USING GENOME-LEVEL PROCESSES TO SOLVE TRICKY NODES IN PHYLOGENETICS

@JordiPaps

USING GENOME-LEVEL PROCESSES TO SOLVE TRICKY NODES IN PHYLOGENETICS

@JordiPaps

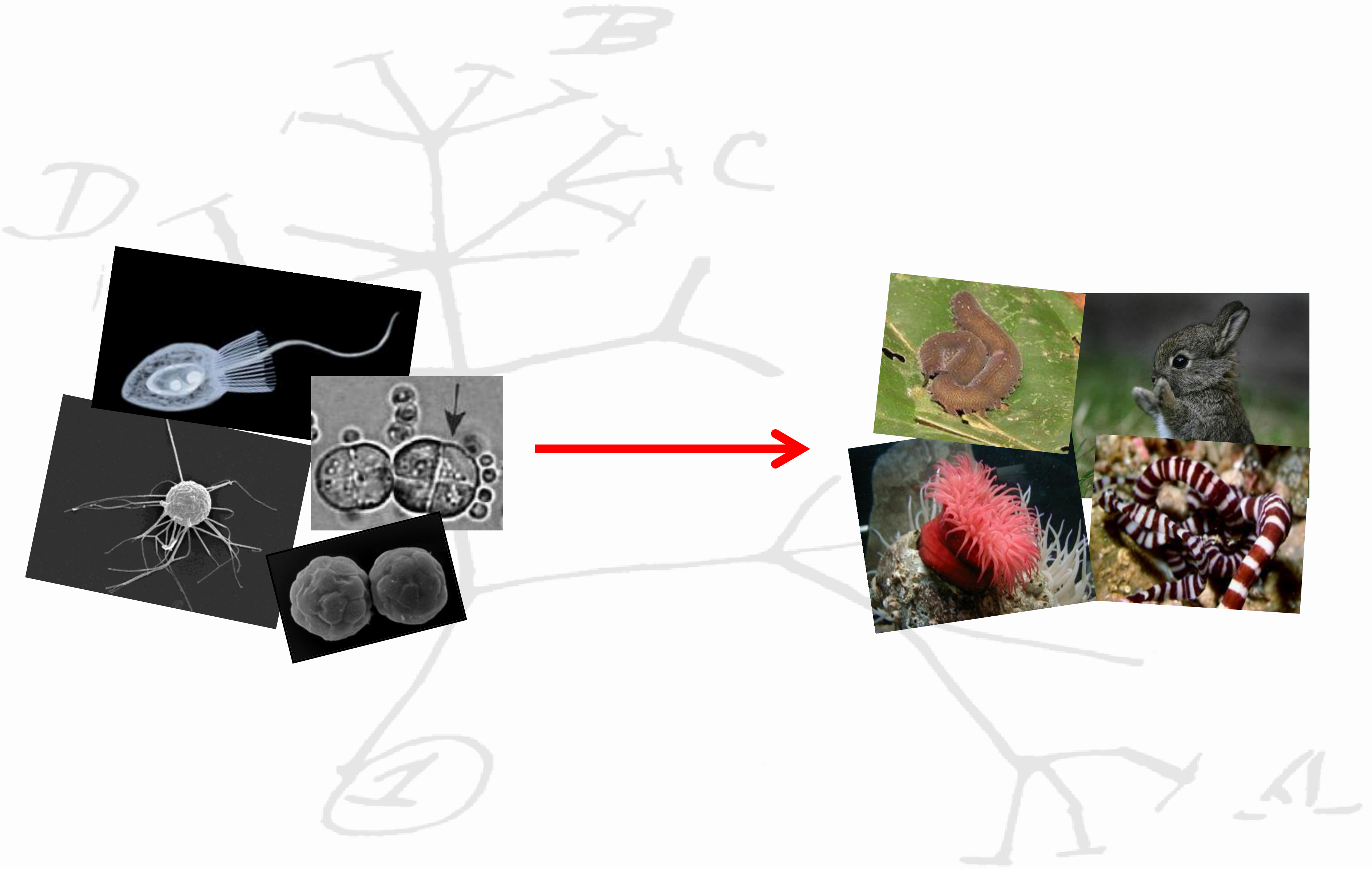
USING GENOME-LEVEL PROCESSES TO SOLVE ~~TRICKY NODES~~ IN PHYLOGENETICS

LBA

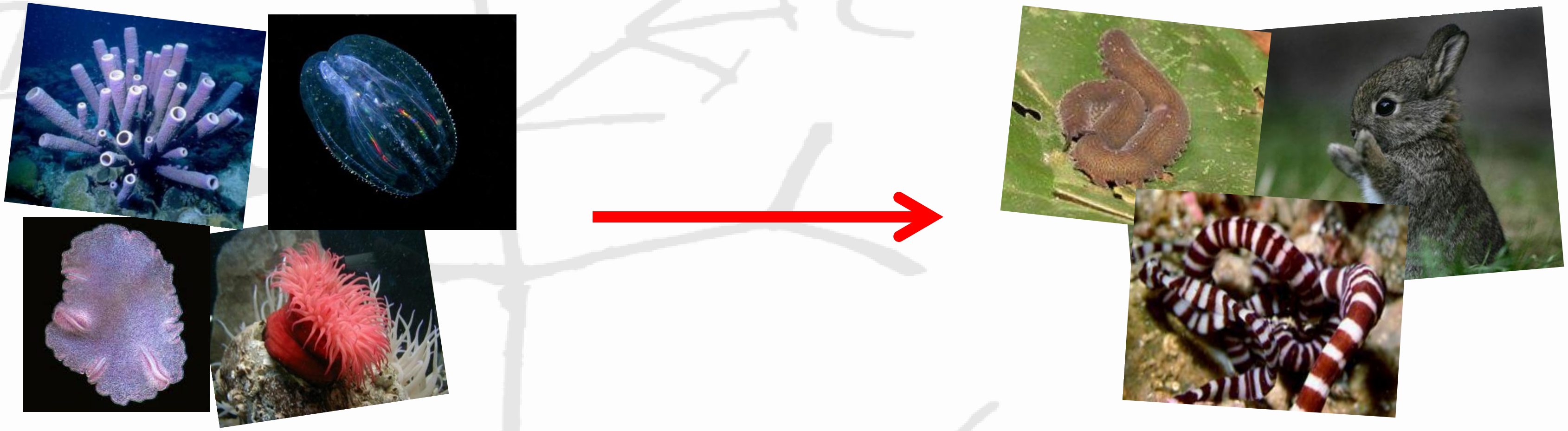
@JordiPaps

Animals

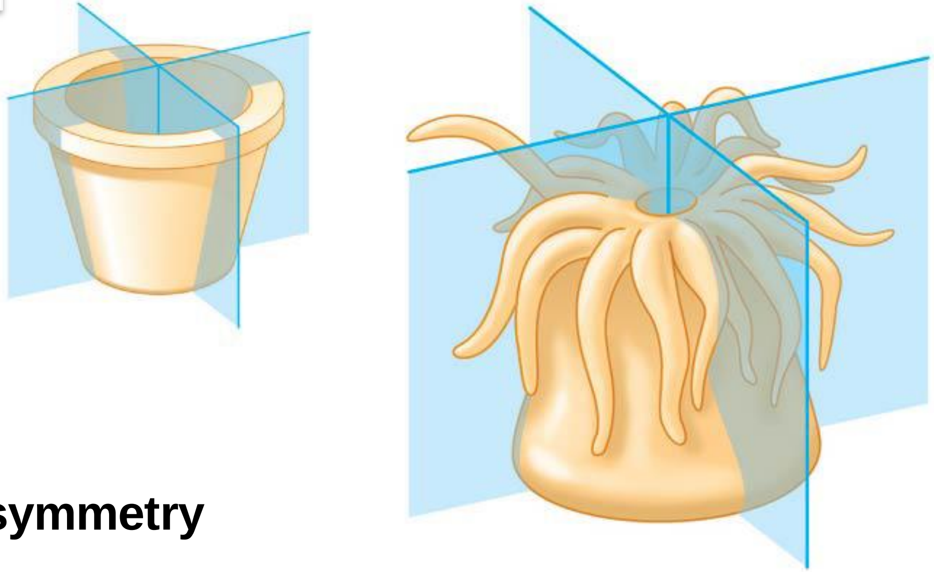
From unicellular to multicellular...



