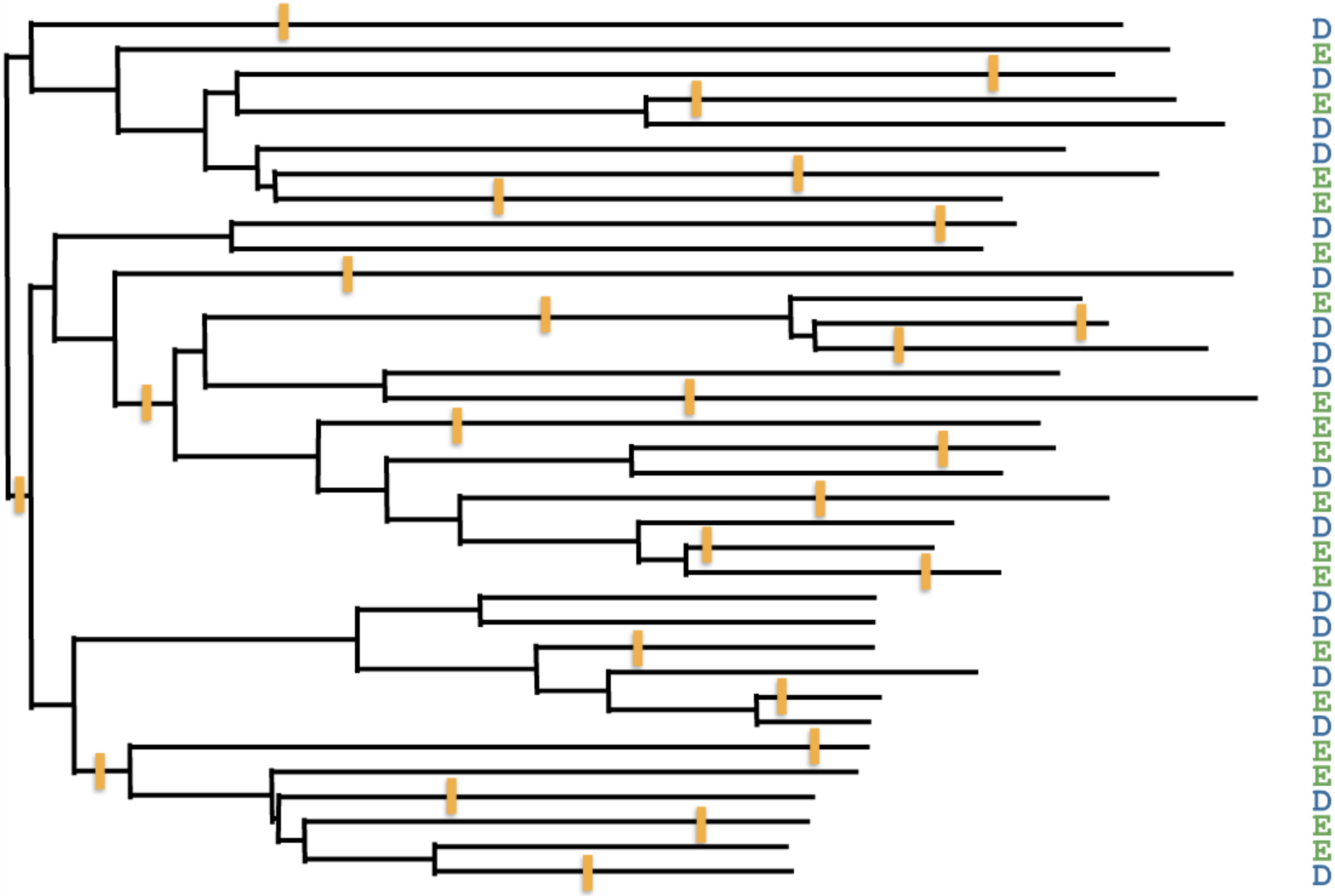
When considering only 4 species, these 7 multiple substitutions suggest an incorrect phylogeny.



Breaking the long branches with 35 species helps detecting all the 25 substitutions.

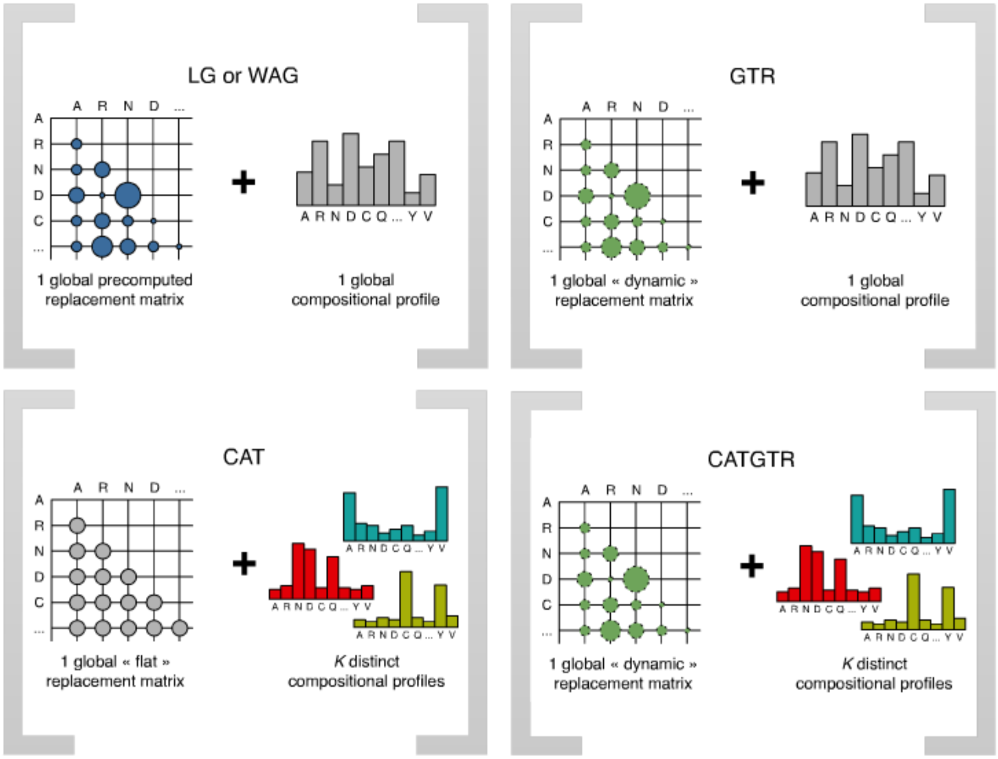


When considering only 4 species, these 7 multiple substitutions suggest an incorrect phylogeny.

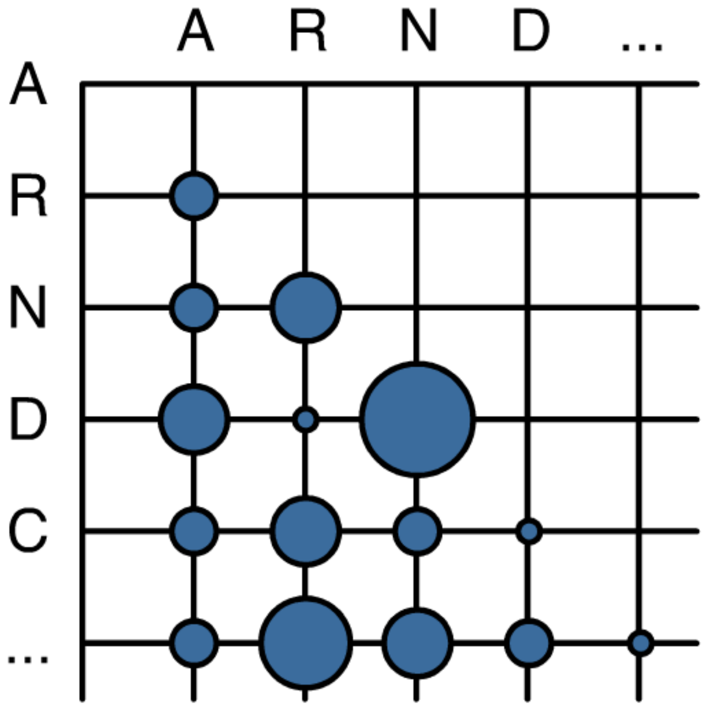


Breaking the long branches with 35 species helps detecting all the 25 substitutions.

We need better models!