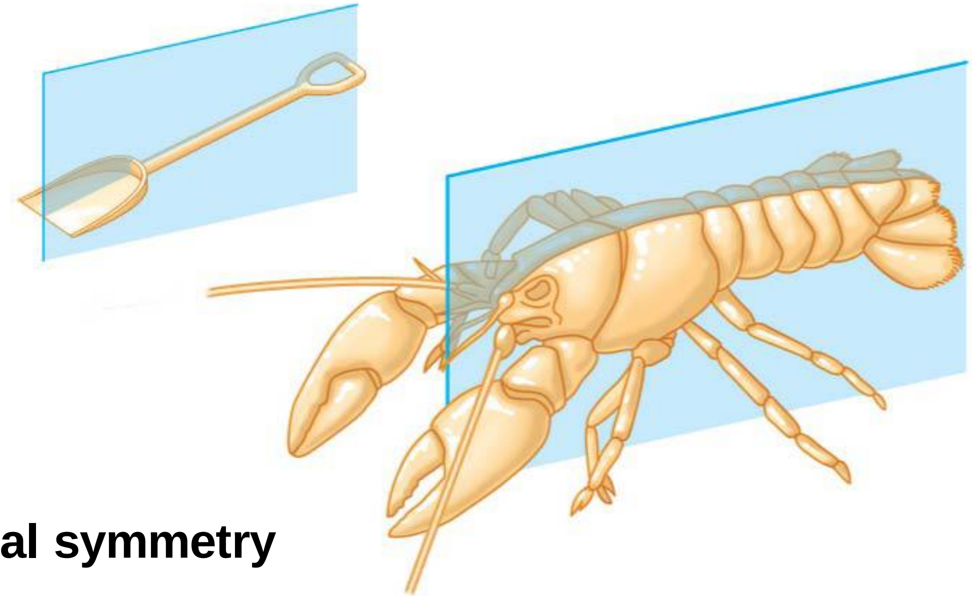
From radial symmetry to bilateral



Bilaterian innovations

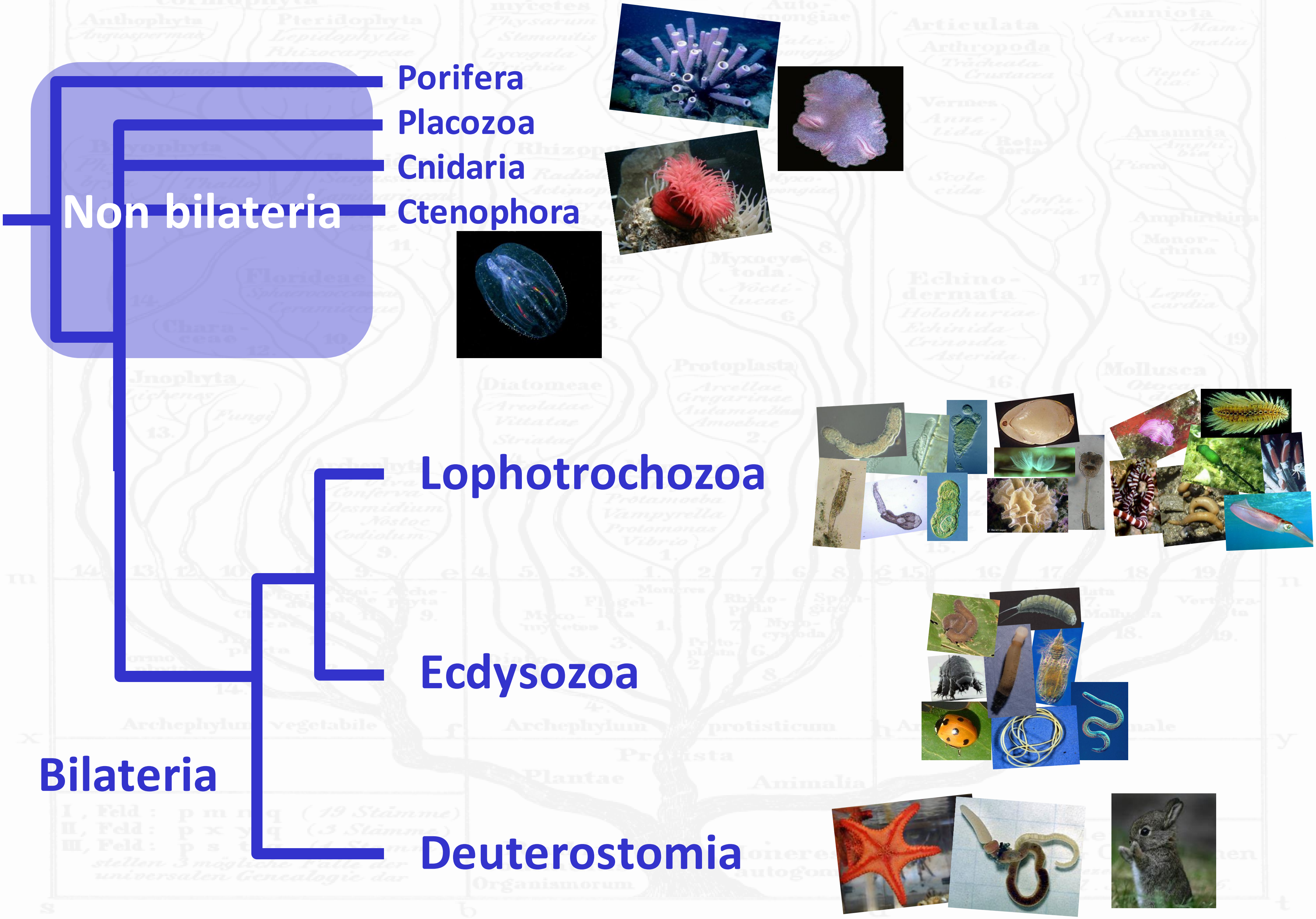


(a) Radial symmetry



(b) Bilateral symmetry

The New Animal Phylogeny



Porifera

Placozoa

Cnidaria

Ctenophora

Non bilateria

Lophotrochozoa

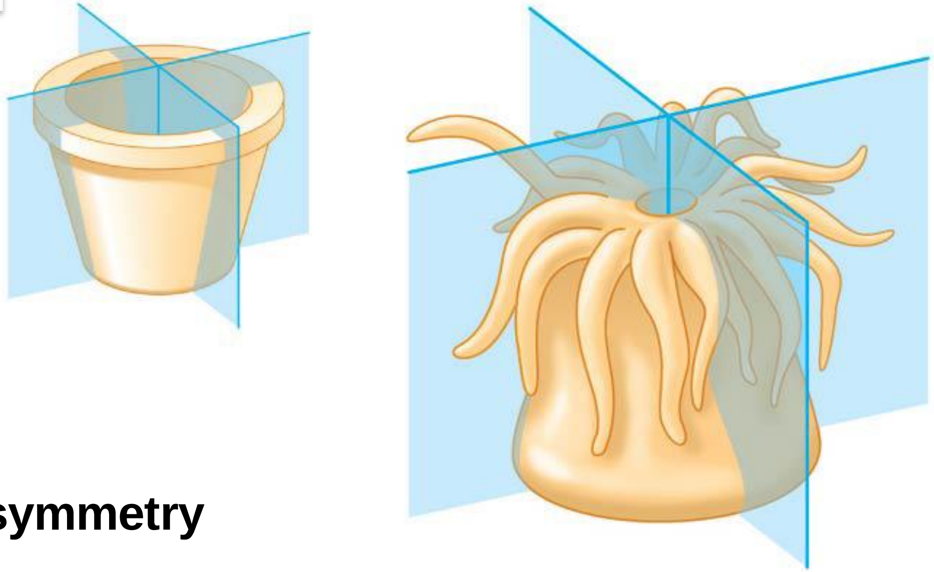
Ecdysozoa

Deuterostomia

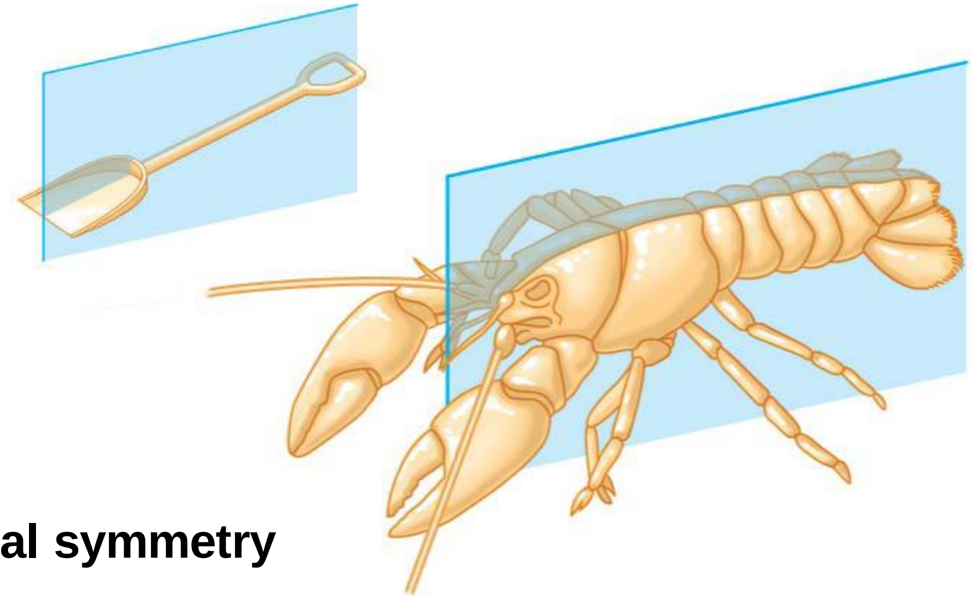
Bilateria

Bilaterian innovations

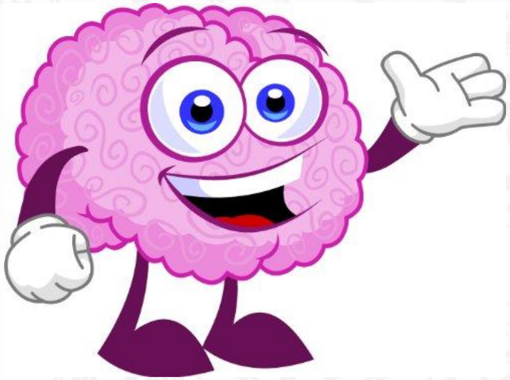
- Symmetry
- Cephalization
- Mesoderm



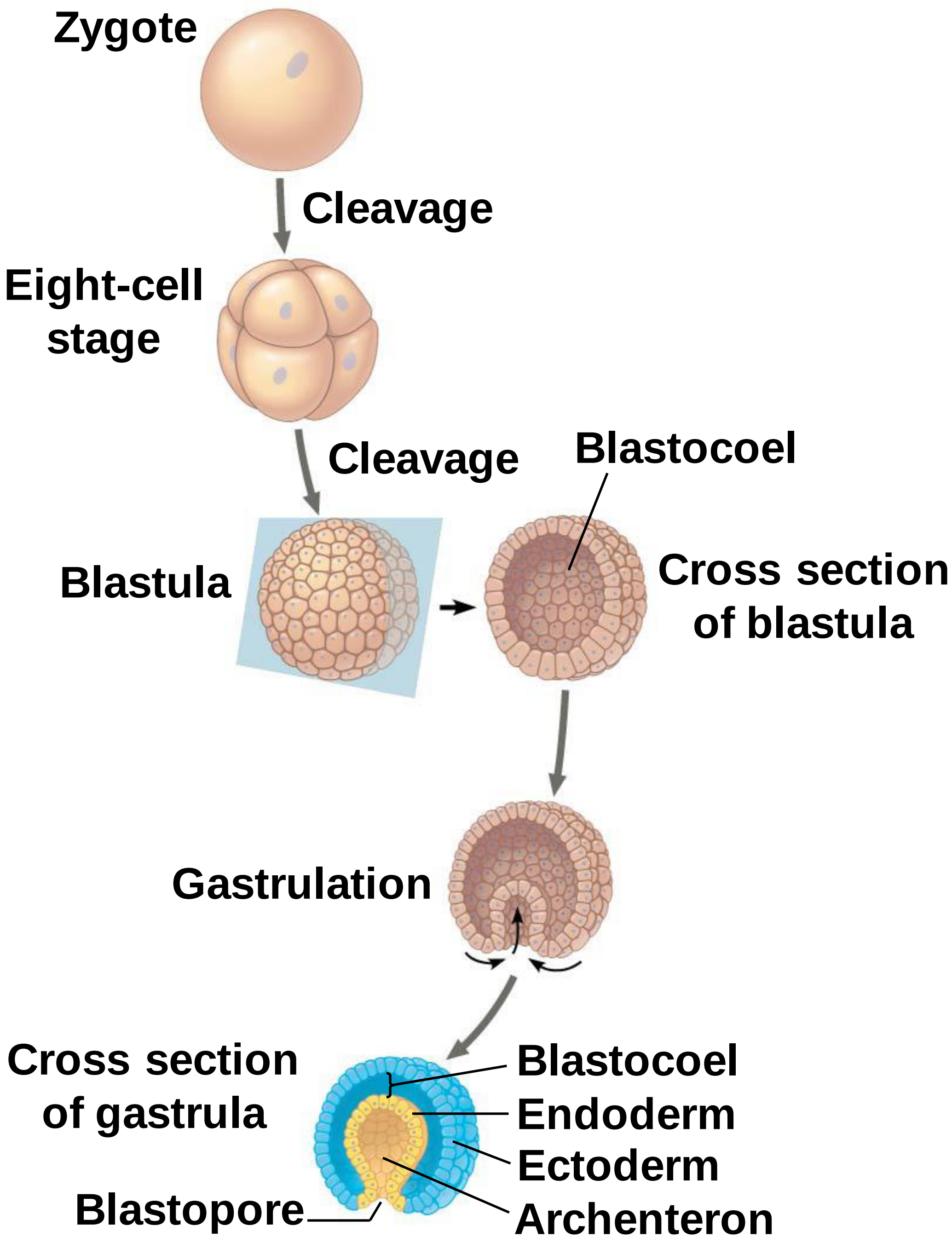
(a) Radial symmetry



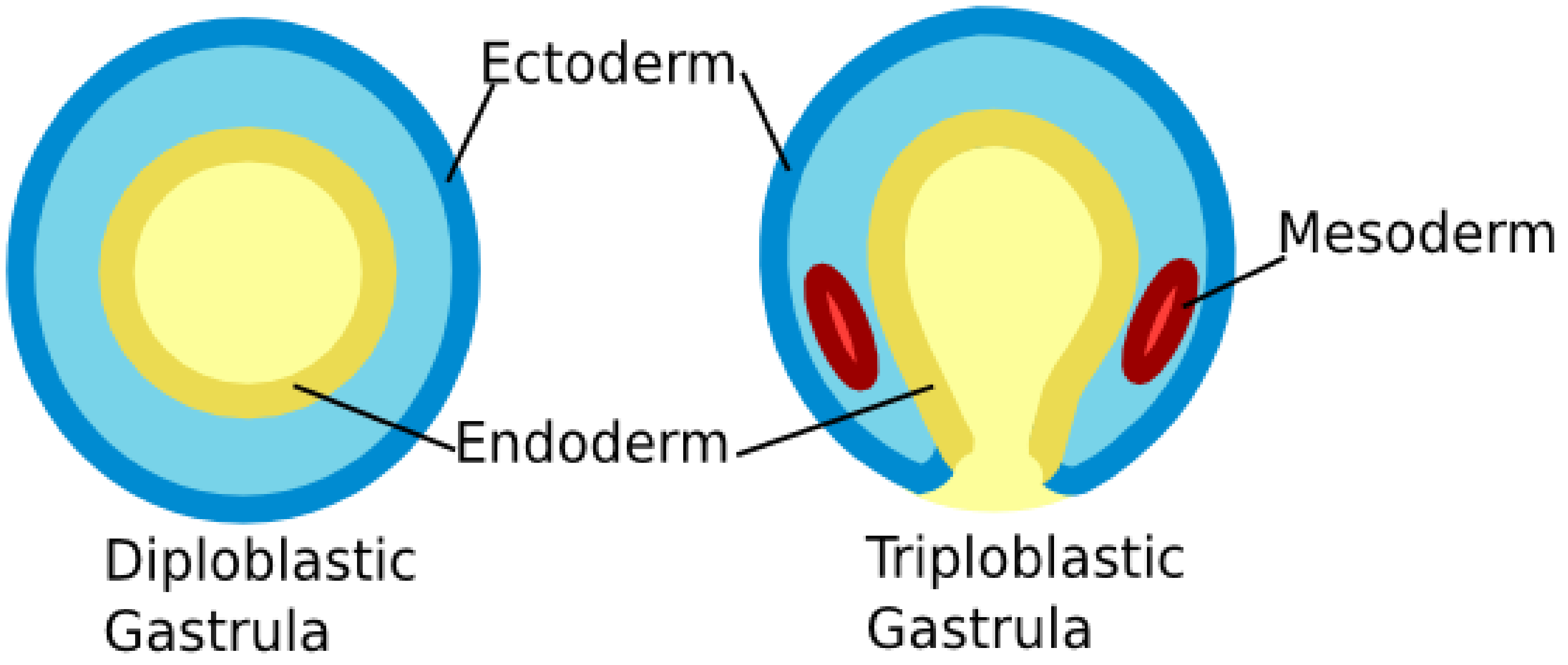
(b) Bilateral symmetry



Mesoderm


Zygote
Cleavage
Eight-cell stage
Cleavage
Blastocoel
Blastula
Cross section of blastula
Gastrulation
Cross section of gastrula
Blastocoel
Endoderm
Ectoderm
Archenteron
Blastopore

Mesoderm


Ectoderm
Mesoderm
Endoderm
Diploblastic Gastrula
Triploblastic Gastrula

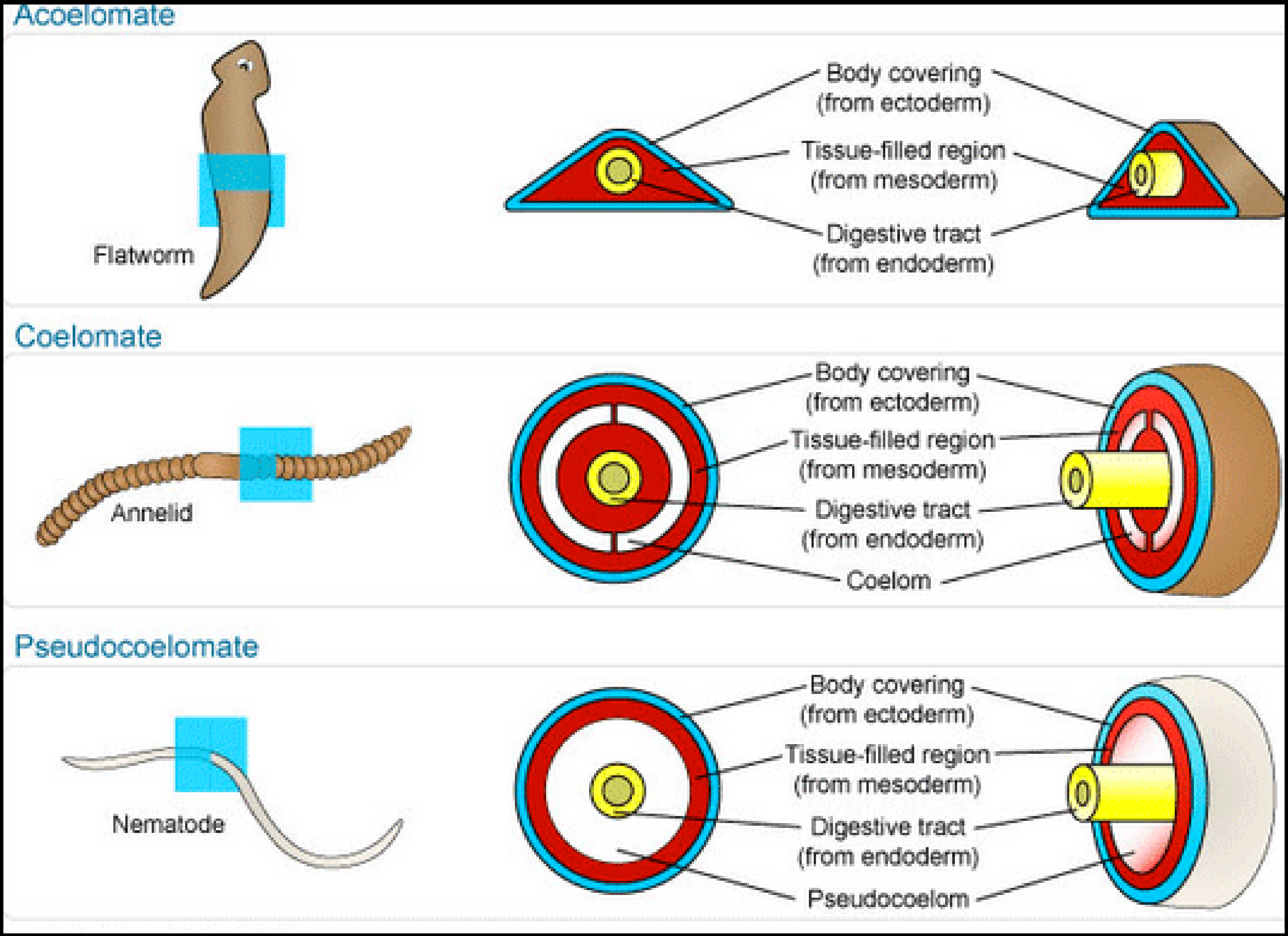
Embryonic layers

- **Endoderm: digestive** tract (i.e. guts) and organs derived from it (i.e. **lungs**) as well as **secretory organs** (liver, pancreas, thyroid, thymus…).

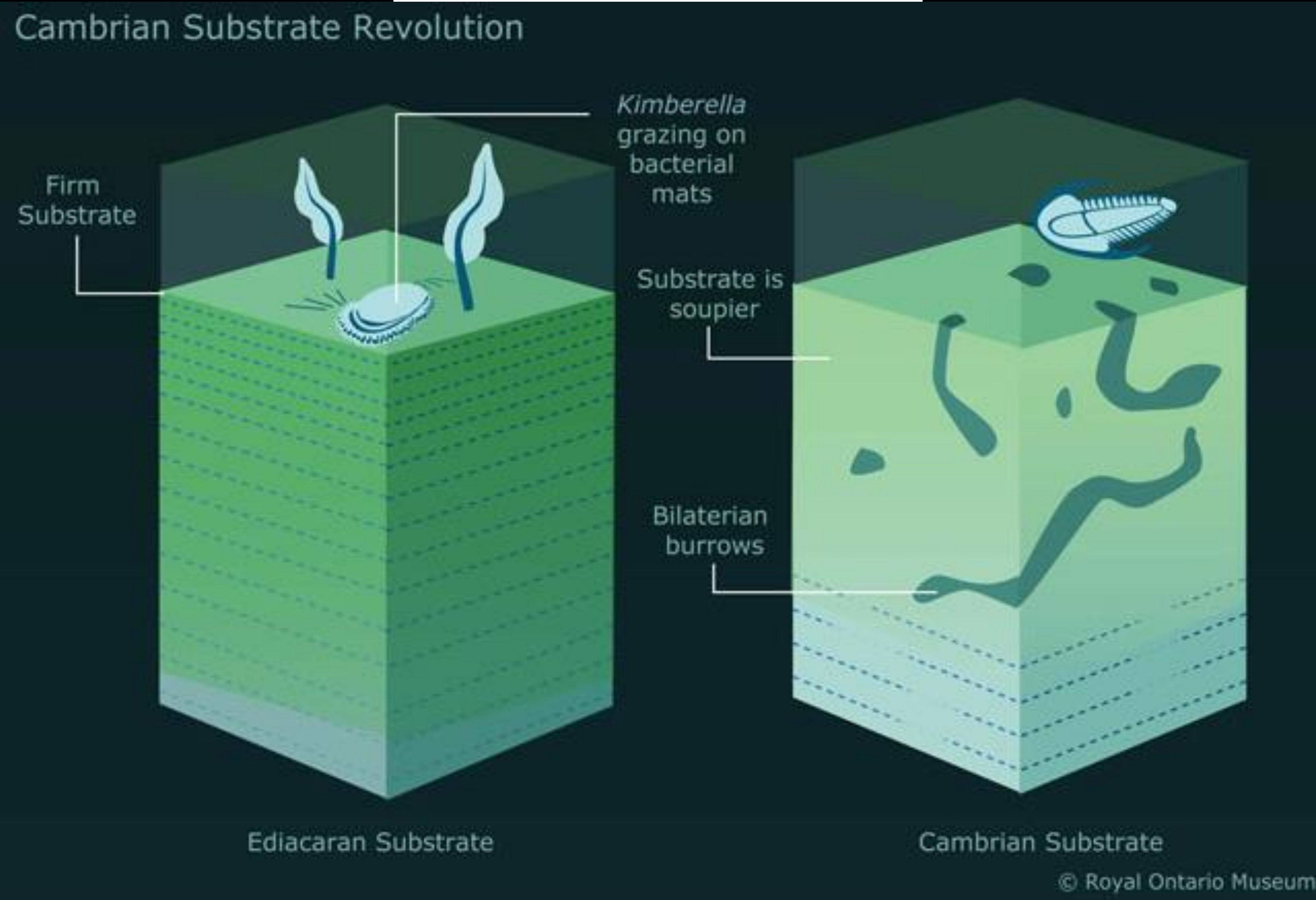
- **Ectoderm: epidermis** and **nervous system**, as well as hair, cornea, tooth enamel, etc.

- **Mesoderm**: **coelom**, **muscles**, **bone/cartilage**, **circulatory**, **spleen**, **kidney**, **reproductive**, etc.

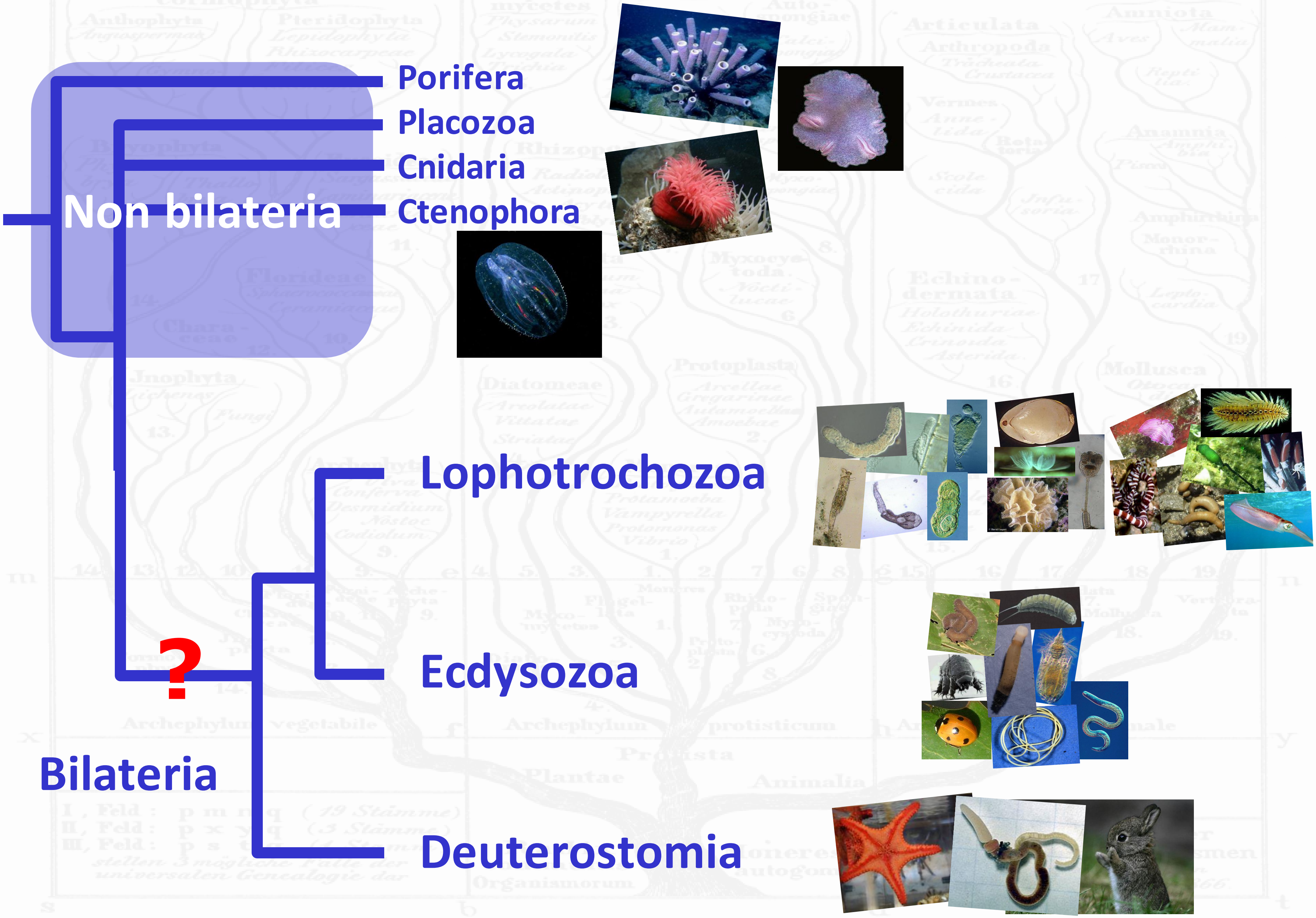
Body cavities



Triggers: Ecological



The New Animal Phylogeny



Non bilateria

Porifera

Placozoa

Cnidaria

Ctenophora

Lophotrochozoa

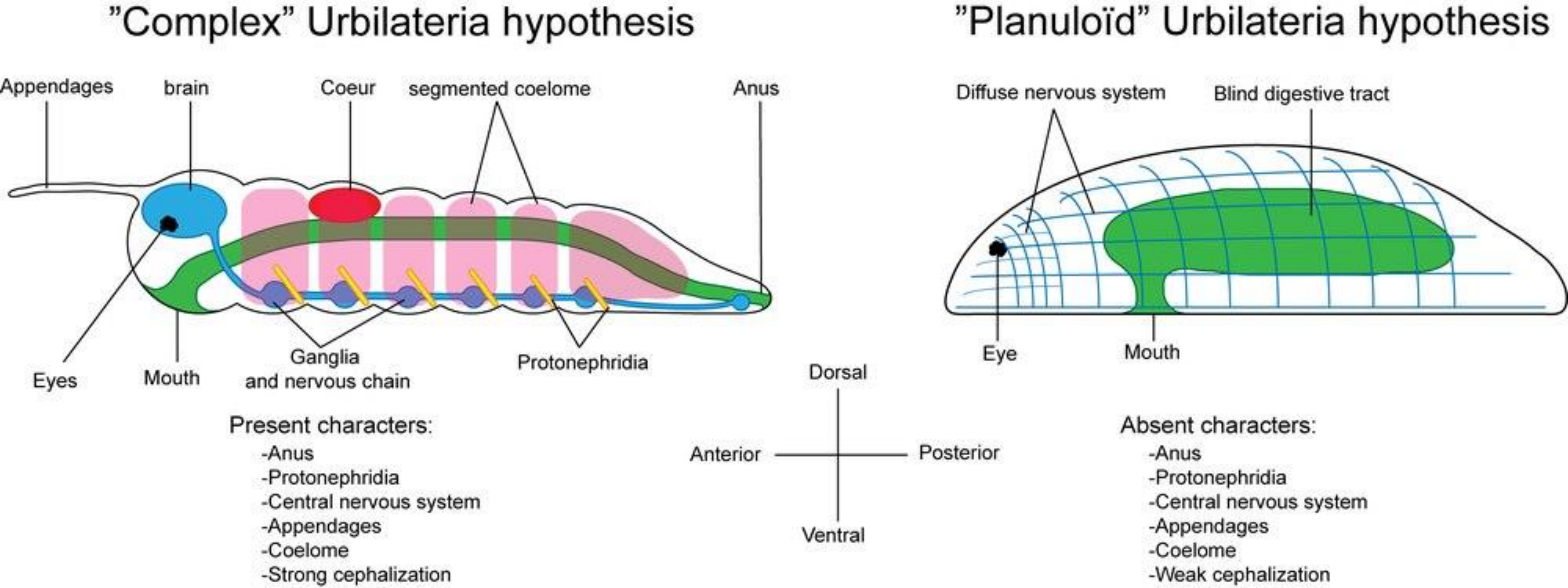
?

Ecdysozoa

Bilateria

Deuterostomia

First bilaterian

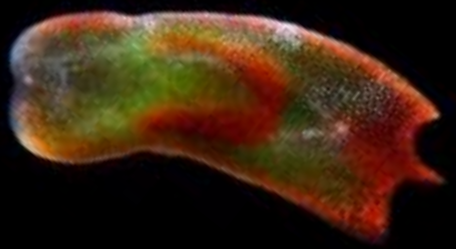


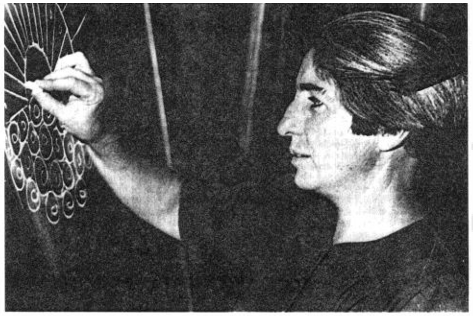
"Complex" Urbilateria hypothesis

"Planuloïd" Urbilateria hypothesis

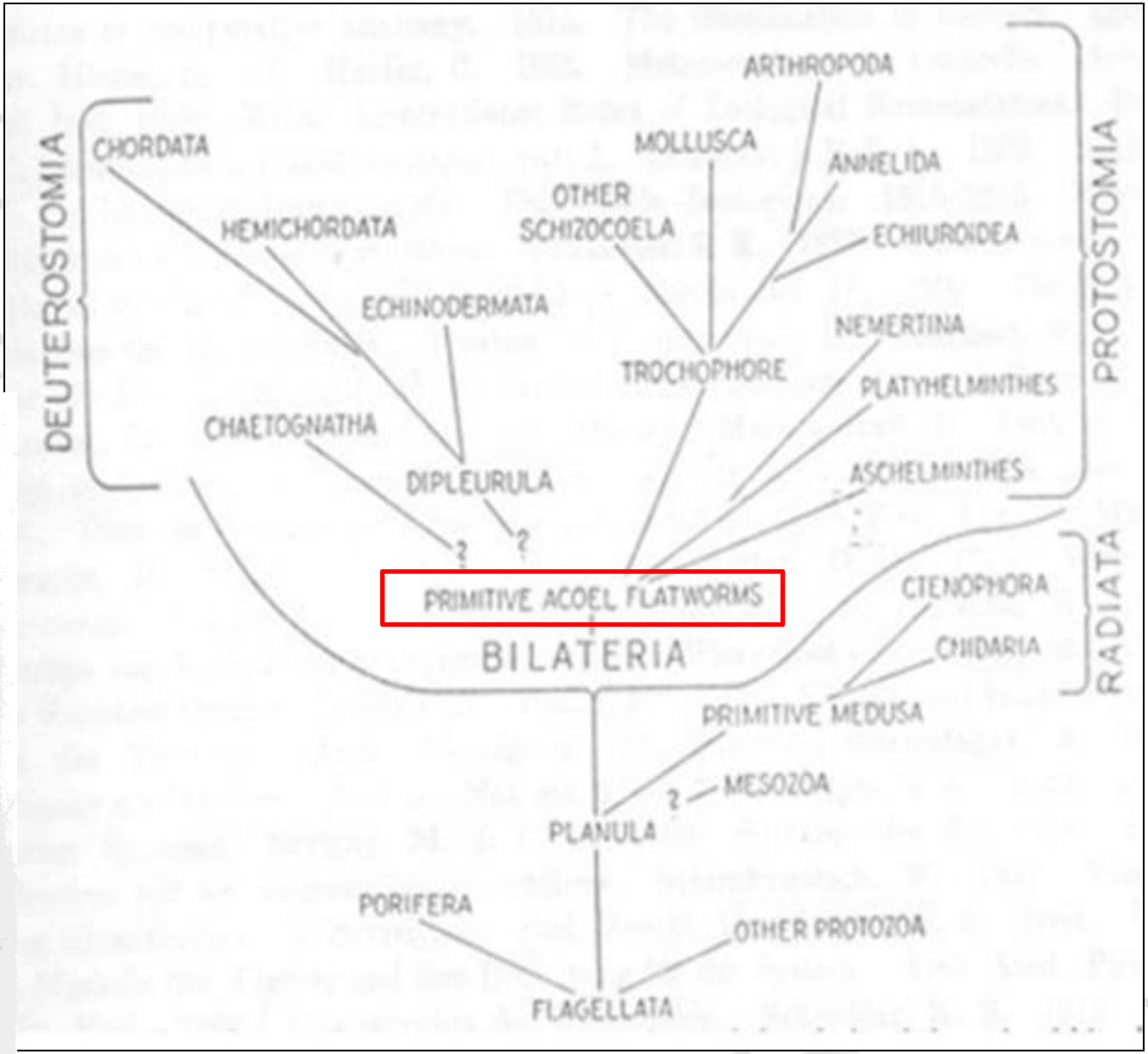
Present characters:
- -Anus
- -Protonephridia
- -Central nervous system
- -Appendages
- -Coelome
- -Strong cephalization