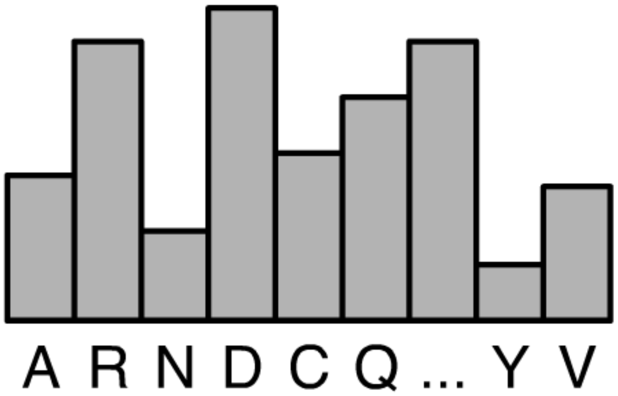


LG or WAG



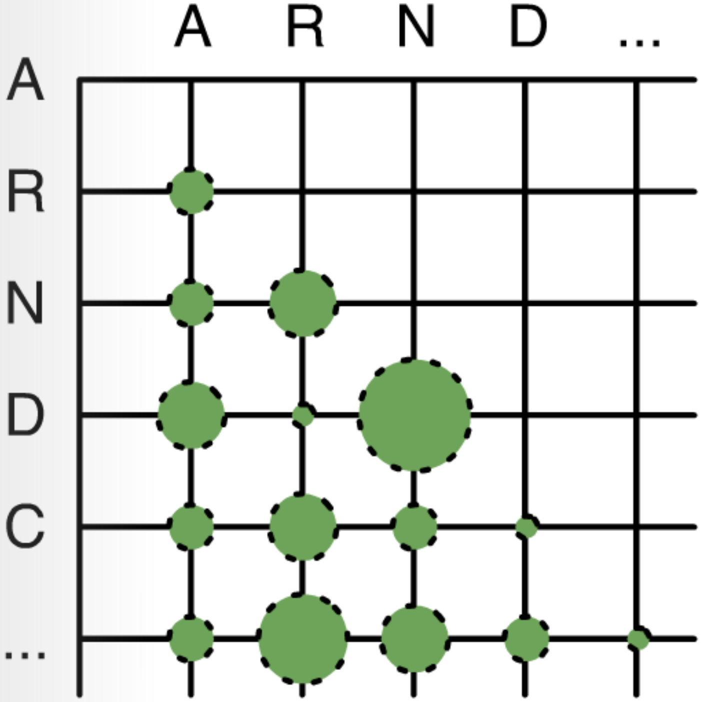
1 global precomputed replacement matrix

+



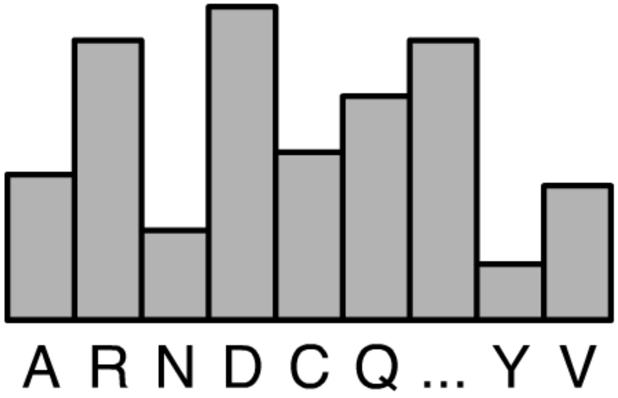
1 global compositional profile

GTR



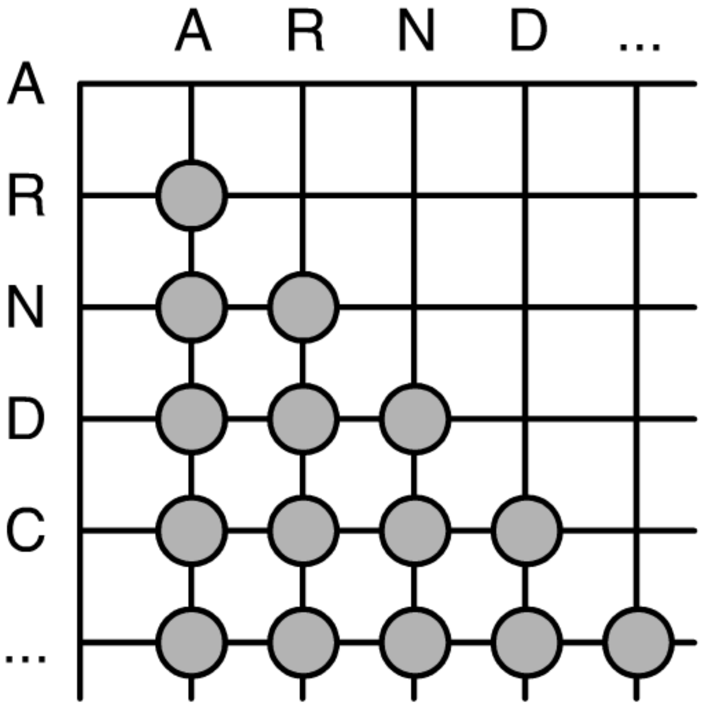
1 global « dynamic » replacement matrix

+



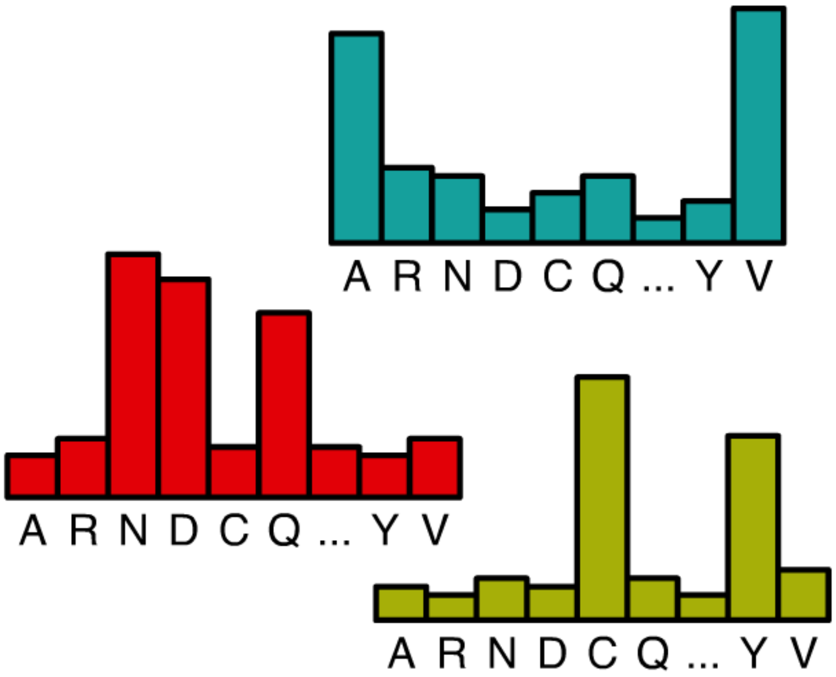
1 global compositional profile

CAT



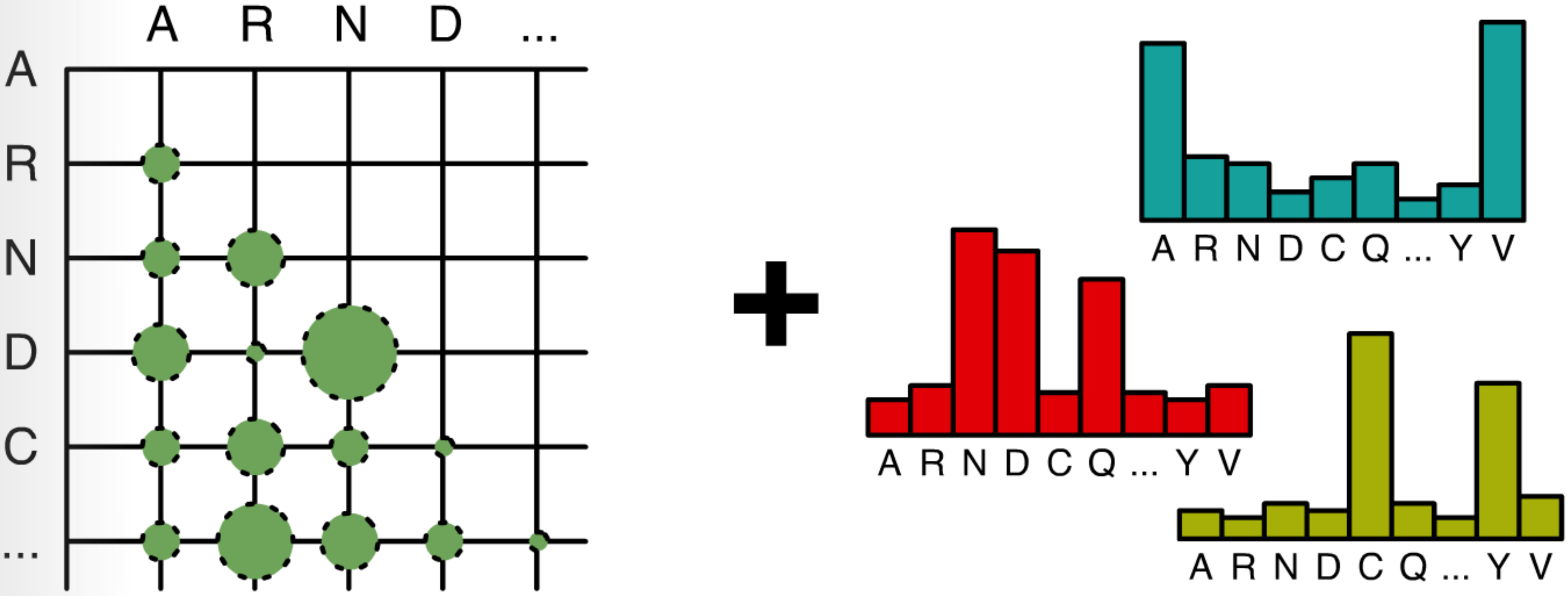
1 global « flat »
replacement matrix

+



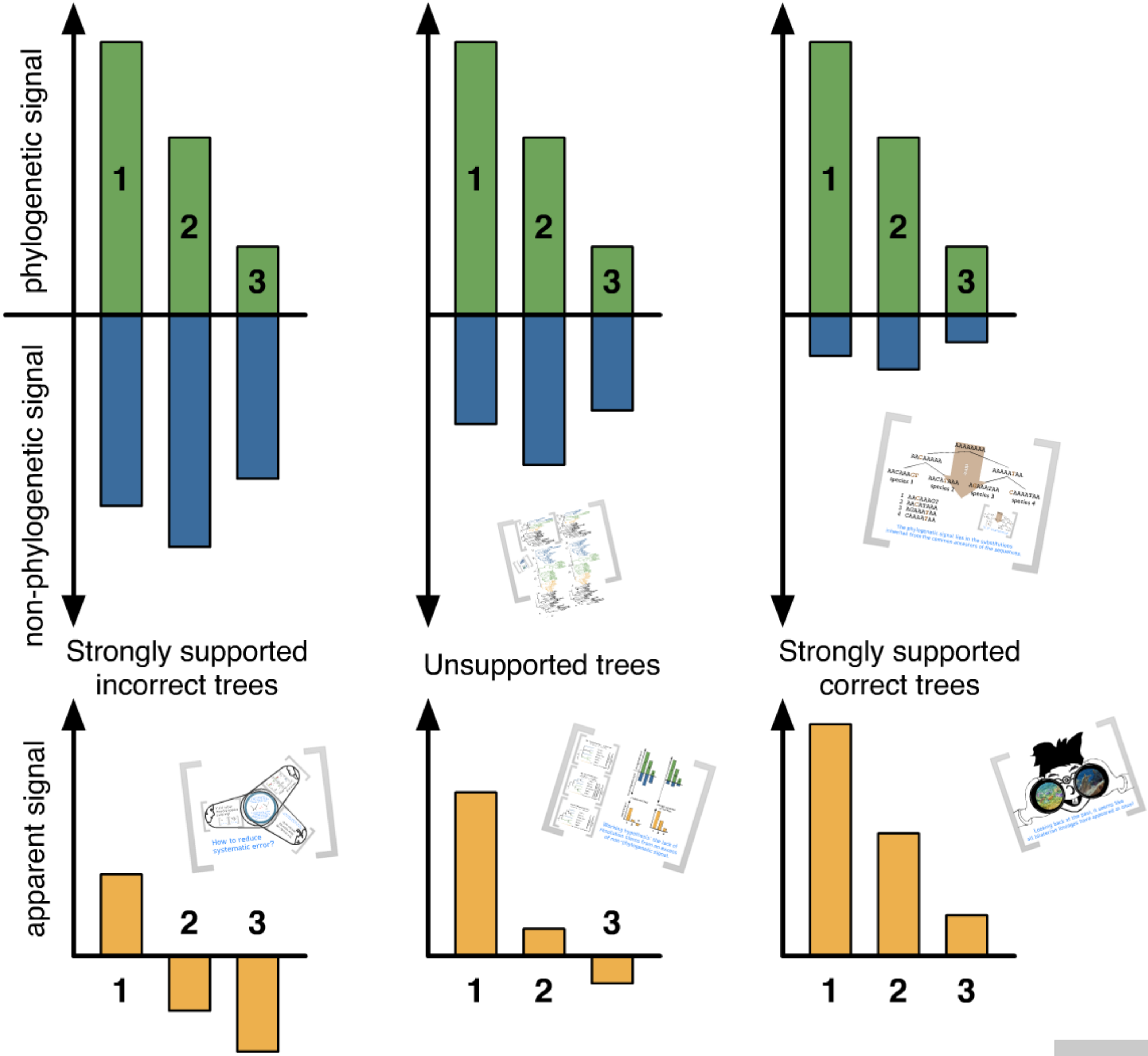
K distinct
compositional profiles

CATGTR

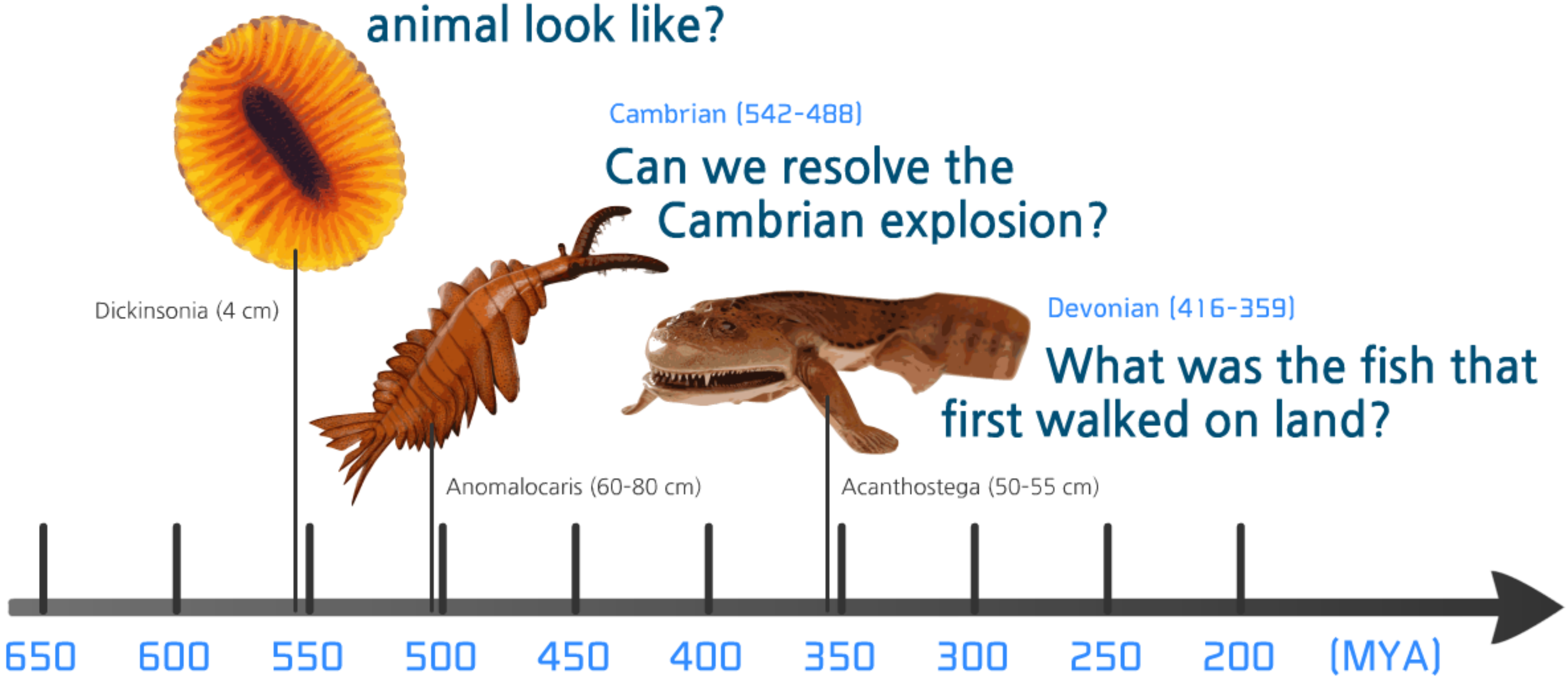


1 global « dynamic » replacement matrix

K distinct compositional profiles


phylogenetic signal
non-phylogenetic signal
1
2
3
Strongly supported incorrect trees
Unsupported trees
Strongly supported correct trees
apparent signal
How to reduce systematic error?
Working hypothesis: the lack of resolution stems from an excess of non-phylogenetic signal.
AAAAAAAA
AACAAAAA
AACAAAGT
species 1
AACATAAA
species 2
AGAAATAA
species 3
AAAAATAA
CAAAATAA
species 4
1 AACAAAGT
2 AACATAAA
3 AGAAATAA
4 CAAAATAA
The phylogenetic signal lies in the substitutions inherited from the common ancestors of the sequences.
Looking back at the past, it seems like all bilaterian lineages have appeared at once!

Computers as Genomic Time Machines for Meeting our Ancestors


Ediacaran (635-542)

What did the first animal look like?

Cambrian (542-488)

Can we resolve the Cambrian explosion?

Devonian (416-359)

What was the fish that first walked on land?

Prof. Denis Baurain — ULg
Bioforum — April 18th, 2013

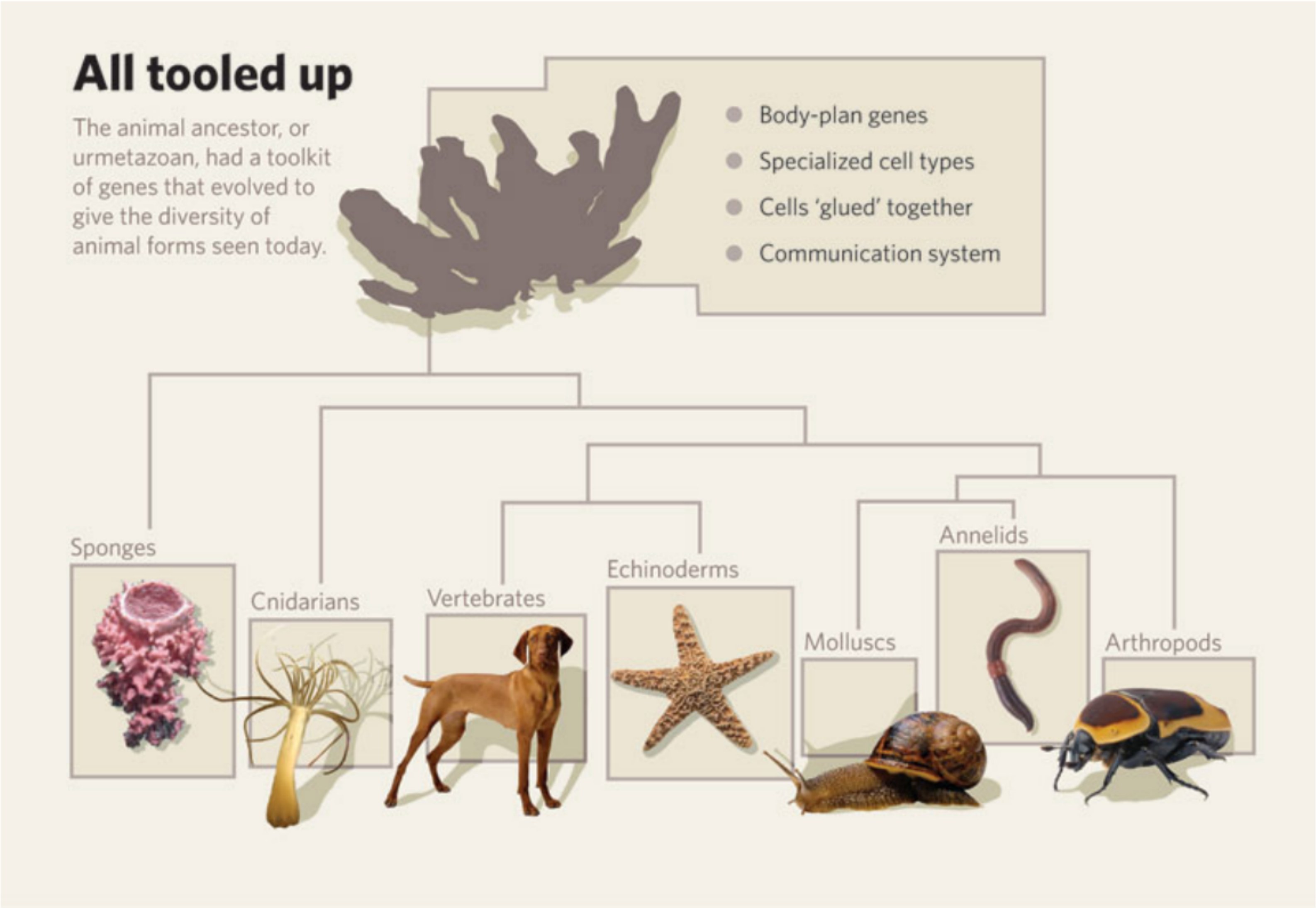
Ediacaran (635-542)

What did the first animal look like?