Absent characters:
- -Anus
- -Protonephridia
- -Central nervous system
- -Appendages
- -Coelome
- -Weak cephalization





Libbie Hyman



***The Invertebrates*, 1940-1967**

Cestoda

Polycladida

Polycladida

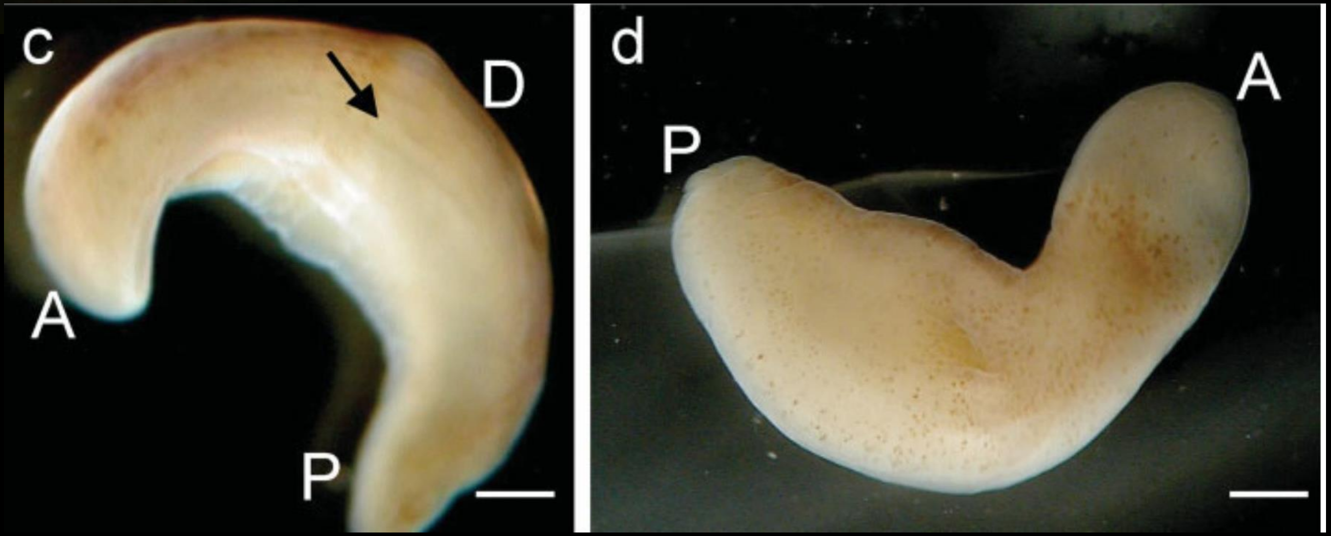
Tricladida

Acoela

Nemertodermatida

Xenoturbellida

The Swedish clog



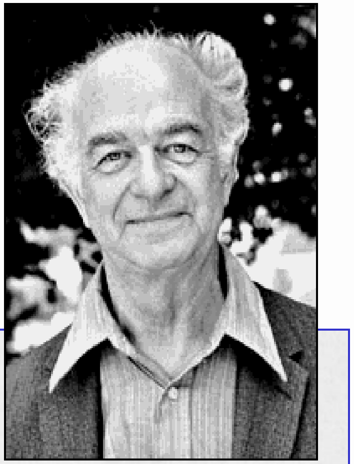
Acoels, nemertodermatids, and xenoturbellids

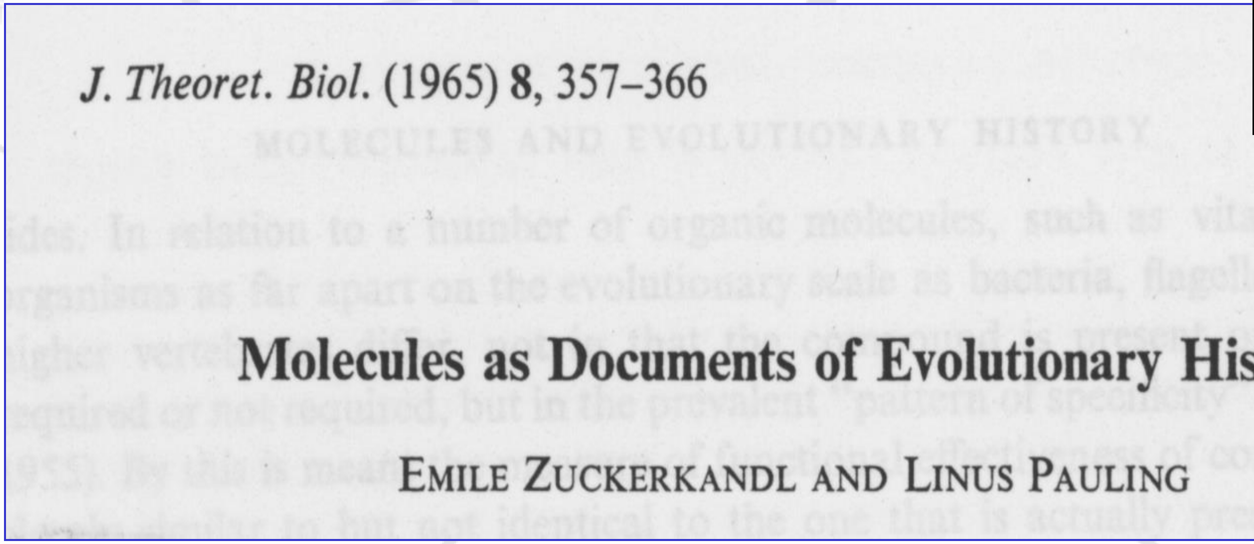
Acoela + Nemertodermatida = Acoelomorpha

Xenoturbellida +Acoelomorpha = Xenacelomorpha

THE MOLECULAR ERA

Linus Pauling





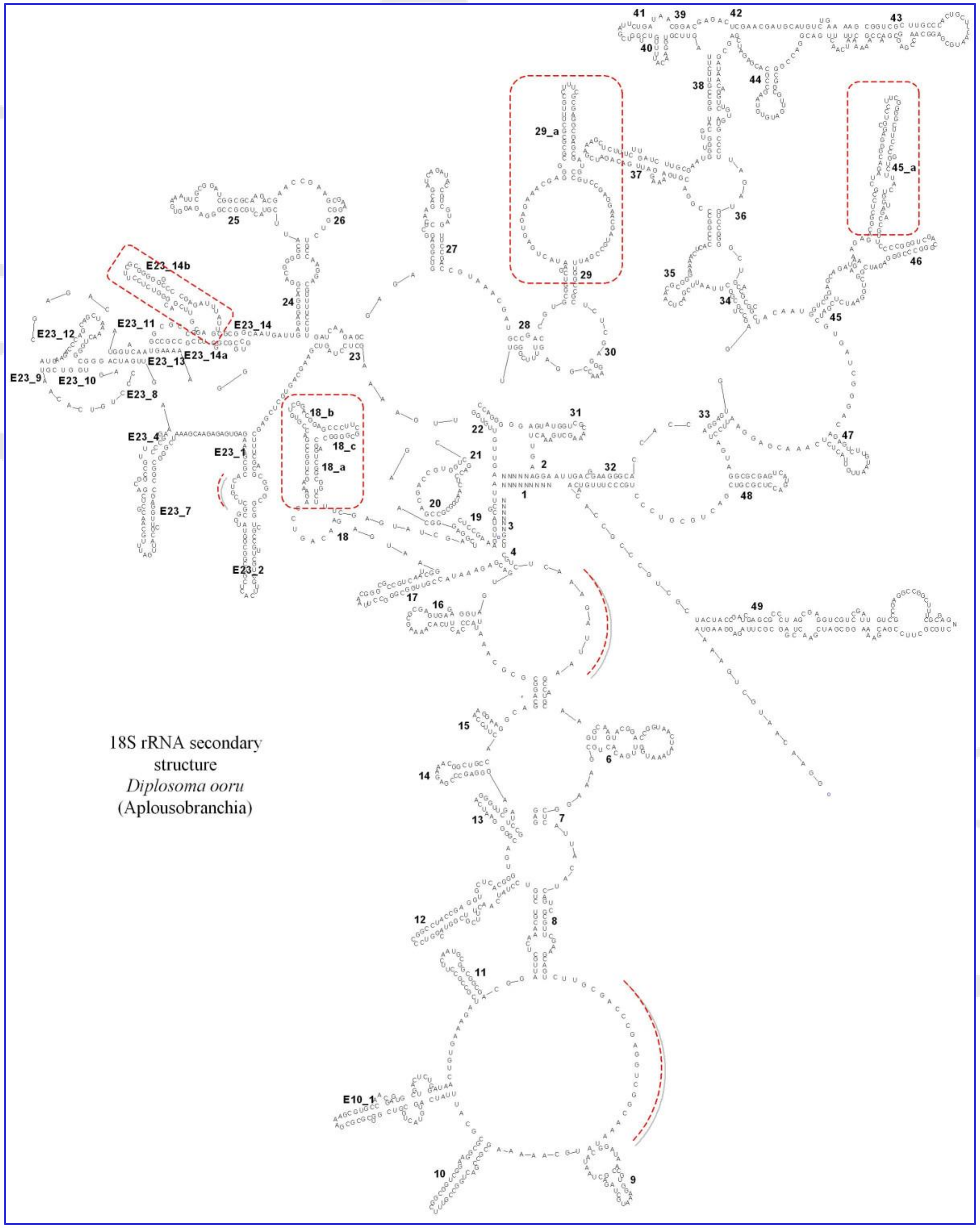
J. Theoret. Biol. (1965) **8**, 357–366

Molecules as Documents of Evolutionary History

EMILE ZUCKERKANDL AND LINUS PAULING

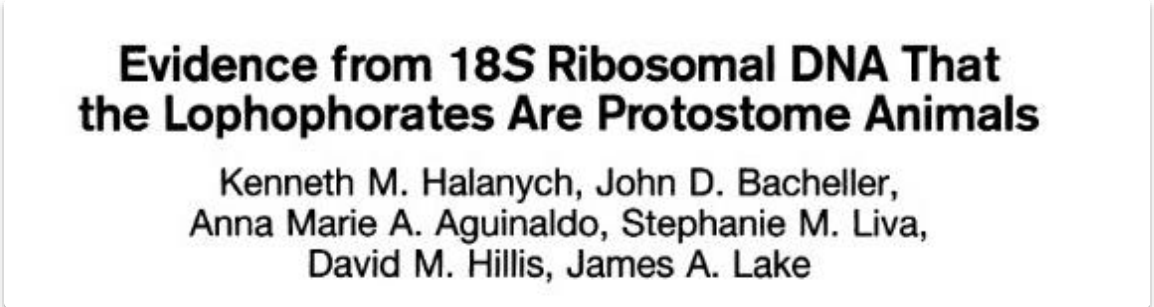
Zuckerland and Pauling 1965

Molecular: 18S



18S rRNA secondary structure
Diplosoma ooru
(Aplousobranchia)

Molecular: 18S



Evidence from 18*S* Ribosomal DNA That the Lophophorates Are Protostome Animals

Kenneth M. Halanych, John D. Bacheller,
Anna Marie A. Aguinaldo, Stephanie M. Liva,
David M. Hillis, James A. Lake

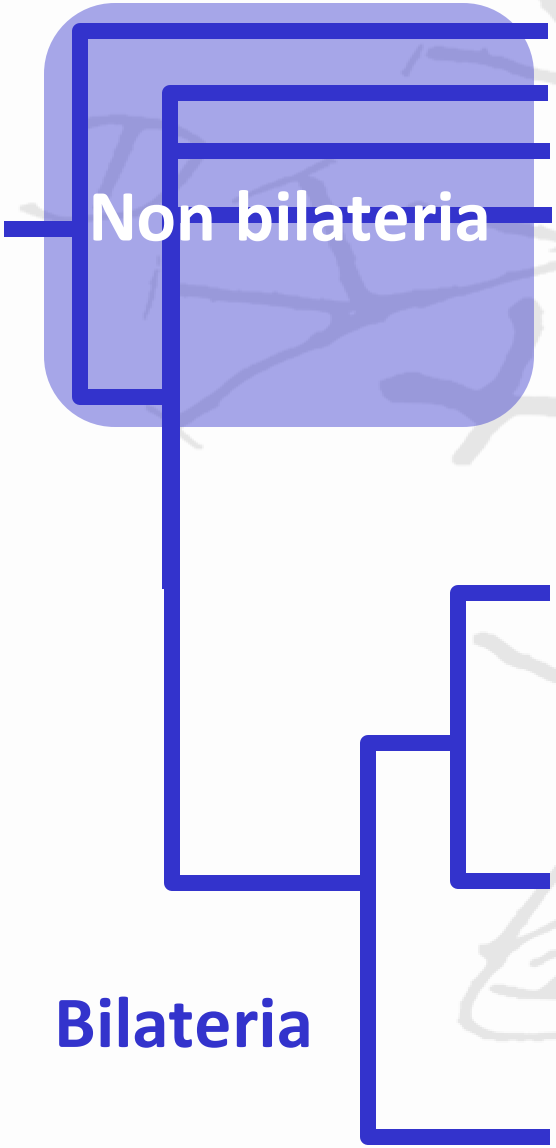
Halanych et al, 1995

Evidence for a clade of nematodes, arthropods and other moulting animals

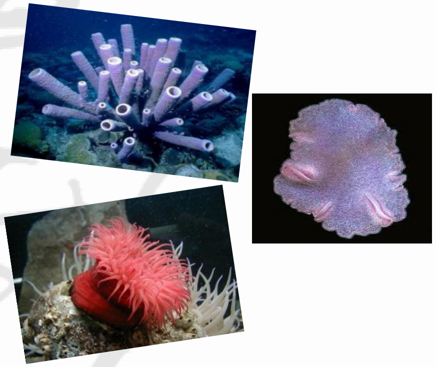
**Anna Marie A. Aguinaldo[*], James M. Turbeville[†],
Lawrence S. Linford[*], Maria C. Rivera[*],
James R. Garey[‡], Rudolf A. Raff[§] & James A. Lake[*]**

Aguinaldo et al, 1997

The New Animal Phylogeny



Porifera
Placozoa
Cnidaria
Ctenophora

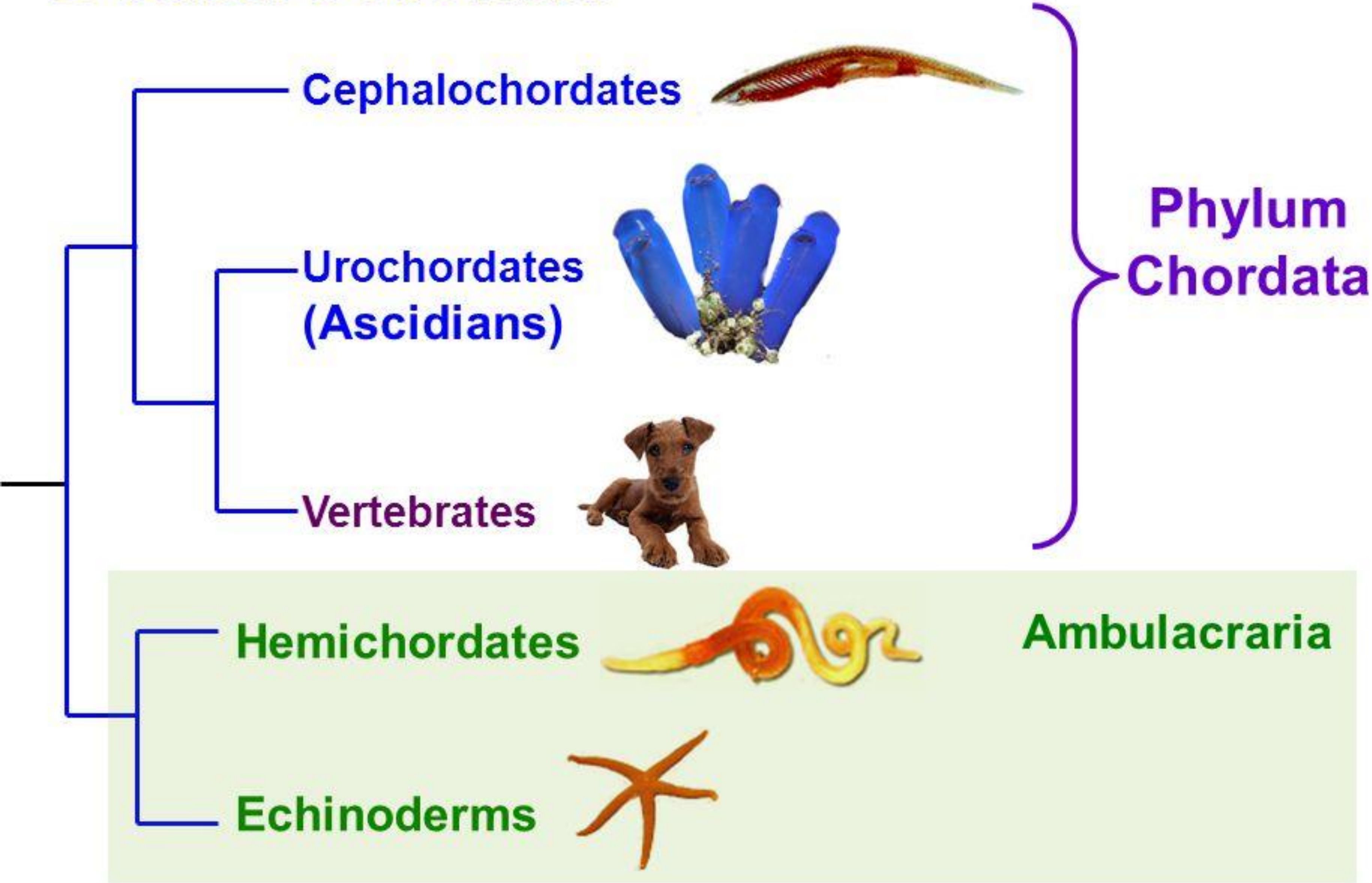


Lophotrochozoa

Ecdysozoa

Deuterostomia

Deuterostomia



Cephalochordates

Urochordates
(Ascidians)

Vertebrates

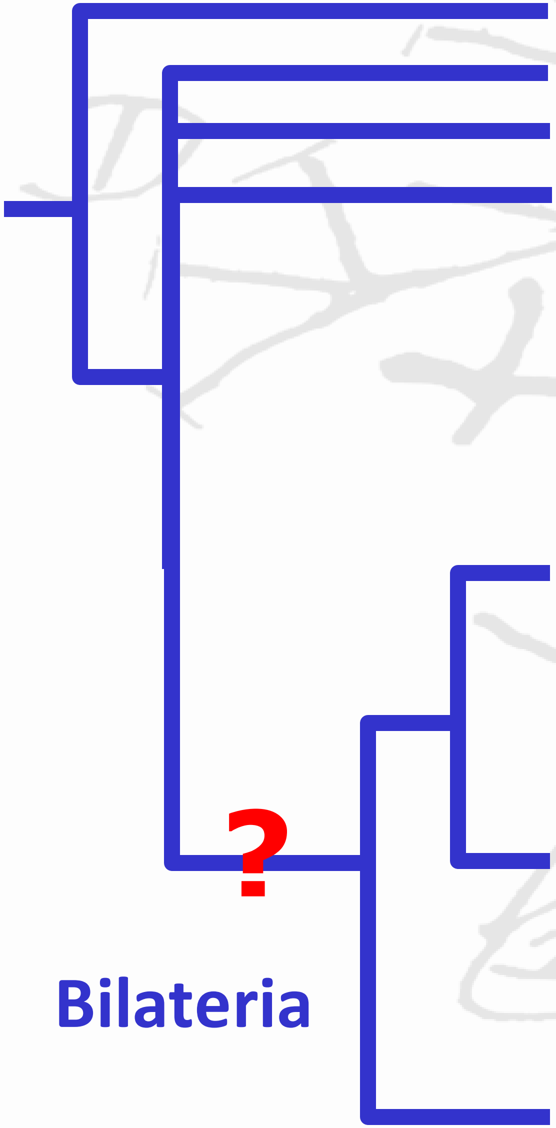
Phylum
Chordata

Hemichordates

Echinoderms

Ambulacraria

The New Animal Phylogeny



Porifera

Placozoa

Cnidaria

Ctenophora

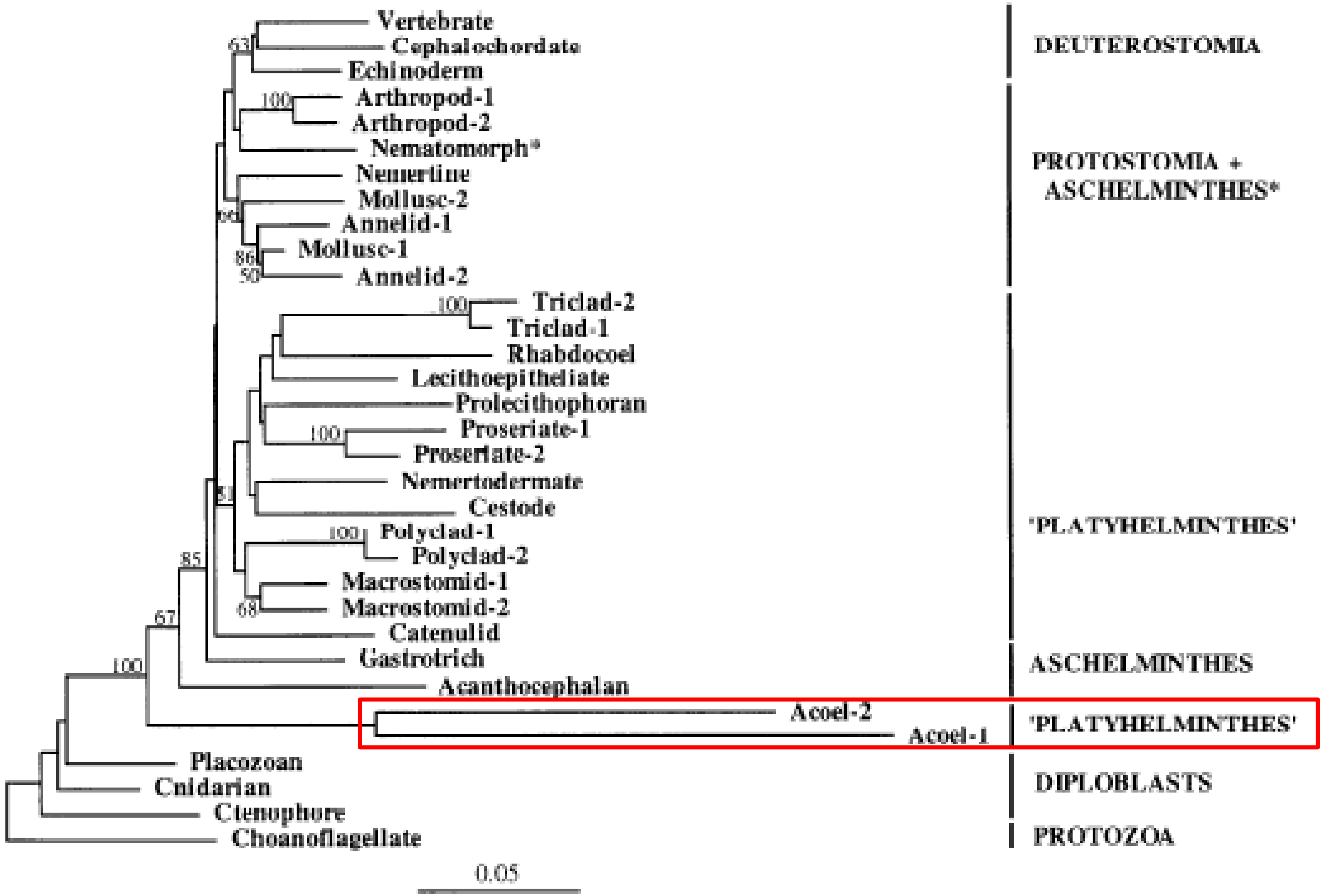
Lophotrochozoa

Ecdysozoa

Bilateria

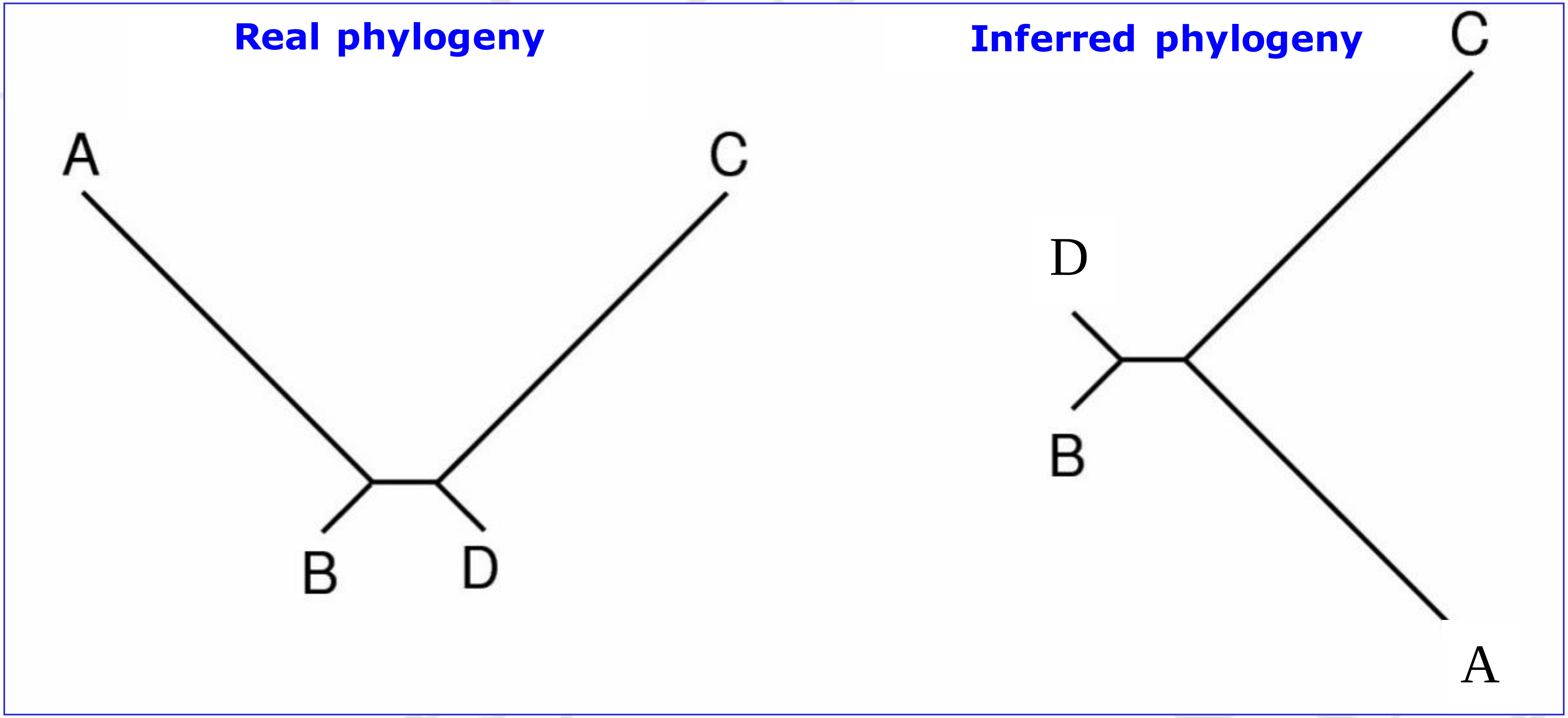
Deuterostomia

Long Branch Atraction

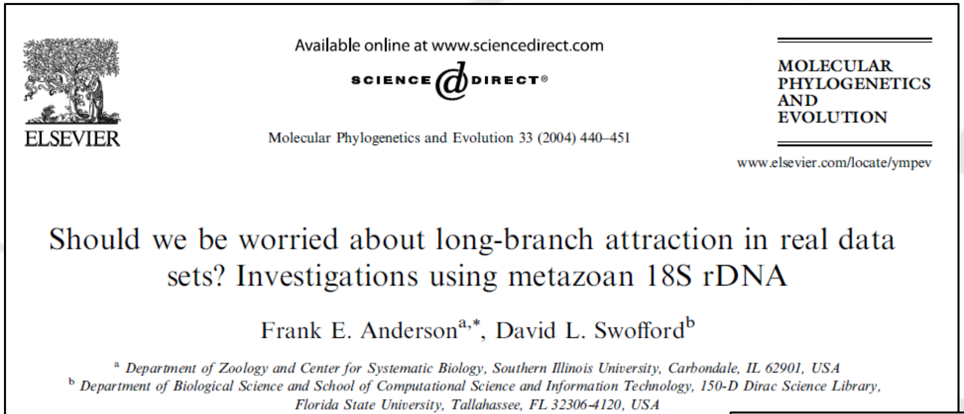


Carranza et al, 1997

Long Branch Atraction



Avoiding LBA



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Molecular Phylogenetics and Evolution 33 (2004) 440–451