Cambrian (54

Can we

Camb



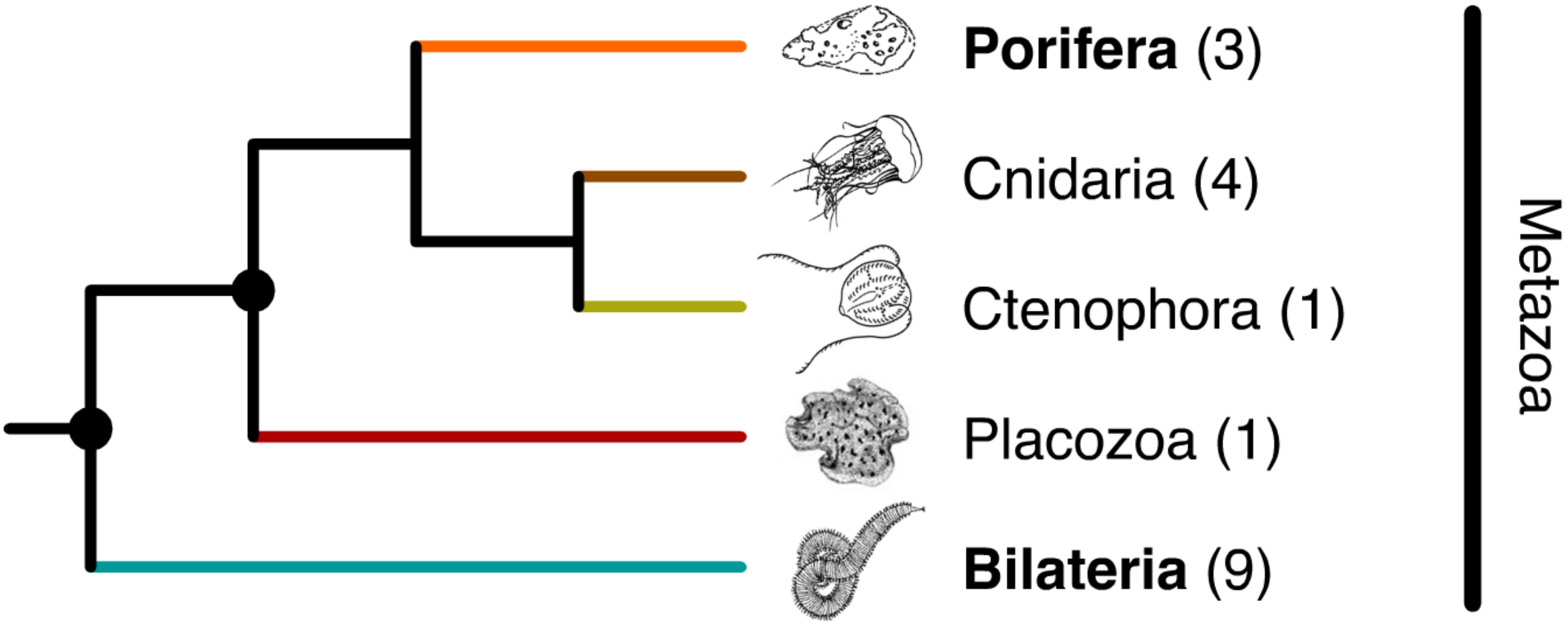
All tooled up

The animal ancestor, or urmetazoan, had a toolkit of genes that evolved to give the diversity of animal forms seen today.

The common ancestor of all animals is called the Urmetazoan. What did it look like?

Schierwater et al. (2009)

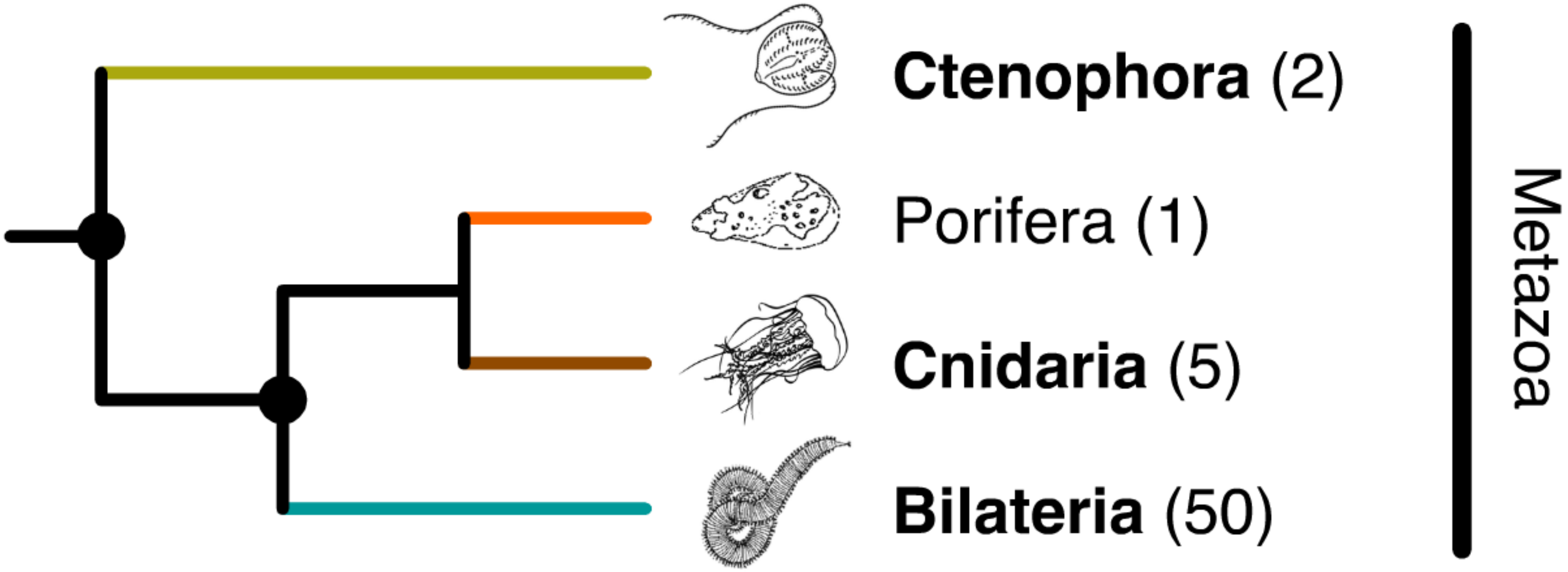
15 mitochondrial and 34 nuclear genes
GTR and WAG models



It was already very complex! All lineages are thus evolutionarily simplified except bilaterians.

Dunn et al. (2008)

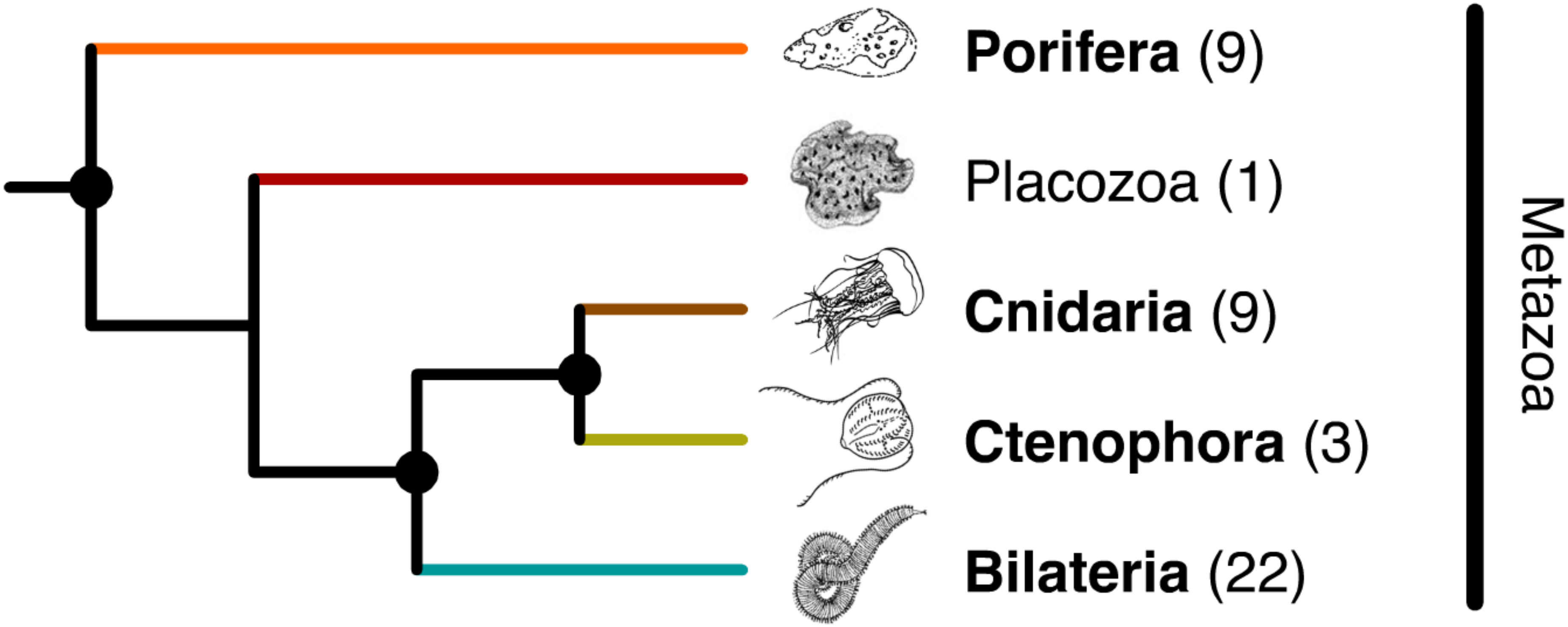
6 mitochondrial and 144 nuclear genes
WAG and CAT models



It featured several complex characters (e.g., neurons)! Or these have evolved convergently…

Philippe et al. (2009)

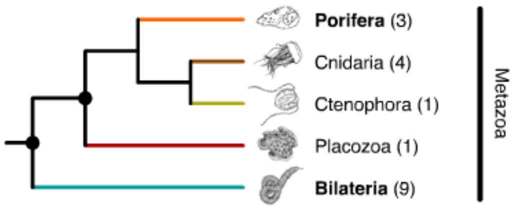
128 nuclear genes
CAT model



It might have been quite simple. Hey! Which one of these 3 strongly supported trees is correct?

Schierwater et al. (2009)

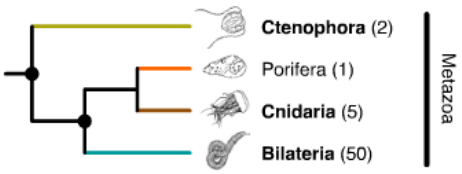
15 mitochondrial and 34 nuclear genes
GTR and WAG models



It was already very complex! All lineages are thus evolutionarily simplified except bilaterians.

Dunn et al. (2008)

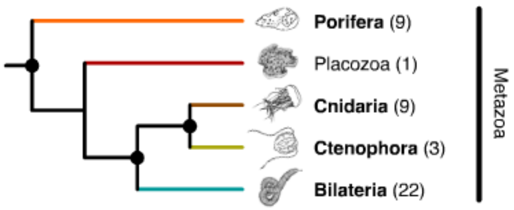
6 mitochondrial and 144 nuclear genes
WAG and CAT models



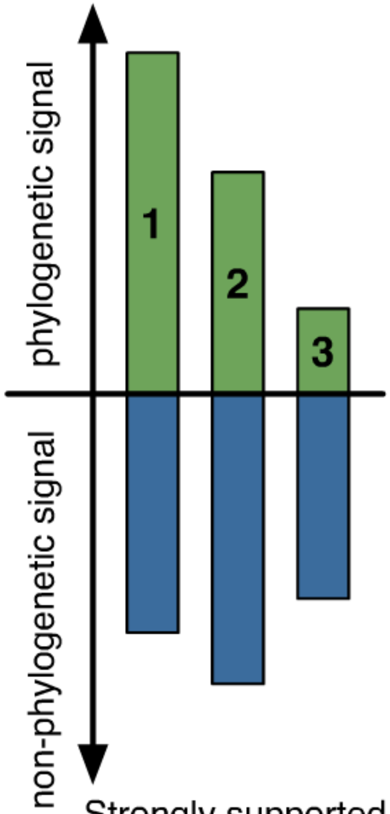
It featured several complex characters (e.g., neurons)! Or these have evolved convergently...

Philippe et al. (2009)

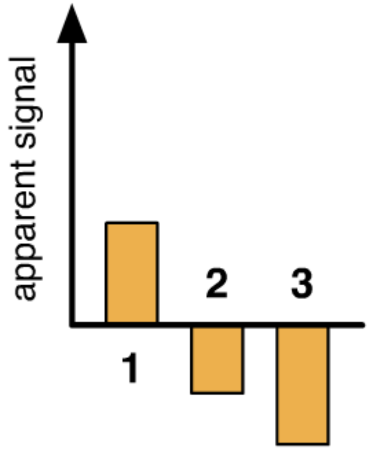
128 nuclear genes
CAT model



It might have been quite simple. Hey! Which one of these 3 strongly supported trees is correct?

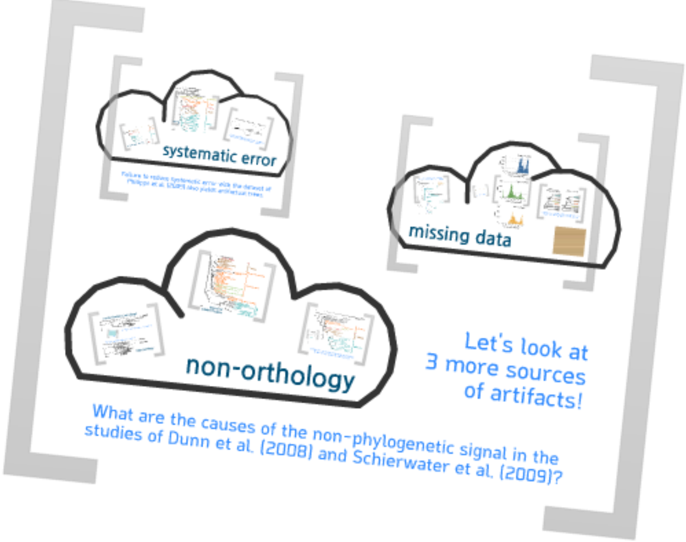


Strongly supported incorrect trees



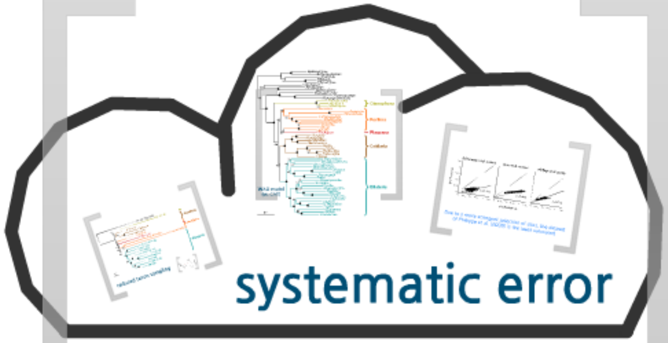
For nodes with a scarce phylogenetic signal, even small amounts of non-phylogenetic signal may dominate and yield an incorrect tree.

Ironically, these nodes are those for which phylogenomics would be the most useful!



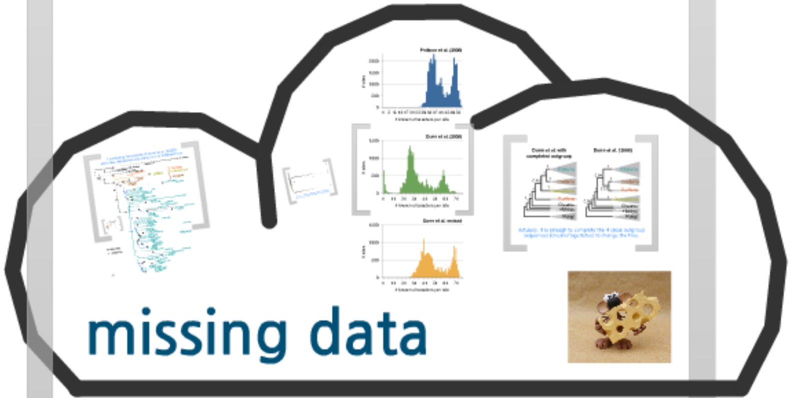
Let's look at 3 more sources of artifacts!

What are the causes of the non-phylogenetic signal in the studies of Dunn et al. (2008) and Schierwater et al. (2009)?



systematic error

Failure to reduce systematic error with the dataset of Philippe et al. (2009) also yields artifactual trees.



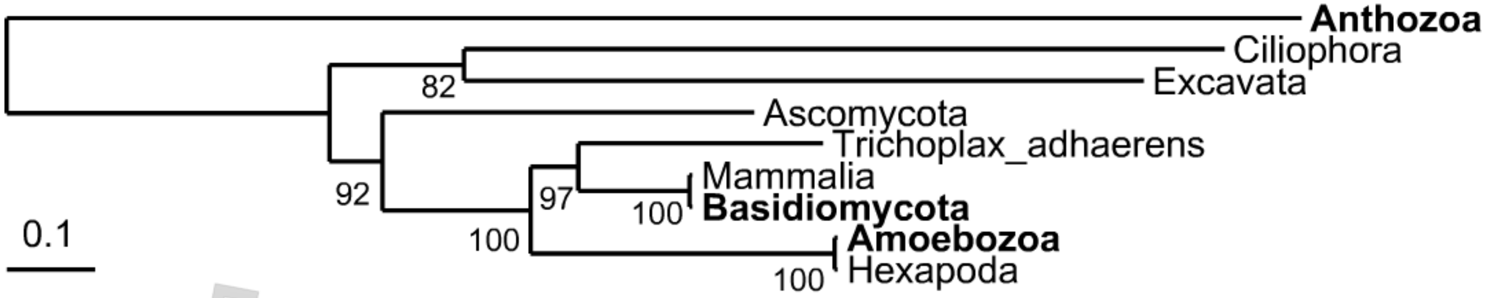
missing data

non-orthology

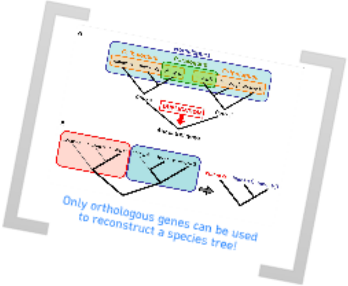
Let's look at
3 more sources
of artifacts!

What are the causes of the non-phylogenetic signal in the studies of Dunn et al. (2008) and Schierwater et al. (2009)?

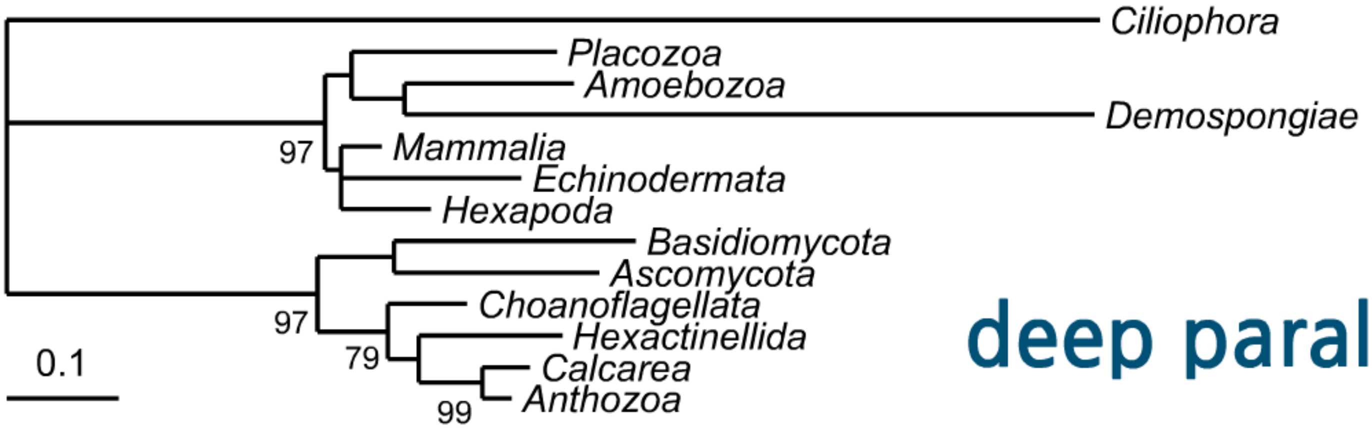
contaminations (xenology)



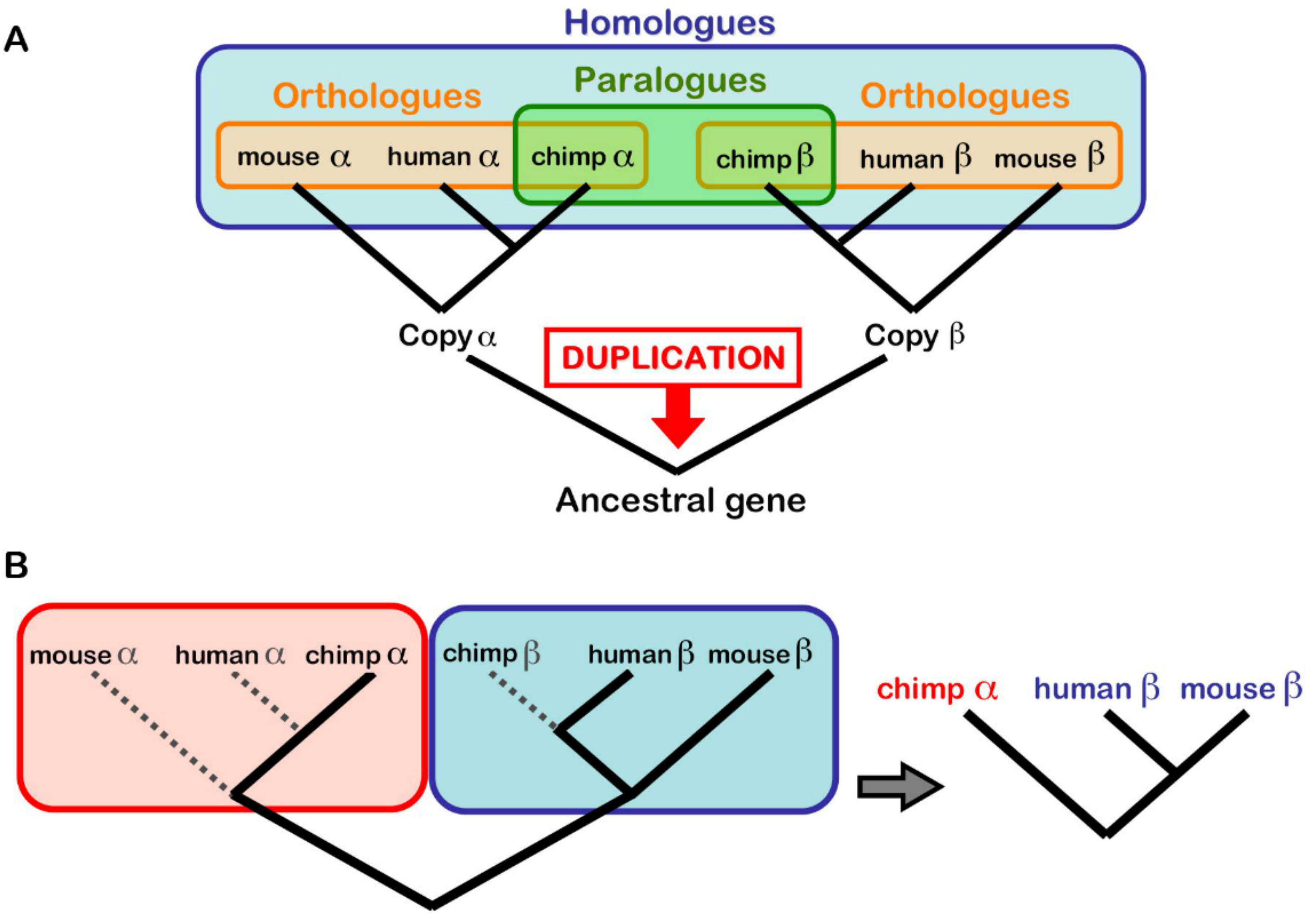
RP3 gene in Schierwater et al. (2009)


Only orthologous genes can be used to reconstruct a species tree!

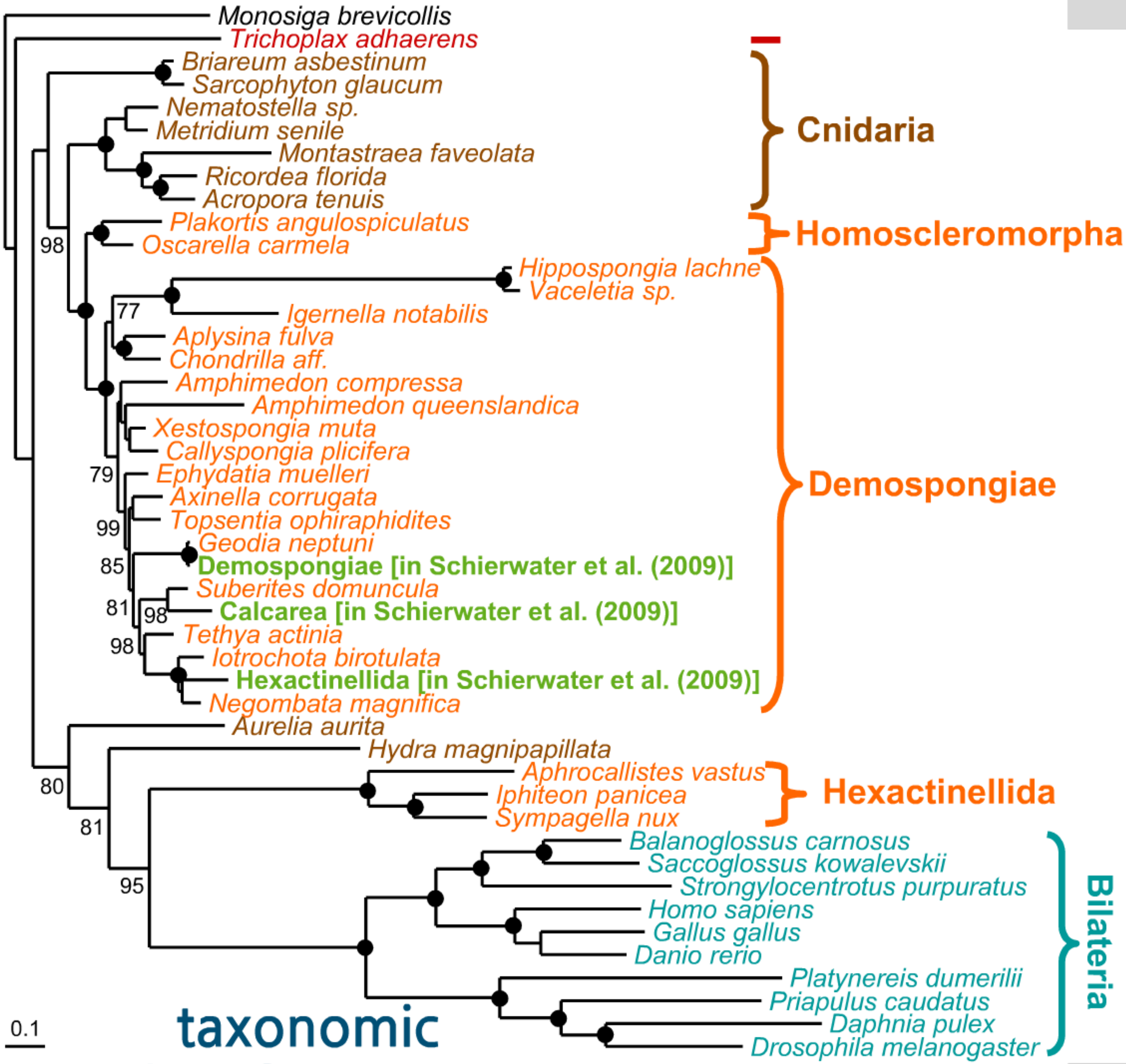
CDC gene in Schierwater et al. (2009)



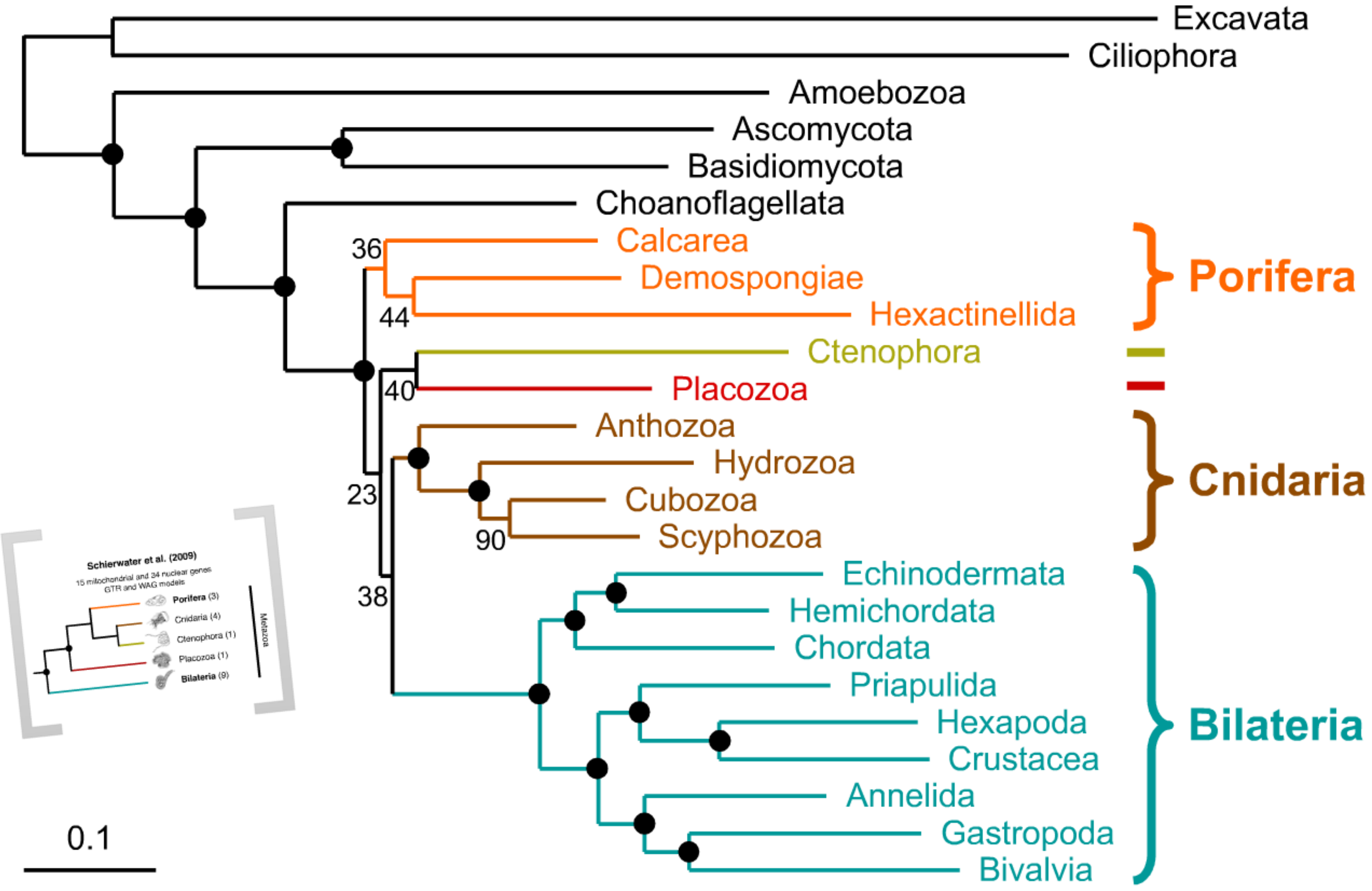
deep paralogy



Only orthologous genes can be used to reconstruct a species tree!