MOLECULAR PHYLOGENETICS AND EVOLUTION

www.elsevier.com/locate/ympev

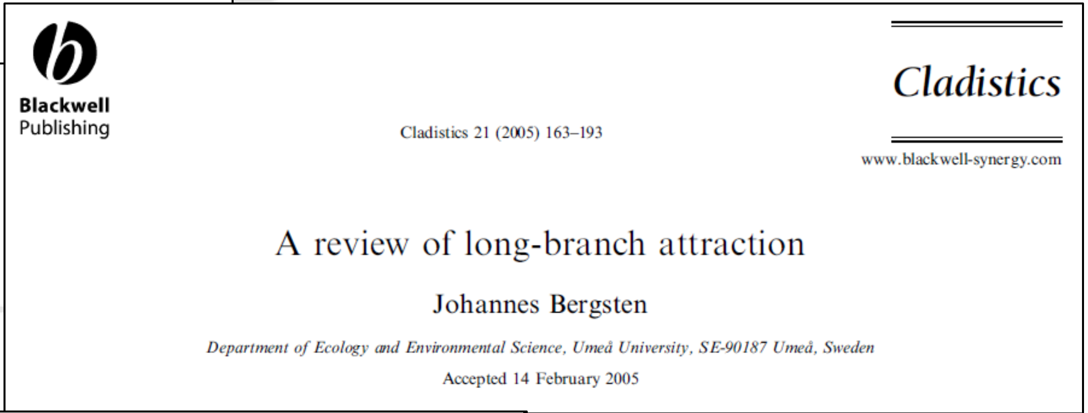
Should we be worried about long-branch attraction in real data sets? Investigations using metazoan 18S rDNA

Frank E. Anderson[a,*], David L. Swofford[b]

[a] Department of Zoology and Center for Systematic Biology, Southern Illinois University, Carbondale, IL 62901, USA

[b] Department of Biological Science and School of Computational Science and Information Technology, 150-D Dirac Science Library, Florida State University, Tallahassee, FL 32306-4120, USA

Received 4 February 2004; revised 1 June 2004



Blackwell Publishing

Cladistics

Cladistics 21 (2005) 163–193

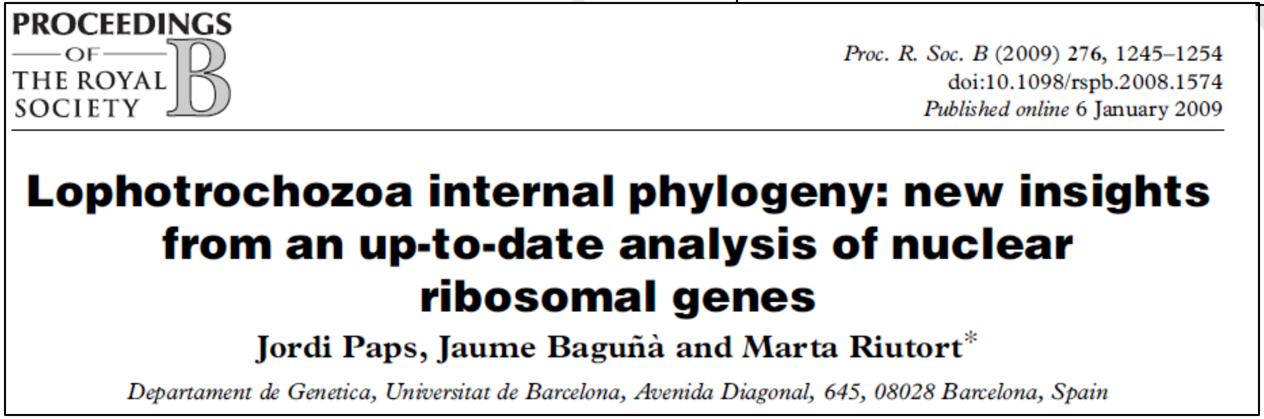
www.blackwell-synergy.com

A review of long-branch attraction

Johannes Bergsten

Department of Ecology and Environmental Science, Umeå University, SE-90187 Umeå, Sweden

Accepted 14 February 2005



PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2009) 276, 1245–1254
doi:10.1098/rspb.2008.1574

Lophotrochozoa internal phylogeny: new insights from an up-to-date analysis of nuclear ribosomal genes

Jordi Paps, Jaume Baguñà and Marta Riutort*

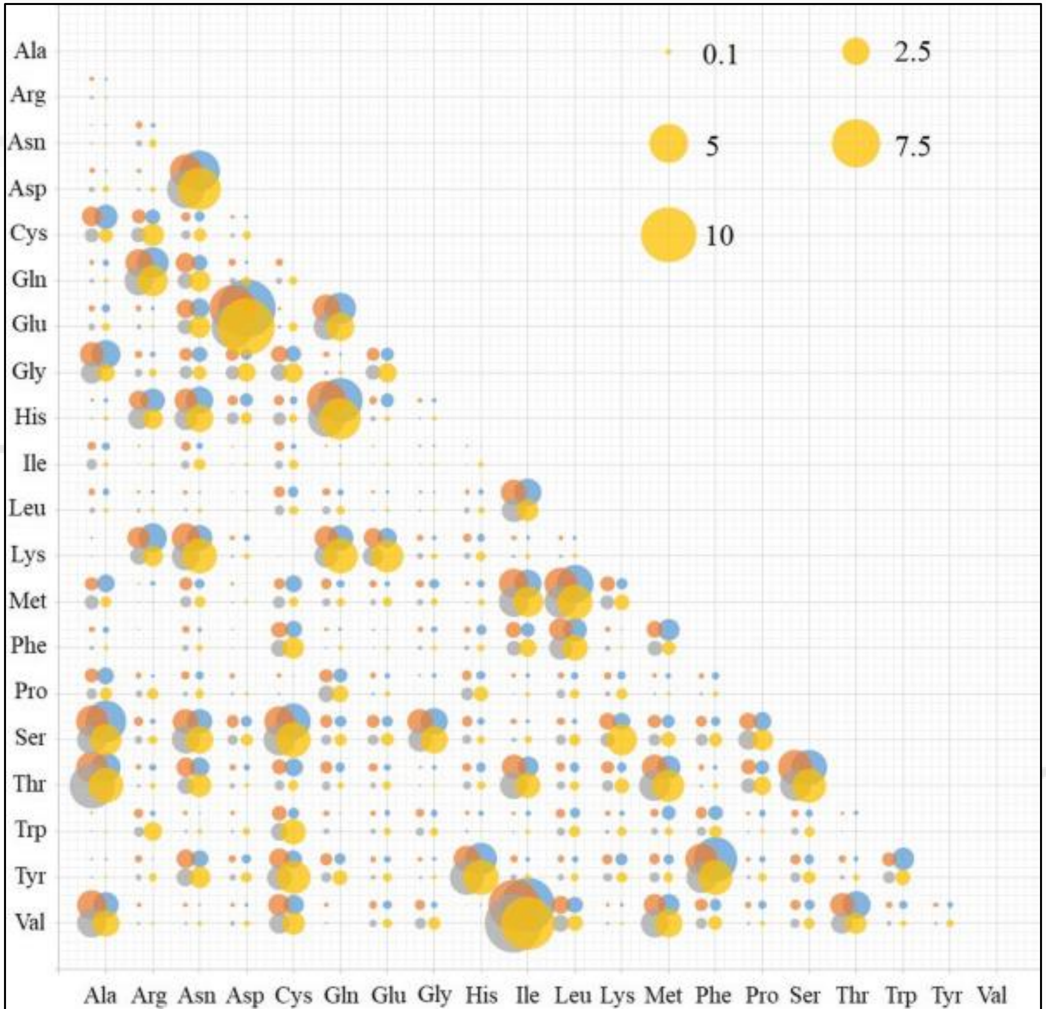
Departament de Genetica, Universitat de Barcelona, Avenida Diagonal, 645, 08028 Barcelona, Spain

Avoiding LBA

1. Use methods less sensitive to **LBA**, such as **ML** or **BI**
   - Use models less sensitive to **LBA**
   - In the case of AA and BI, use CAT over LG

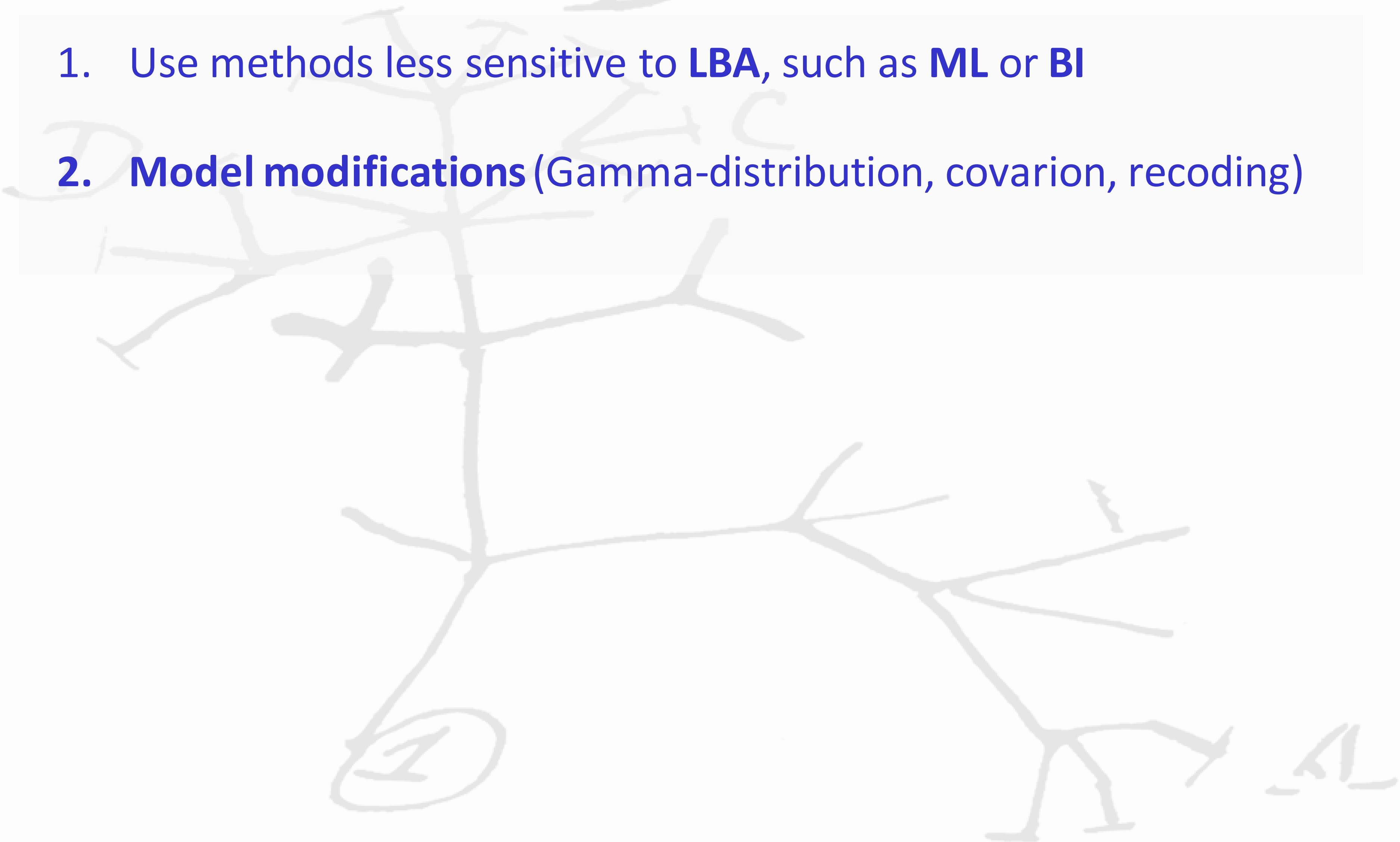
GTR

$$\mathbf{Q} = \begin{bmatrix} \cdot & \alpha\pi_T & \beta\pi_C & \gamma\pi_G \\ \alpha\pi_A & \cdot & \rho\pi_C & \sigma\pi_G \\ \beta\pi_A & \rho\pi_T & \cdot & \tau\pi_G \\ \gamma\pi_A & \sigma\pi_T & \tau\pi_C & \cdot \end{bmatrix}$$

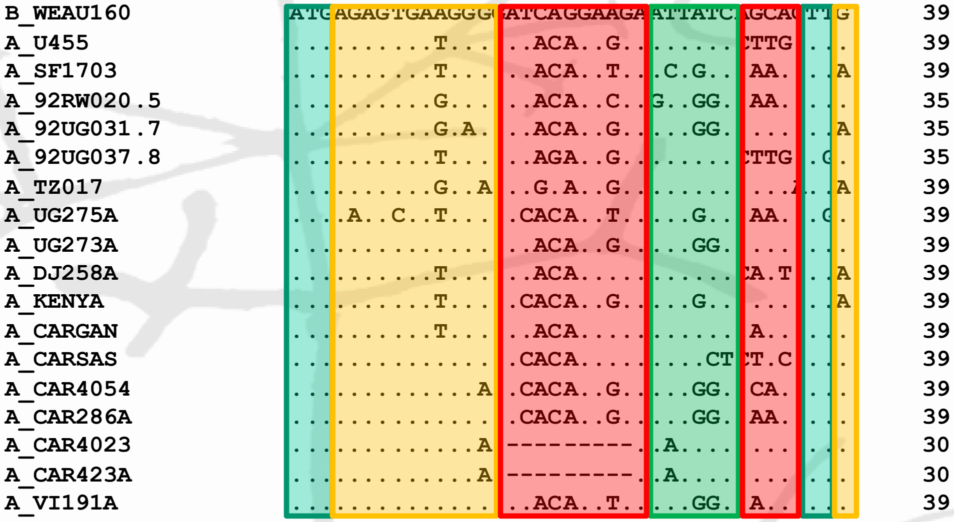


Avoiding LBA

1. Use methods less sensitive to **LBA**, such as **ML** or **BI**
2. **Model modifications** (Gamma-distribution, covarion, recoding)