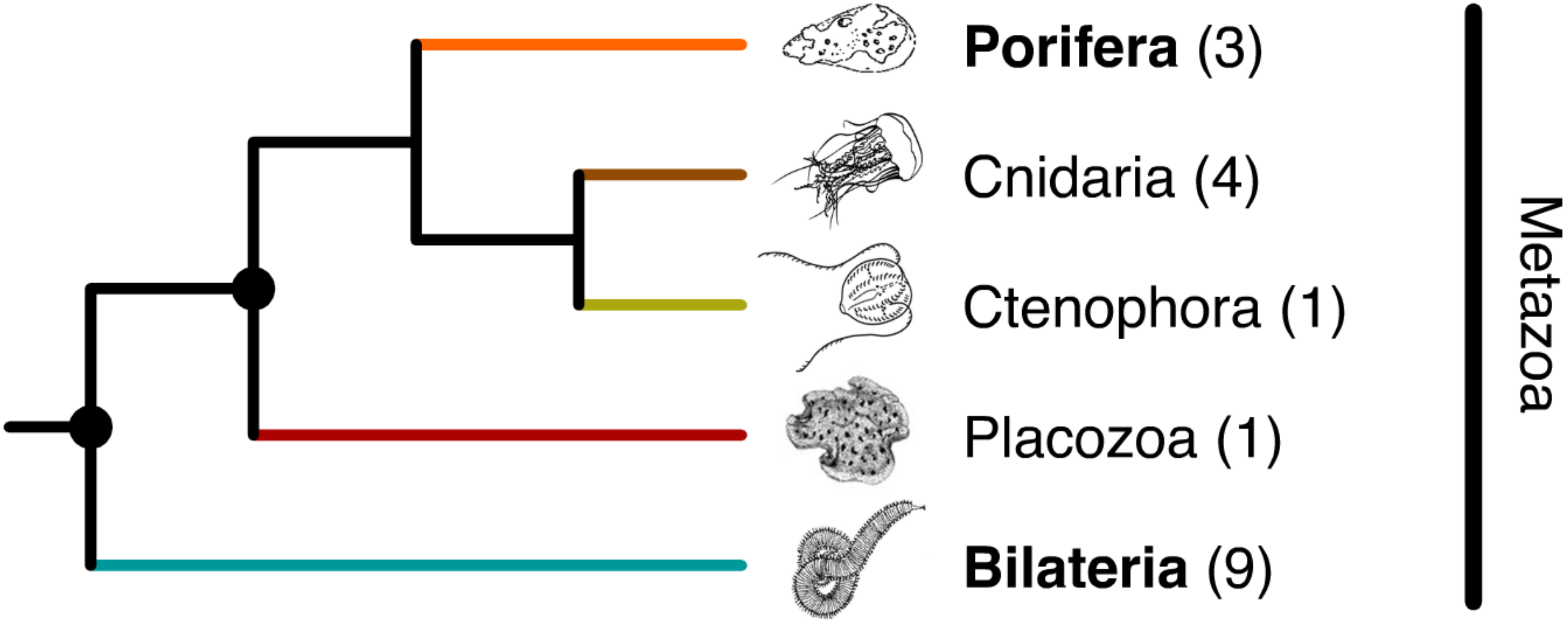
taxonomic misidentification

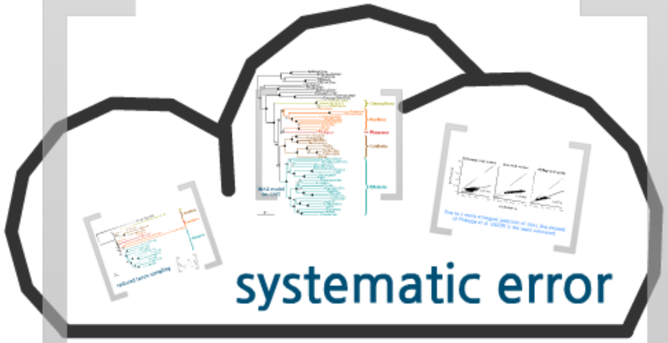


Purging the dataset of Schierwater et al. (2009)
of its errors yields a very different tree.

Schierwater et al. (2009)

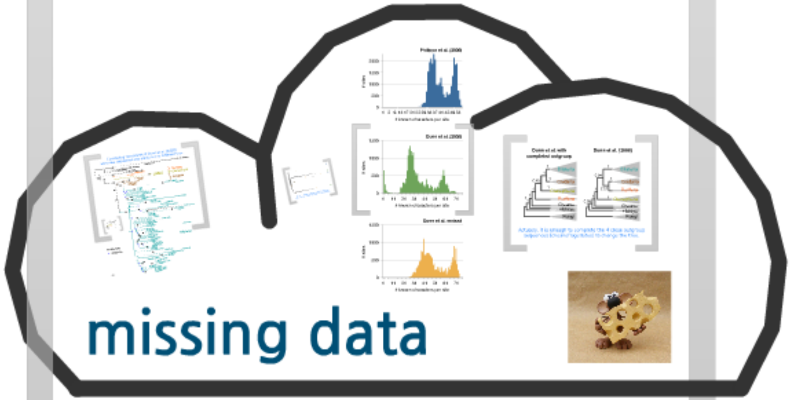
15 mitochondrial and 34 nuclear genes
GTR and WAG models



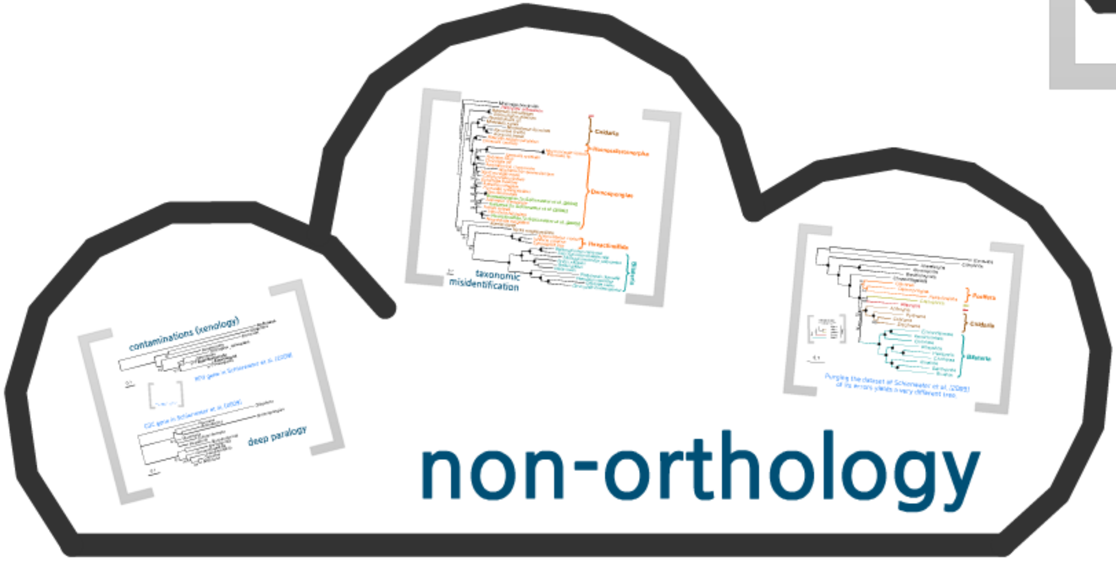


systematic error

Failure to reduce systematic error with the dataset of Philippe et al. (2009) also yields artifactual trees.



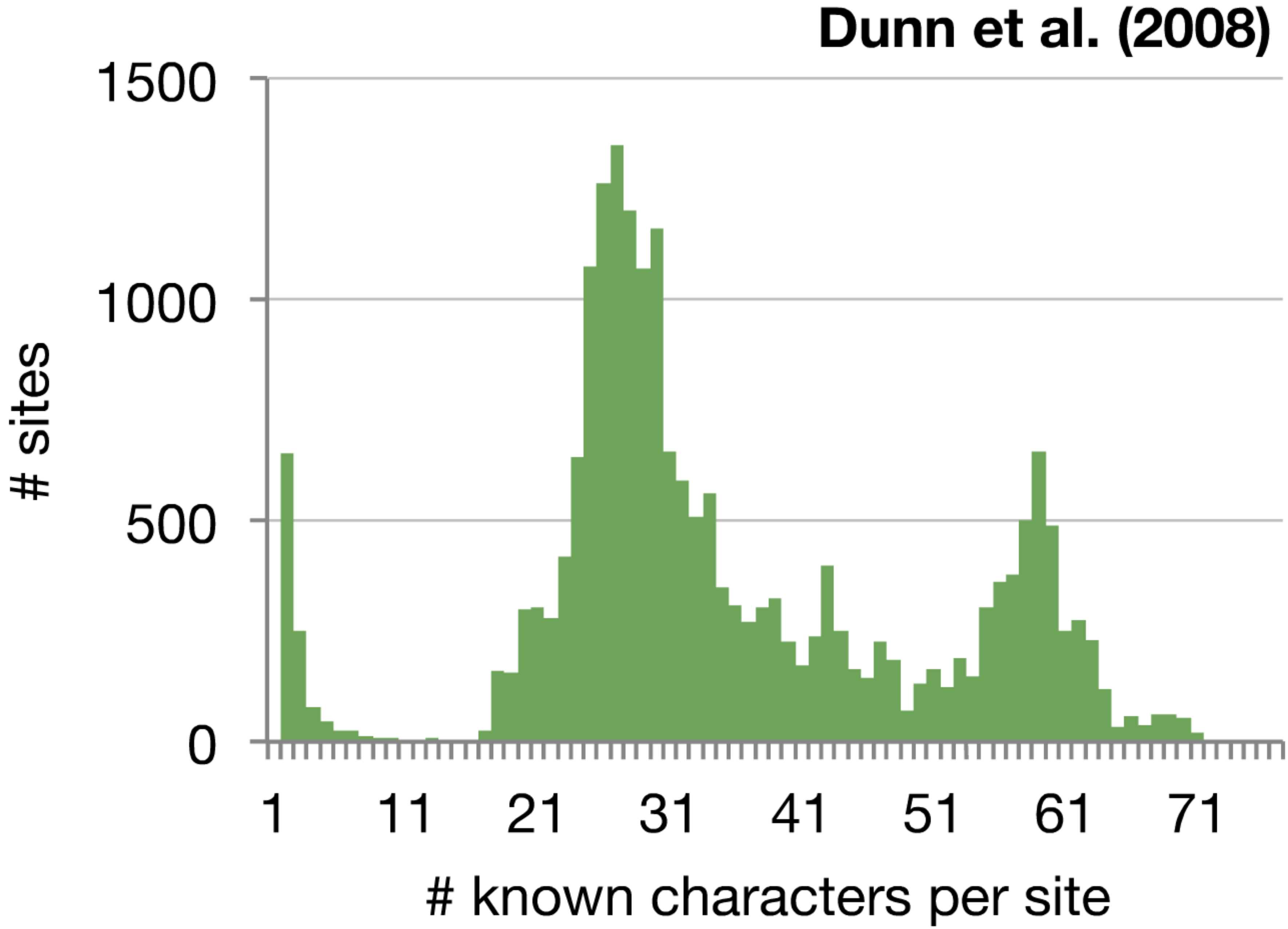
missing data

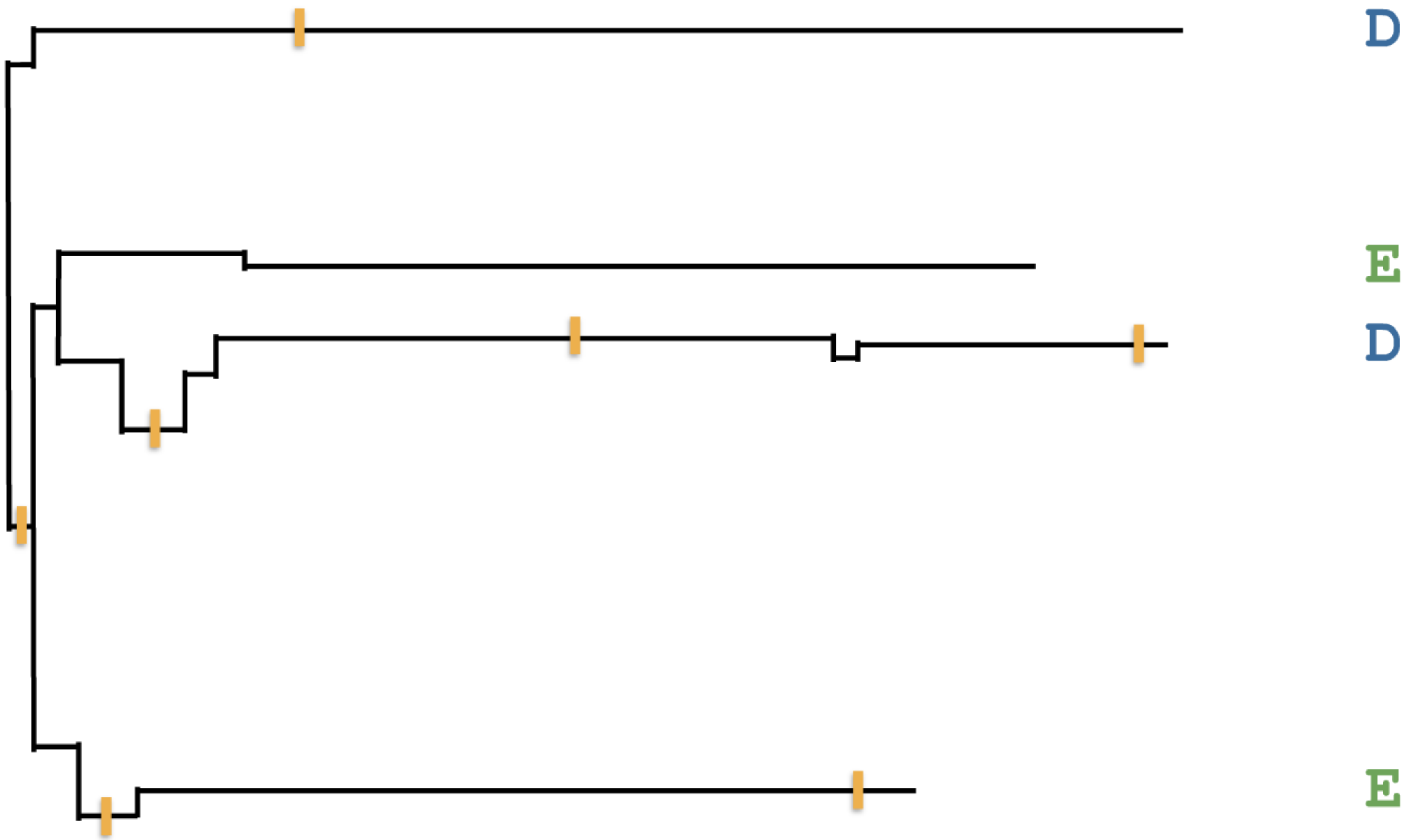


non-orthology

Let's look at
3 more sources
of artifacts!

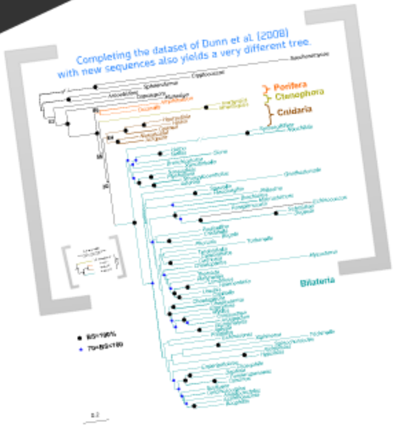
What are the causes of the non-phylogenetic signal in the studies of Dunn et al. (2008) and Schierwater et al. (2009)?


Dunn et al. (2008)
sites
1500
1000
500
0
1
11
21
31
41
51
61
71
known characters per site

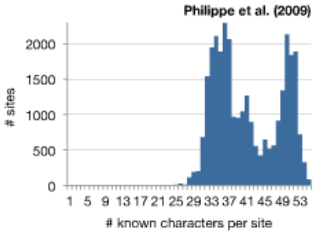


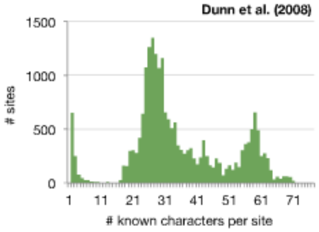
Missing data exacerbate the systematic error by reducing the number of species effectively available.

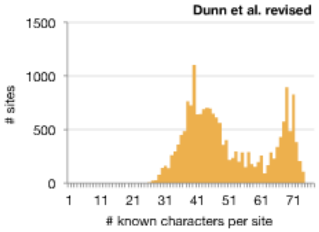
missing data

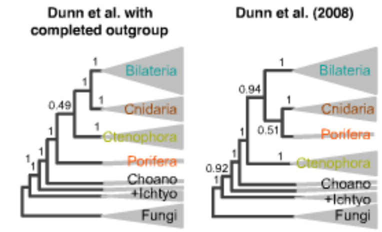


Completing the dataset of Dunn et al. (2008) with new sequences also yields a very different tree.



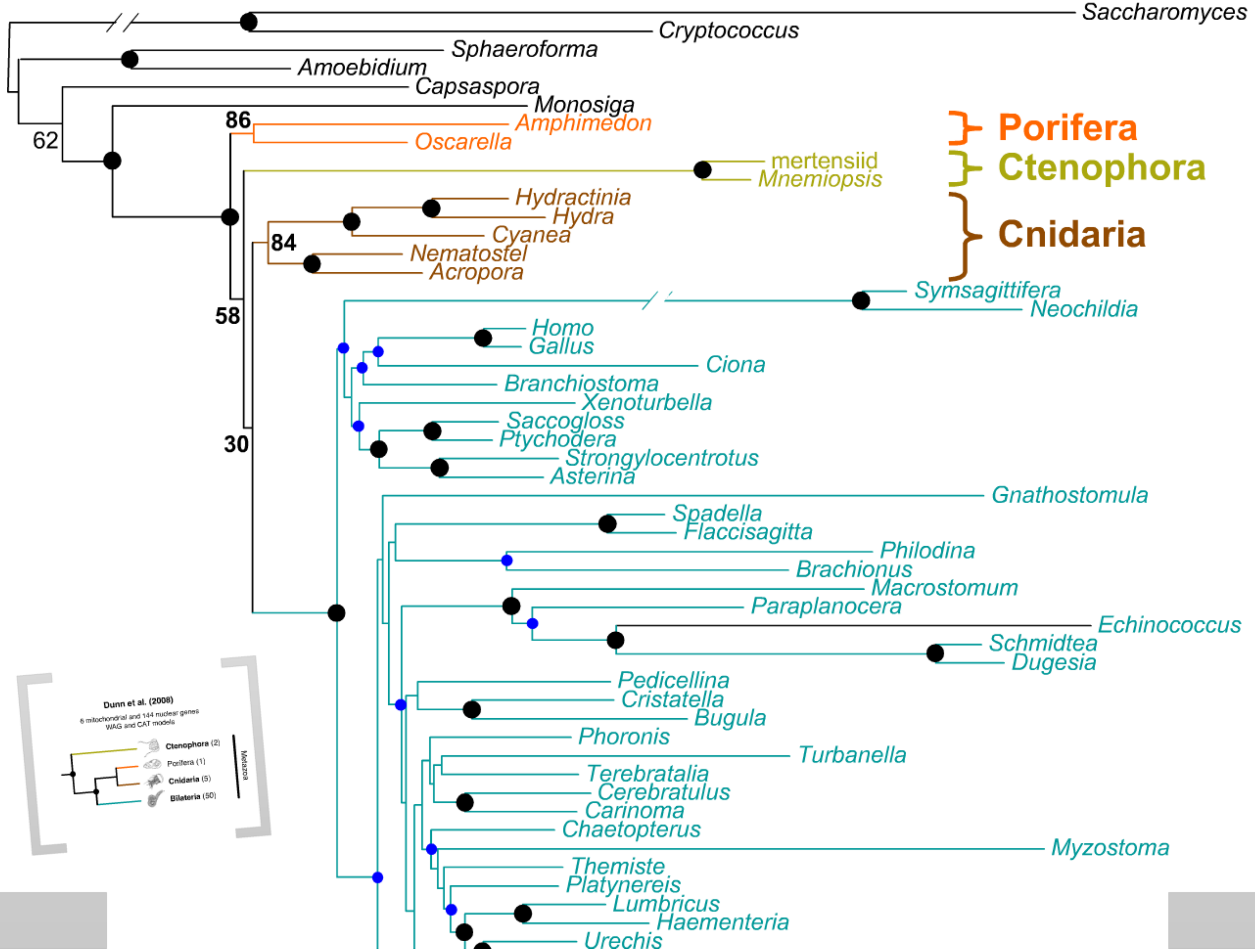






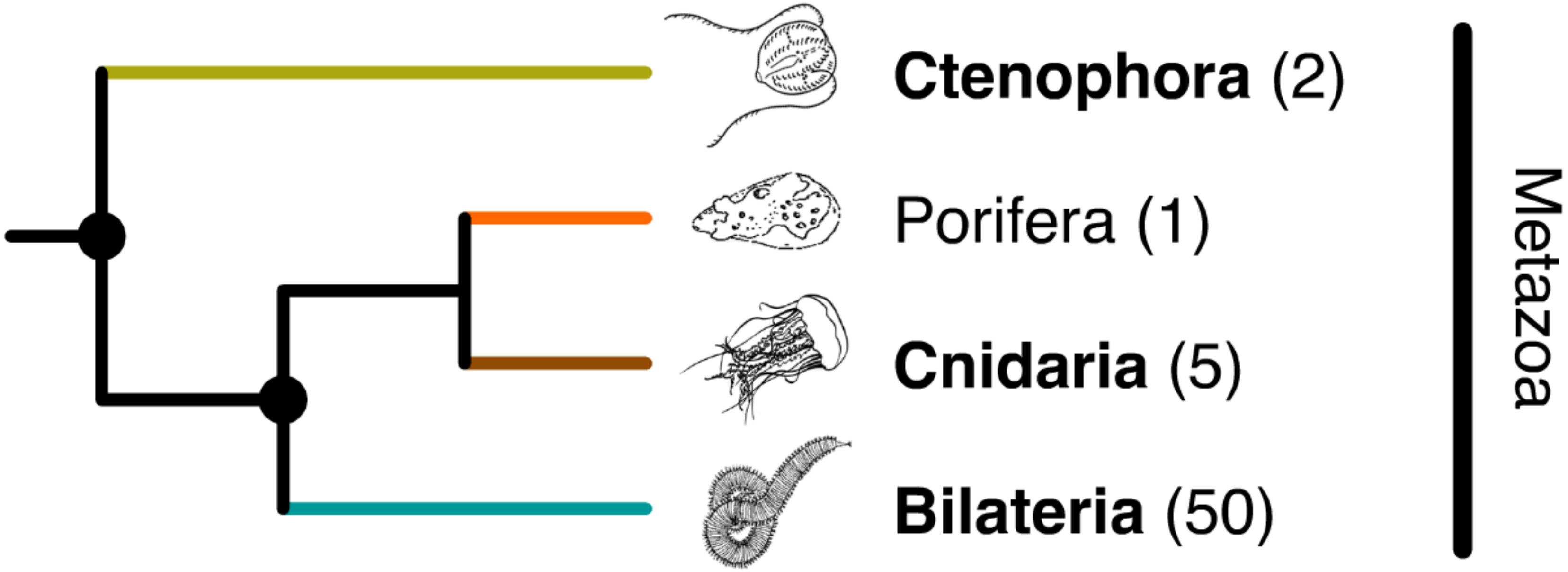
Actually, it is enough to complete the 4 close outgroup sequences (choanoflagellates) to change the tree.

Completing the dataset of Dunn et al. (2008) with new sequences also yields a very different tree.



Dunn et al. (2008)

6 mitochondrial and 144 nuclear genes
WAG and CAT models



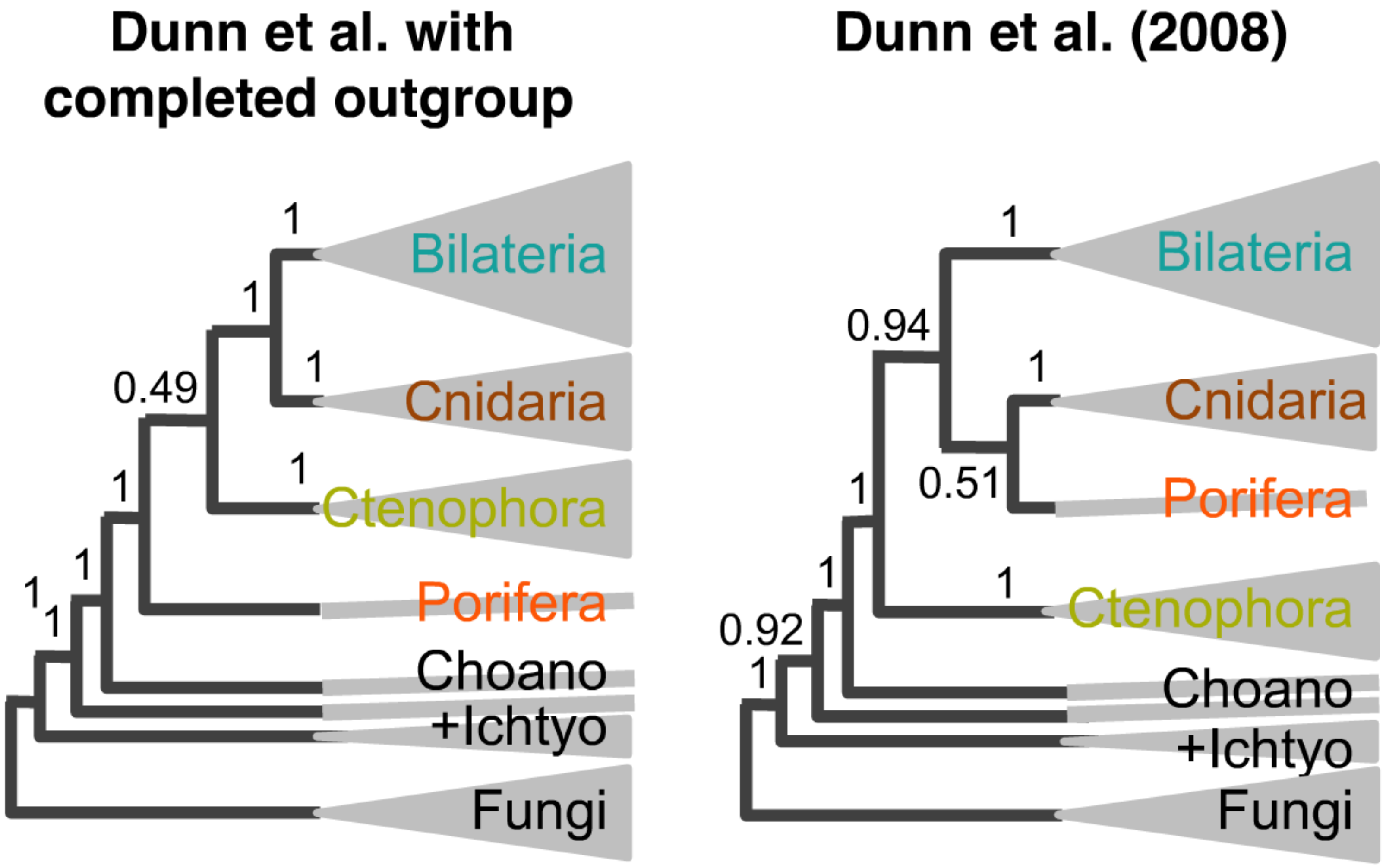
Ctenophora (2)

Porifera (1)

Cnidaria (5)

Bilateria (50)

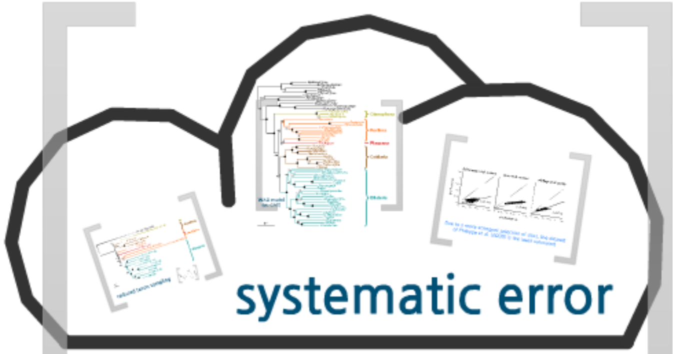
Metazoa



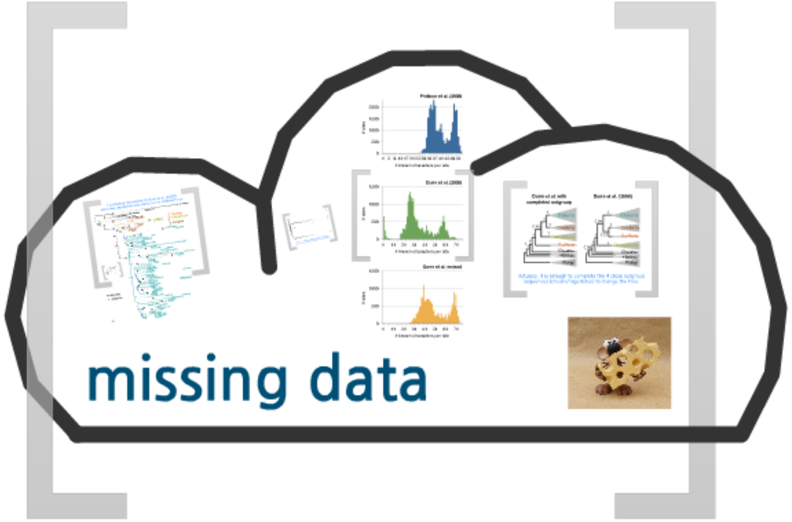
Dunn et al. with completed outgroup

Dunn et al. (2008)

Actually, it is enough to complete the 4 close outgroup sequences (choanoflagellates) to change the tree.