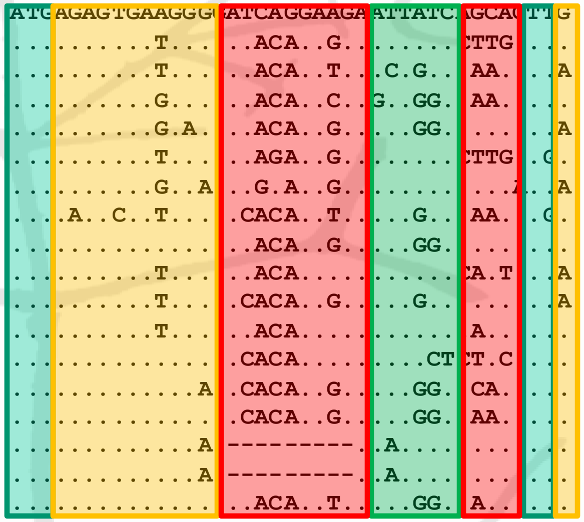


Gamma distribution



```
B_WEAU160     ATGAGAGTGAAGGGGATCAGGAAGAATTATCAGCAGTTG   39
A_U455        ..........T......ACA..G........CTTG....   39
A_SF1703      ..........T......ACA..T...C.G...AA....A   39
A_92RW020.5   ..........G......ACA..C..G..GG..AA.....   35
A_92UG031.7   ..........G.A....ACA..G.....GG........A   35
A_92UG037.8   ..........T......AGA..G........CTTG..G.   35
A_TZ017       ..........G..A...G.A..G............A..A   39
A_UG275A      ....A..C..T.....CACA..T.....G...AA...G.   39
A_UG273A      .................ACA..G.....GG.........   39
A_DJ258A      ..........T......ACA...........CA.T...A   39
A_KENYA       ..........T.....CACA..G.....G.........A   39
A_CARGAN      ..........T......ACA............A......   39
A_CARSAS      ................CACA.........CTCT.C....   39
A_CAR4054     .............A..CACA..G.....GG..CA.....   39
A_CAR286A     ................CACA..G.....GG..AA.....   39
A_CAR4023     .............A.---------..A............   30
A_CAR423A     .............A.---------..A............   30
A_VI191A      .................ACA..T.....GG..A......   39
```

Gamma distribution



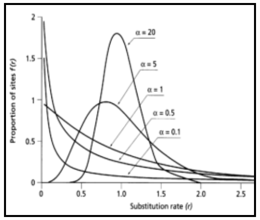
```
B_WEAU160    ATGAGAGTGAAGGGGATCAGGAAGAATTATCAGCAGTTG   39
A_U455       ...........T.....ACA..G.........CTTG....   39
A_SF1703     ...........T.....ACA..T...C.G...AA.....A   39
A_92RW020.5  ...........G.....ACA..C..G..GG..AA......   35
A_92UG031.7  ...........G.A....ACA..G.....GG.........A   35
A_92UG037.8  ...........T.....AGA..G.........CTTG..G.   35
A_TZ017      ...........G..A...G.A..G..............A..A   39
A_UG275A     ....A..C..T......CACA..T.....G....AA...G.   39
A_UG273A     ...............ACA..G.....GG...........   39
A_DJ258A     ...........T.....ACA.............CA.T...A   39
A_KENYA      ...........T.....CACA..G.....G..........A   39
A_CARGAN     ...........T.....ACA...............A.....   39
A_CARSAS     ................CACA...........CTCT.C....   39
A_CAR4054    .............A..CACA..G.....GG..CA.......   39
A_CAR286A    ................CACA..G.....GG..AA.......   39
A_CAR4023    .............A.---------..A..............   30
A_CAR423A    .............A.---------..A..............   30
A_VI191A     ................ACA..T.....GG..A.........   39
```

GTR

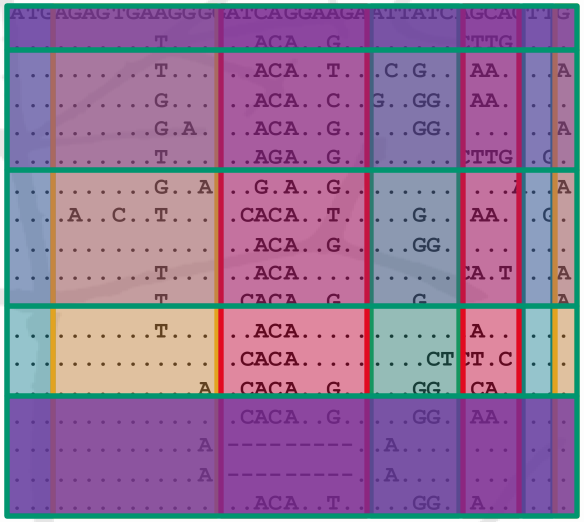
$$
Q = \begin{bmatrix} \cdot & \alpha\pi_T & \beta\pi_C & \gamma\pi_G \\ \alpha\pi_A & \cdot & \rho\pi_C & \sigma\pi_G \\ \beta\pi_A & \rho\pi_T & \cdot & \tau\pi_G \\ \gamma\pi_A & \sigma\pi_T & \tau\pi_C & \cdot \end{bmatrix}
$$

X

Gamma



Heterotachy & Covarion



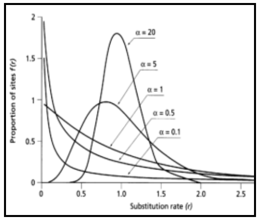
```
B_WEAU160      ATG AGAGTGAAGGGG ATCAGGAAGA ATTATC AGCAG TTG   39
A_U455         ... .......T.... ..ACA..G.. ...... CTTG  ...   39
A_SF1703       ... .......T.... ..ACA..T.. .C.G.. AA.   ..A   39
A_92RW020.5    ... .......G.... ..ACA..C.. G..GG. AA.   ...   35
A_92UG031.7    ... .......G.A.. ..ACA..G.. ...GG. ...   ..A   35
A_92UG037.8    ... .......T.... ..AGA..G.. ...... CTTG  .G.   35
A_TZ017        ... .......G..A. ..G.A..G.. ...... ...A  ..A   39
A_UG275A       ... .A..C..T.... .CACA..T.. ...G.. AA.   .G.   39
A_UG273A       ... ............ ..ACA..G.. ...GG. ...   ...   39
A_DJ258A       ... .......T.... ..ACA..... ...... CA.T  ..A   39
A_KENYA        ... .......T.... .CACA..G.. ...G..       ..A   39
A_CARGAN       ... .......T.... ..ACA..... ...... A..   ...   39
A_CARSAS       ... ............ .CACA..... ....CT CT.C  ...   39
A_CAR4054      ... ..........A. .CACA..G.. ...GG. CA.   ...   39
A_CAR286A      ... ............ .CACA..G.. ...GG. AA.   ...   39
A_CAR4023      ... ..........A. ---------. .A.... ...   ...   30
A_CAR423A      ... ..........A. ---------. .A.... ...   ...   30
A_VI191A       ... ............ ..ACA..T.. ...GG. A..   ...   39
```

GTR

$$Q = \begin{bmatrix} \cdot & \alpha\pi_T & \beta\pi_C & \gamma\pi_G \\ \alpha\pi_A & \cdot & \rho\pi_C & \sigma\pi_G \\ \beta\pi_A & \rho\pi_T & \cdot & \tau\pi_G \\ \gamma\pi_A & \sigma\pi_T & \tau\pi_C & \cdot \end{bmatrix}$$

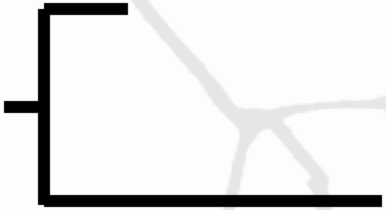
X

Gamma

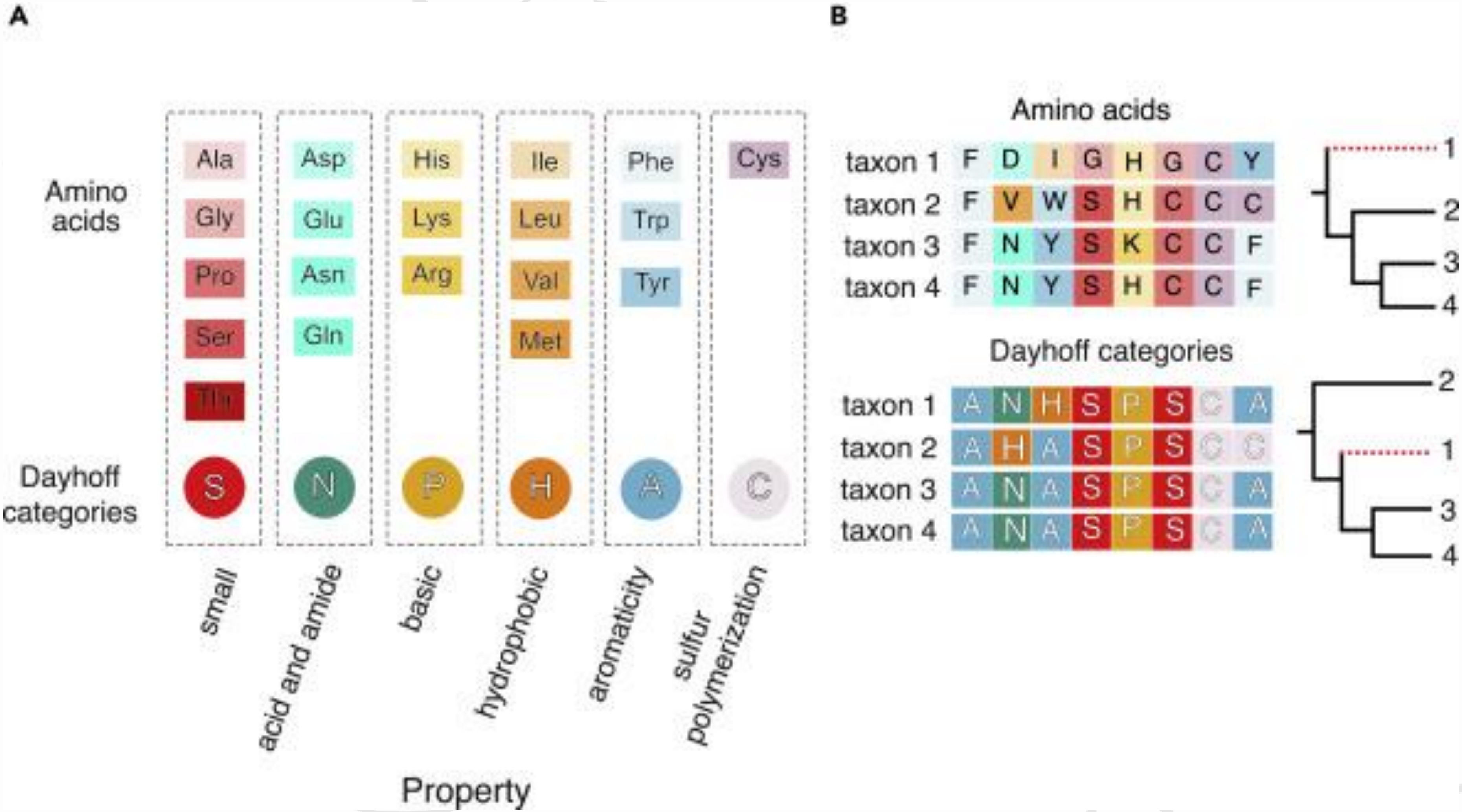


X

Covarion



Recoding



Giacomelli et al, 2022

Recoding

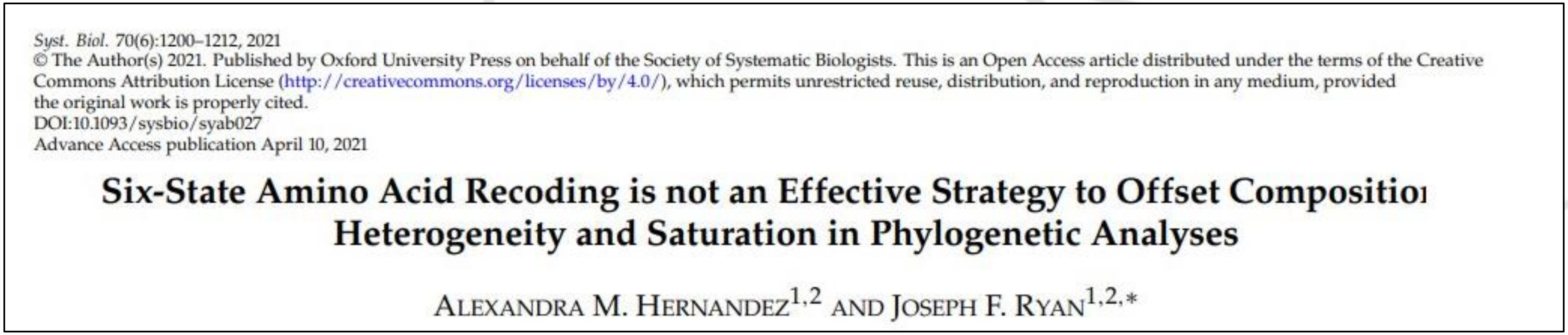
nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-021-22074-7 OPEN

Evidence for sponges as sister to all other animals from partitioned phylogenomics with mixture models and recoding

Anthony K. Redmond[1] & Aoife McLysaght[1]✉



Syst. Biol. 70(6):1200–1212, 2021
© The Author(s) 2021. Published by Oxford University Press on behalf of the Society of Systematic Biologists. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
DOI:10.1093/sysbio/syab027
Advance Access publication April 10, 2021

Six-State Amino Acid Recoding is not an Effective Strategy to Offset Compositio Heterogeneity and Saturation in Phylogenetic Analyses

ALEXANDRA M. HERNANDEZ[1,2] AND JOSEPH F. RYAN[1,2,*]