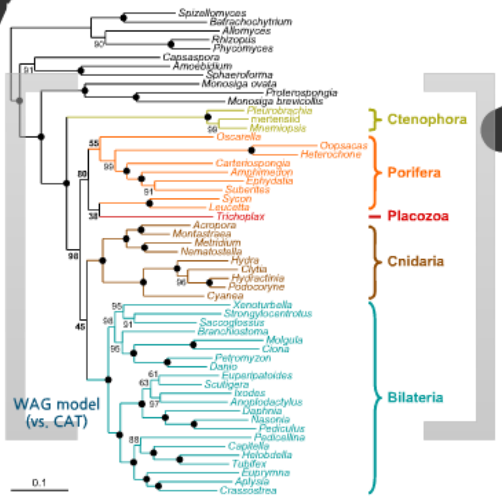
Failure to reduce systematic error with the dataset of Philippe et al. (2009) also yields artifactual trees.



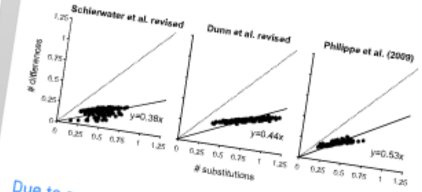
Purging the dataset of Schierwater et al. (2009) of its errors yields a very different tree.

Let's look at
3 more sources
of artifacts!

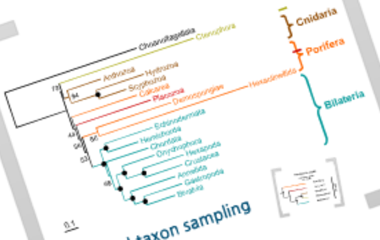
What are the causes of the non-phylogenetic signal in the studies of Dunn et al. (2008) and Schierwater et al. (2009)?



WAG model (vs. CAT)



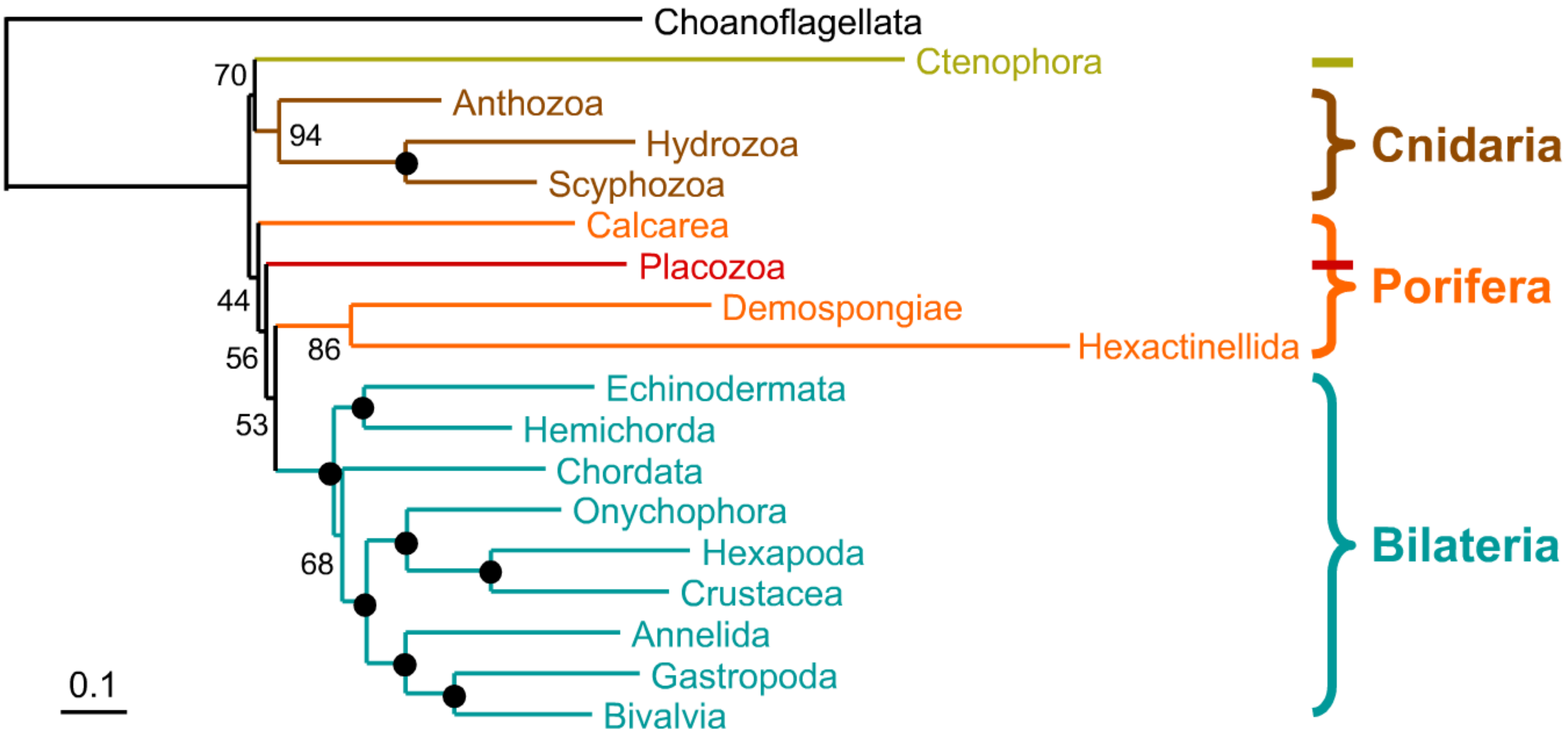
Due to a more stringent selection of sites, the dataset of Philippe et al. (2009) is the least saturated.



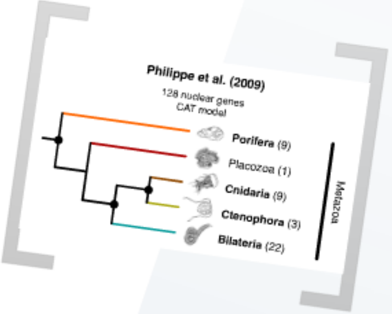
reduced taxon sampling

systematic error

Failure to reduce systematic error with the dataset of Philippe et al. (2009) also yields artifactual trees.

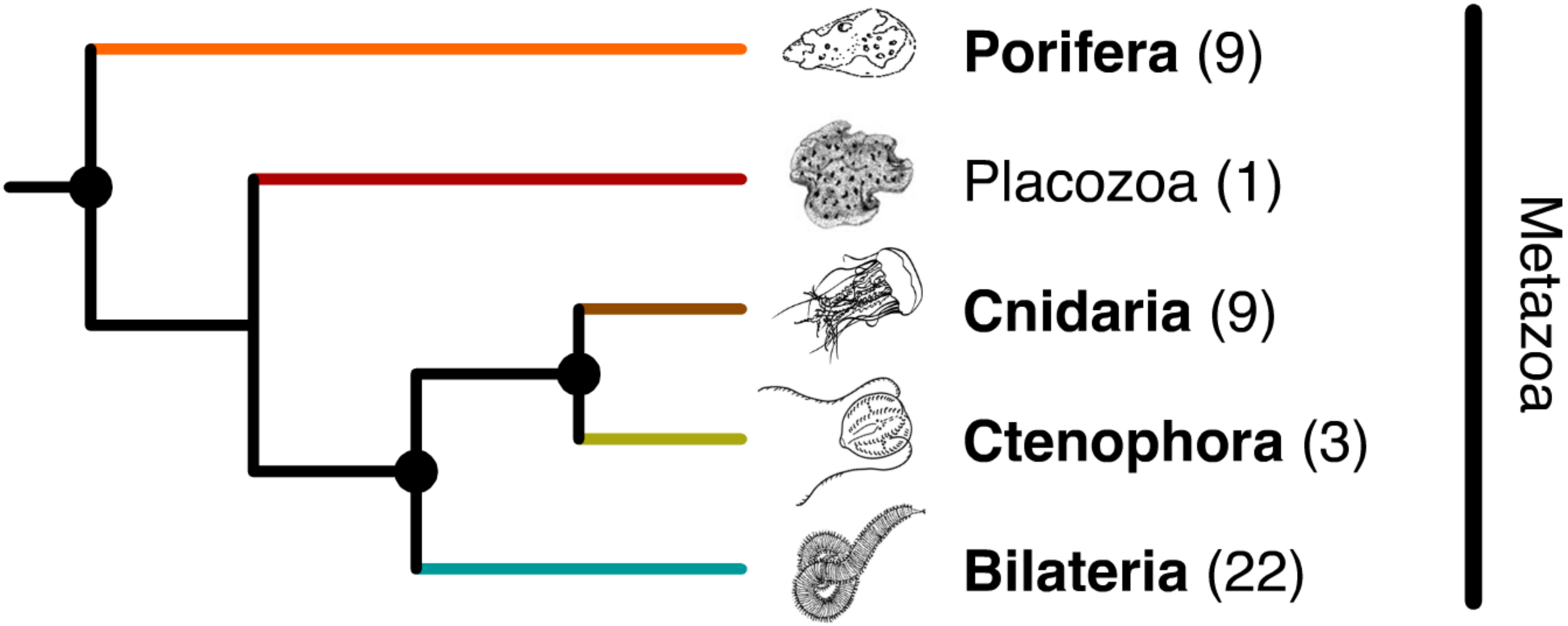


reduced taxon sampling



Philippe et al. (2009)

128 nuclear genes
CAT model



Porifera (9)

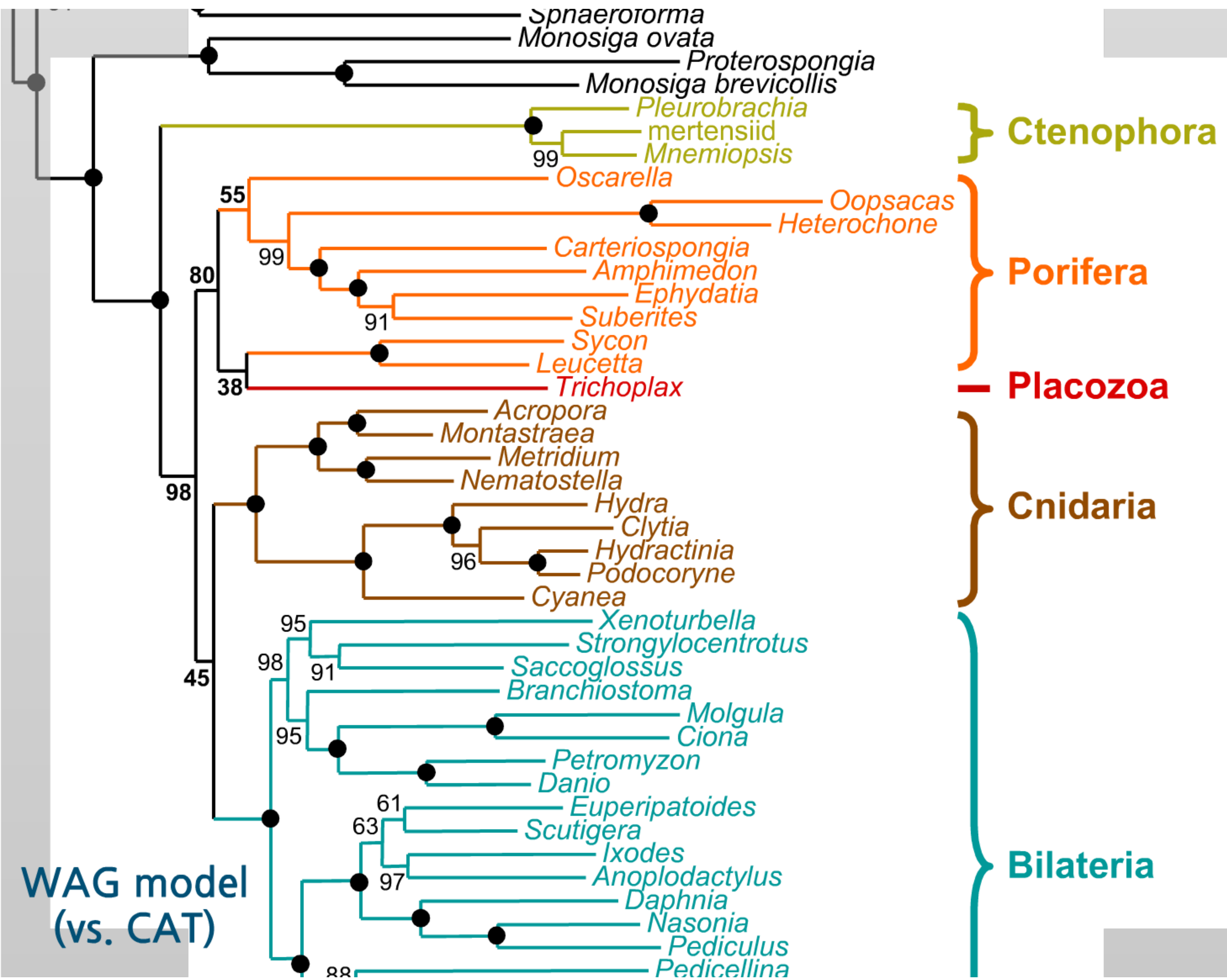
Placozoa (1)

Cnidaria (9)

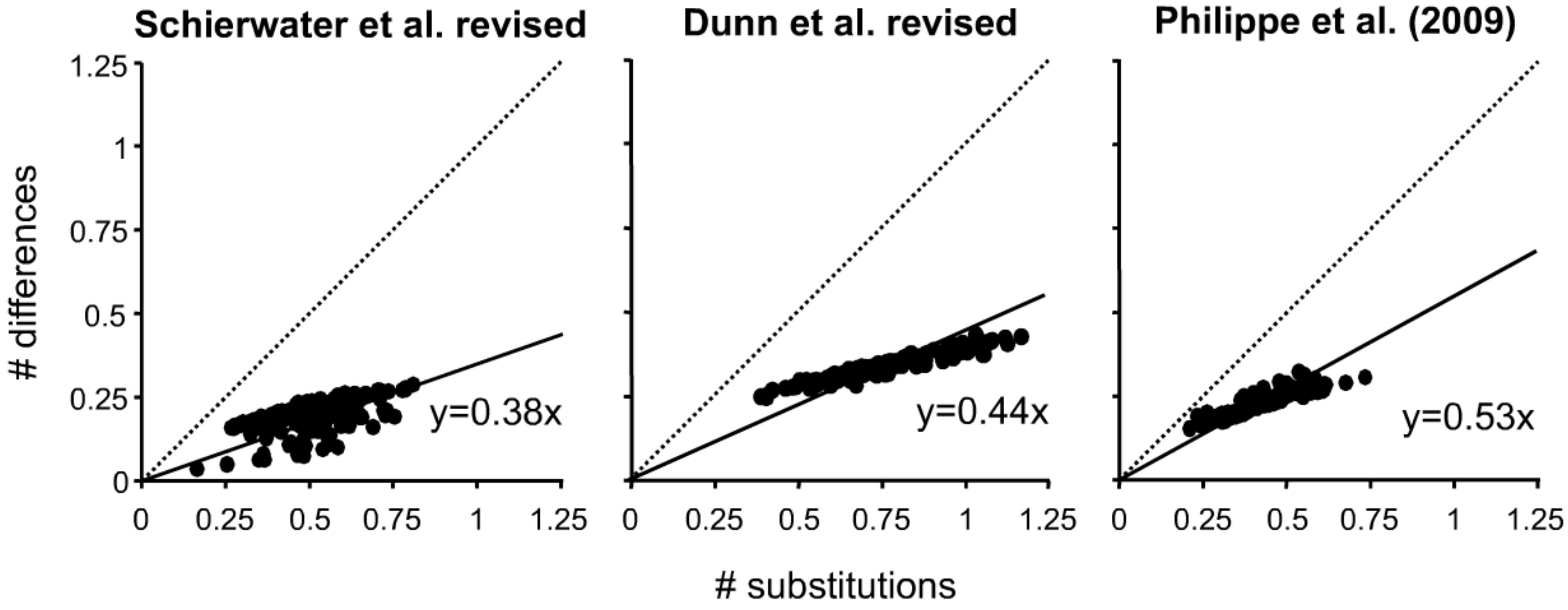
Ctenophora (3)

Bilateria (22)

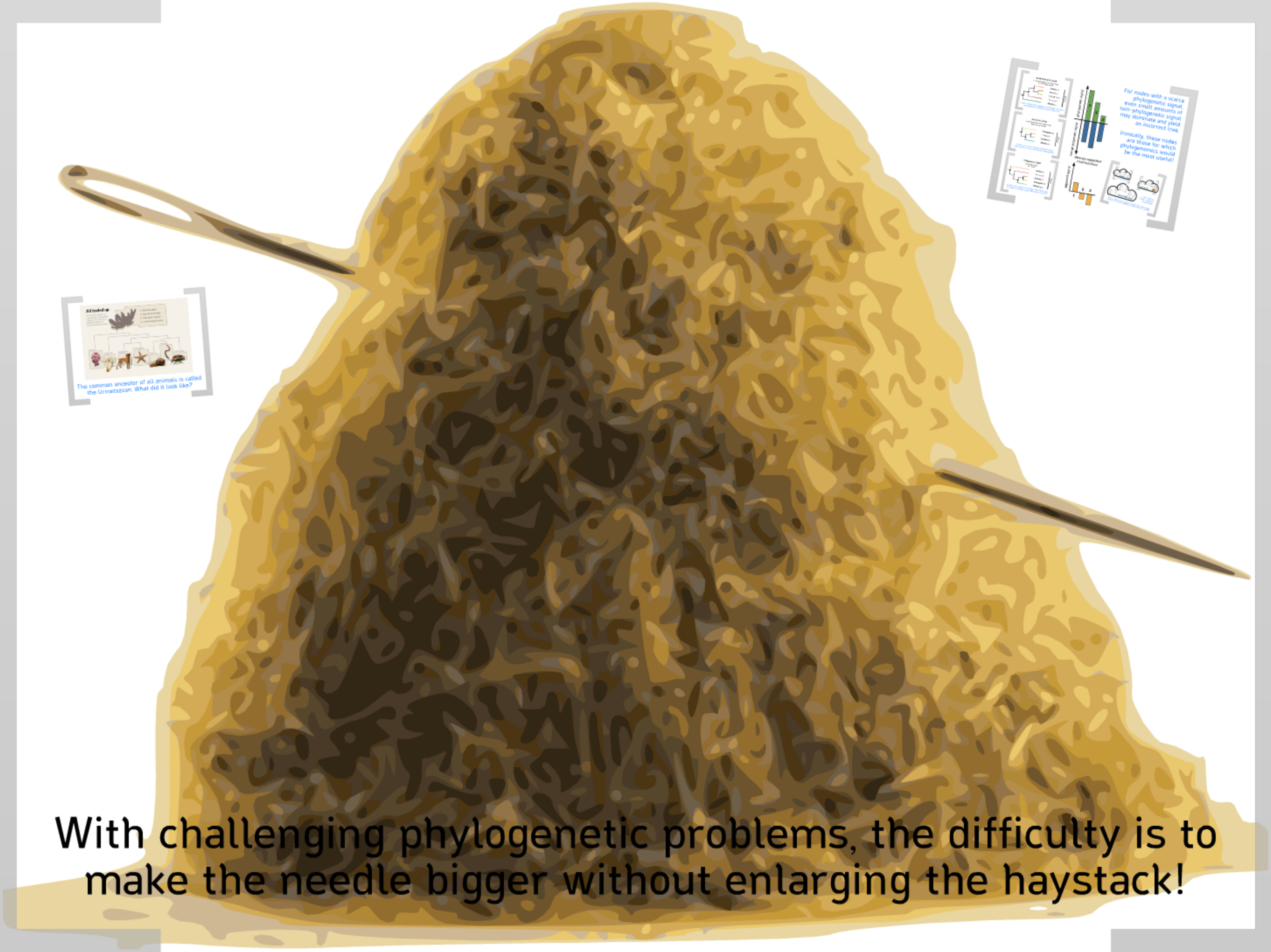
Metazoa



WAG model
(vs. CAT)



Due to a more stringent selection of sites, the dataset of Philippe et al. (2009) is the least saturated.



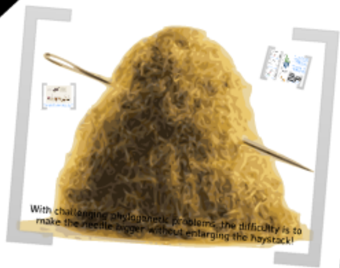
All tooled up

The common ancestor of all animals is called the Urmetazoan. What did it look like?

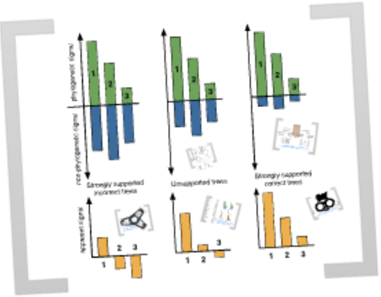
For nodes with a scarce phylogenetic signal, even small amounts of non-phylogenetic signal may dominate and yield an incorrect tree.

Ironically, these nodes are those for which phylogenomics would be the most useful!

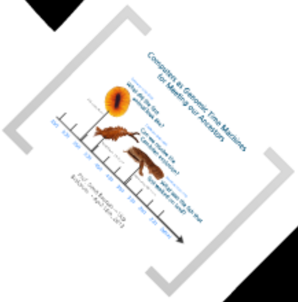
With challenging phylogenetic problems, the difficulty is to make the needle bigger without enlarging the haystack!


With challenging phylogenetic problems, the difficulty is to make the needle bigger without enlarging the haystack!

more sequences
are not enough



phylogenetic vs.
non-phylogenetic signal


Computers as Genomic Time Machines for Meeting our Ancestors

What did the first animal look like?

Can we resolve the Cambrian explosion?

What was the fish that first walked on land?


Latimeria chalumnae

The coelacanth is a large marine fish with fleshy fins that resembles the limbs of terrestrial vertebrates.

phylogenomics can resolve
the evolution of animals

acknowledgment

Université
de Montréal

- Herve Philippe
- Henner Brinkmann
- Beatrice Roure

Starting grant SFRD-12/04 (FSR)

fnrs
LA LIBERTÉ DE CHERCHER

Université
de Liège

references

acknowledgment

Université de Montréal

- Herve Philippe
- Henner Brinkmann
- Beatrice Roure

Starting grant SFRD-12/04 (FSR)

fnrs

Université de Liège

Amemiya CT, Alfoldi J et al. (2013)
The African coelacanth genome provides insights into tetrapod evolution.
Nature 496, 311-316.

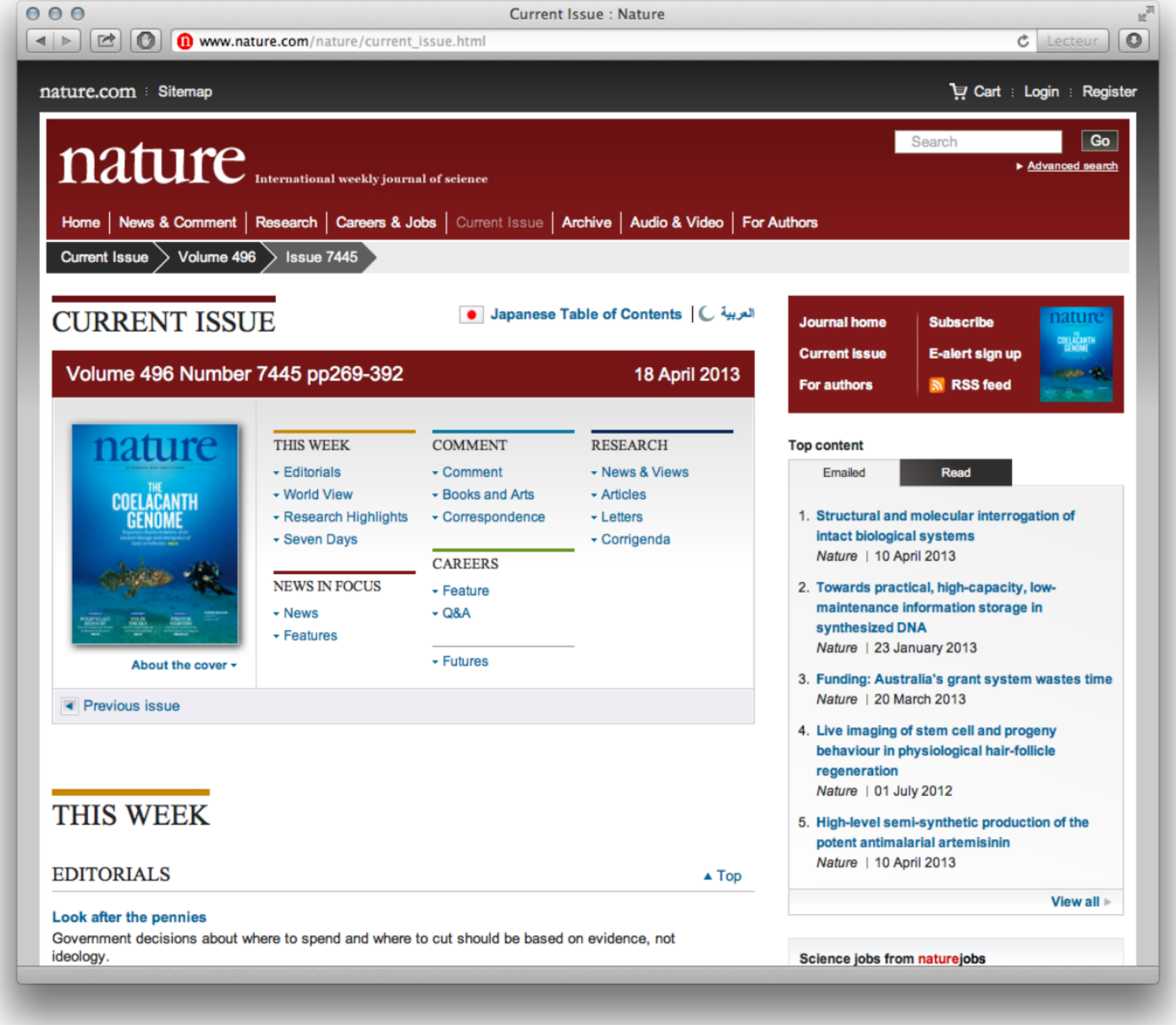
Roure B, Baurain D, Philippe H (2013)
Impact of missing data on phylogenies inferred from empirical phylogenomic data sets.
Mol Biol Evol 30, 197-214.

Philippe H, Brinkmann H, Lavrov DV, Littlewood DT, Manuel M, Wrheide G, Baurain D (2011)
Resolving difficult phylogenetic questions: why more sequences are not enough.
PLoS Biol 9, e1000602.

Baurain D, Philippe H (2010)
Current approaches to phylogenomic reconstruction.
In Evolutionary Genomics and Systems Biology, ed Caetano-Anolles G (Wiley-Blackwell, Hoboken, N.J.), pp 17-41.

Baurain D, Brinkmann H, Philippe H (2007)
Lack of resolution in the animal phylogeny: closely spaced cladogeneses or undetected systematic errors?
Mol Biol Evol 24, 6-9.

All papers are available on http://orbi.ulg.ac.be/



CURRENT ISSUE

Volume 496 Number 7445 pp269-392 18 April 2013

THIS WEEK

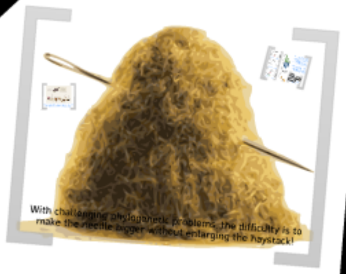
EDITORIALS

Look after the pennies

Government decisions about where to spend and where to cut should be based on evidence, not ideology.

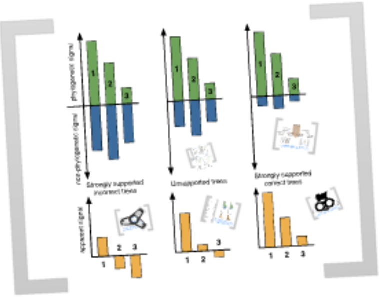
1. **Structural and molecular interrogation of intact biological systems**
   Nature | 10 April 2013
2. **Towards practical, high-capacity, low-maintenance information storage in synthesized DNA**
   Nature | 23 January 2013
3. **Funding: Australia's grant system wastes time**
   Nature | 20 March 2013
4. **Live imaging of stem cell and progeny behaviour in physiological hair-follicle regeneration**
   Nature | 01 July 2012
5. **High-level semi-synthetic production of the potent antimalarial artemisinin**
   Nature | 10 April 2013

Science jobs from naturejobs



With challenging phylogenetic problems, the difficulty is to make the needle bigger without enlarging the haystack!

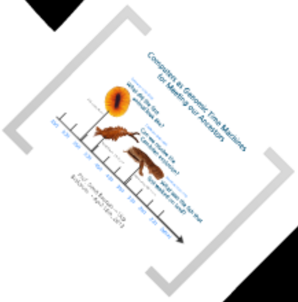
more sequences
are not enough



phylogenetic vs.
non-phylogenetic signal

The coelacanth is a large marine fish with fleshy fins that resembles the limbs of terrestrial vertebrates.

phylogenomics can resolve
the evolution of animals



Computers as Genomic Time Machines for Meeting our Ancestors

What did the first animal look like?

Can we resolve the Cambrian explosion?

What was the fish that first walked on land?