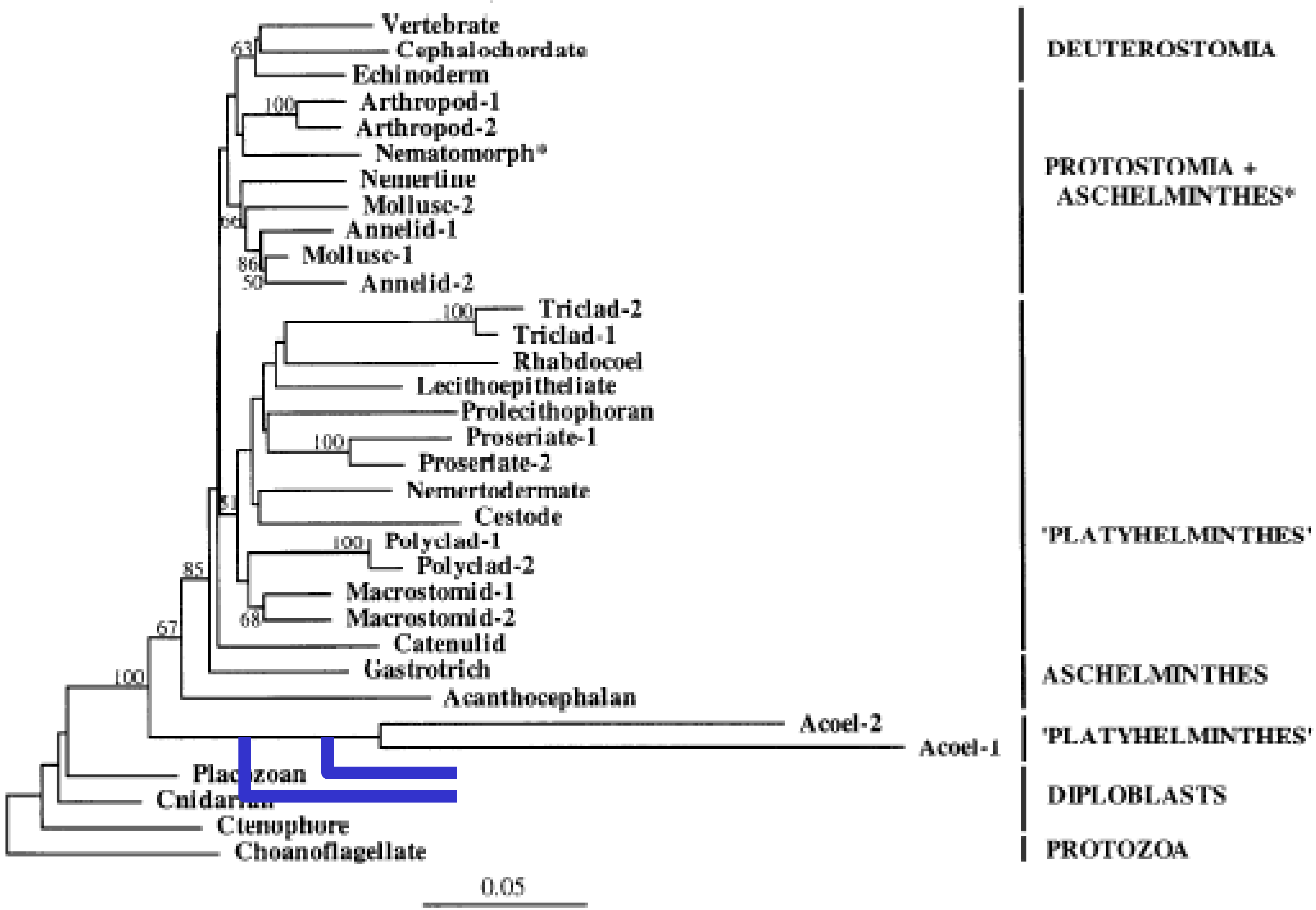
Avoiding LBA

1. Use methods less sensitive to **LBA**, such as **ML** or **BI**

2. **Model modifications** (Gamma-distribution, covarion, recoding)

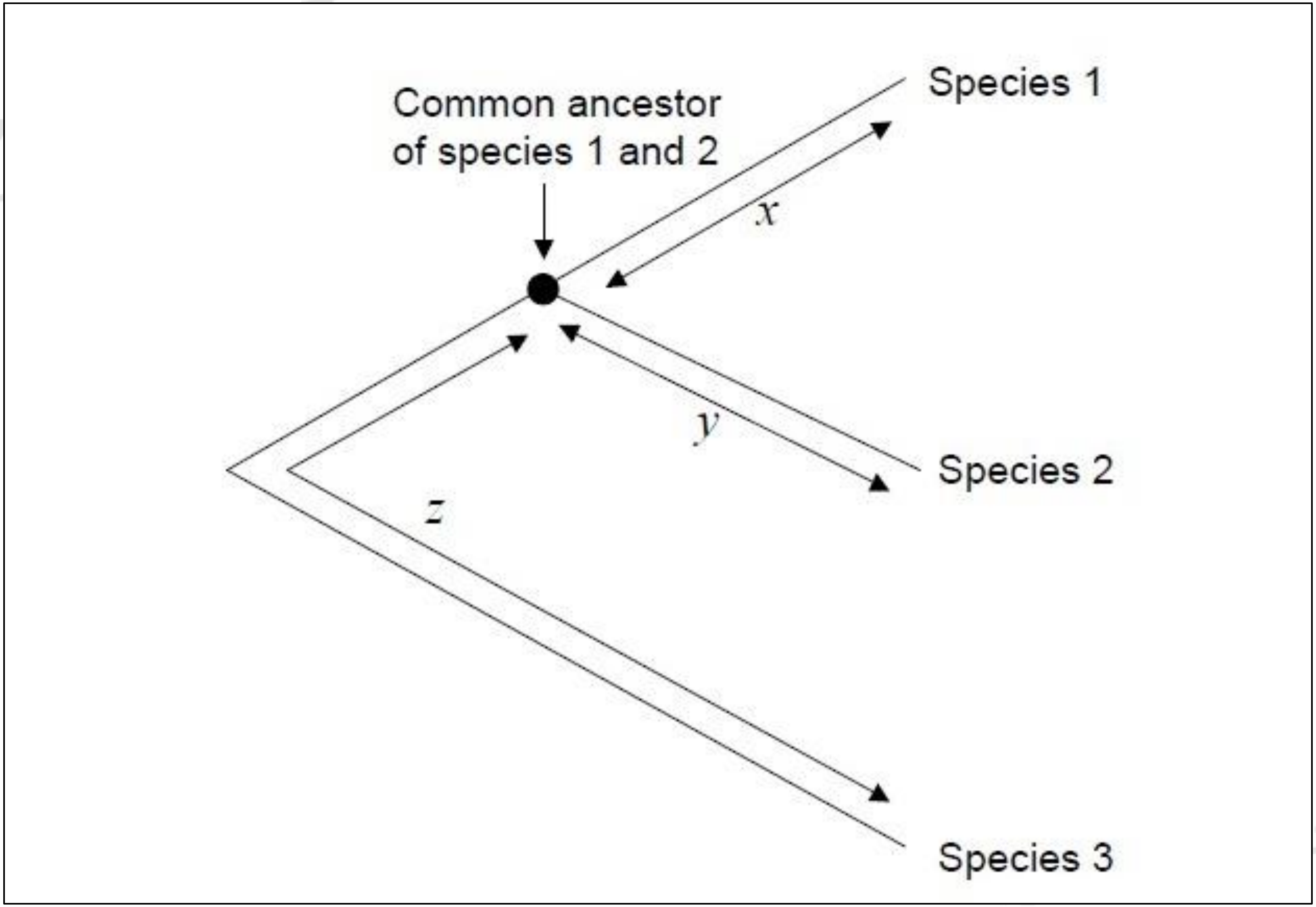
3. Using the **shortest branched representatives** available for each phyla (Relative Rate Test)

Long Branch Atraction



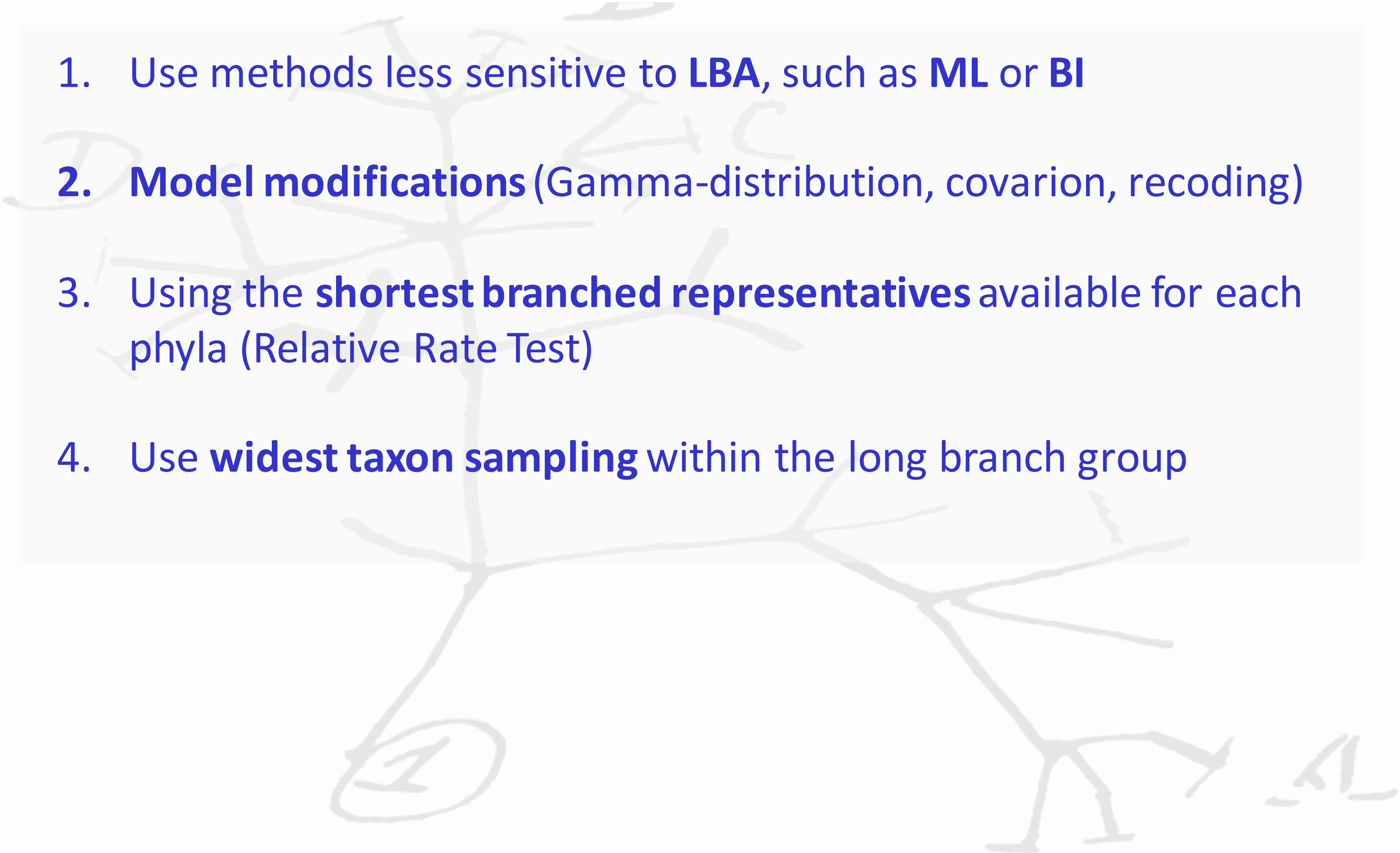
Carranza et al, 1997

Relative Rate Test

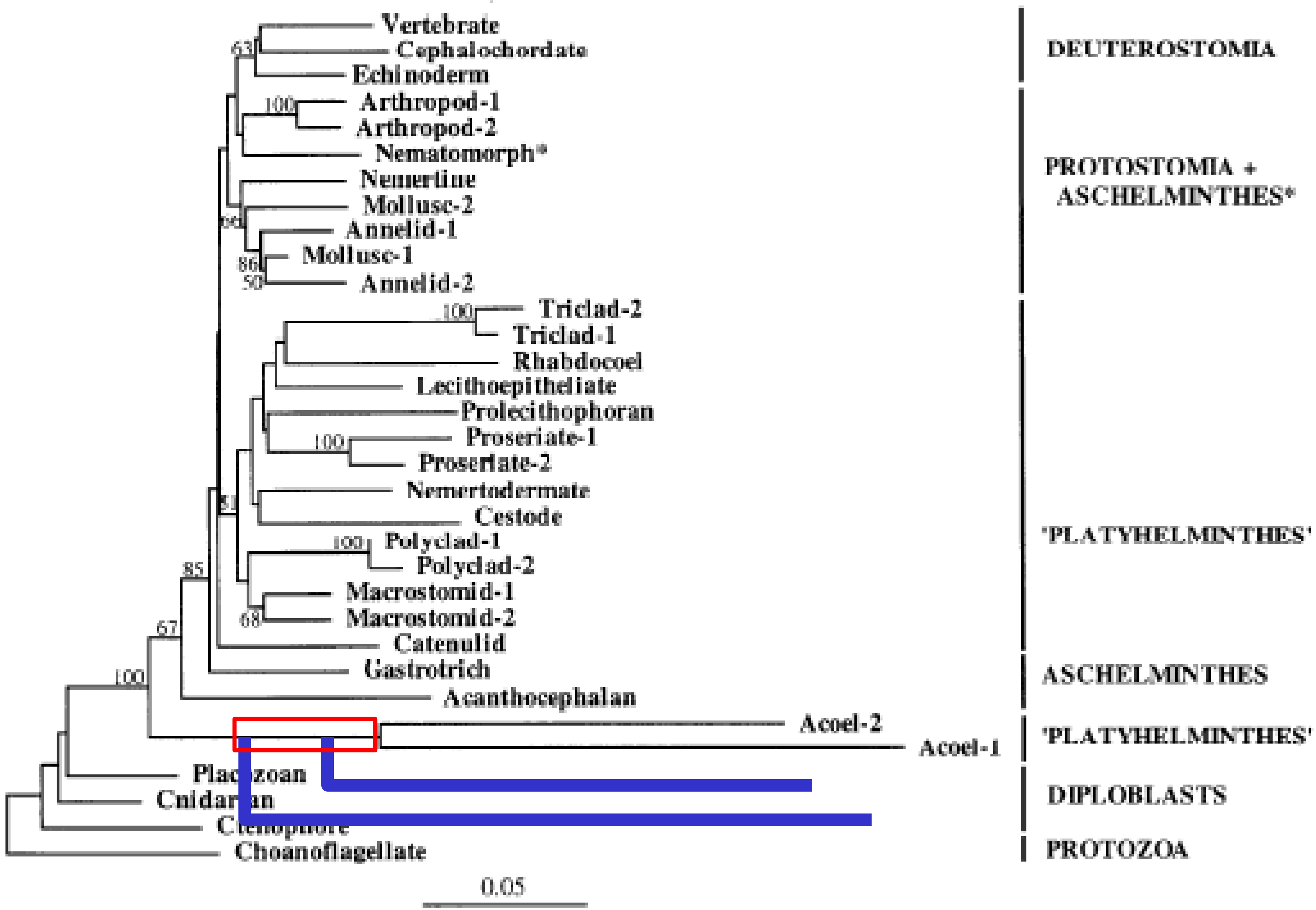


Avoiding LBA

1. Use methods less sensitive to **LBA**, such as **ML** or **BI**
2. **Model modifications** (Gamma-distribution, covarion, recoding)
3. Using the **shortest branched representatives** available for each phyla (Relative Rate Test)
4. Use **widest taxon sampling** within the long branch group

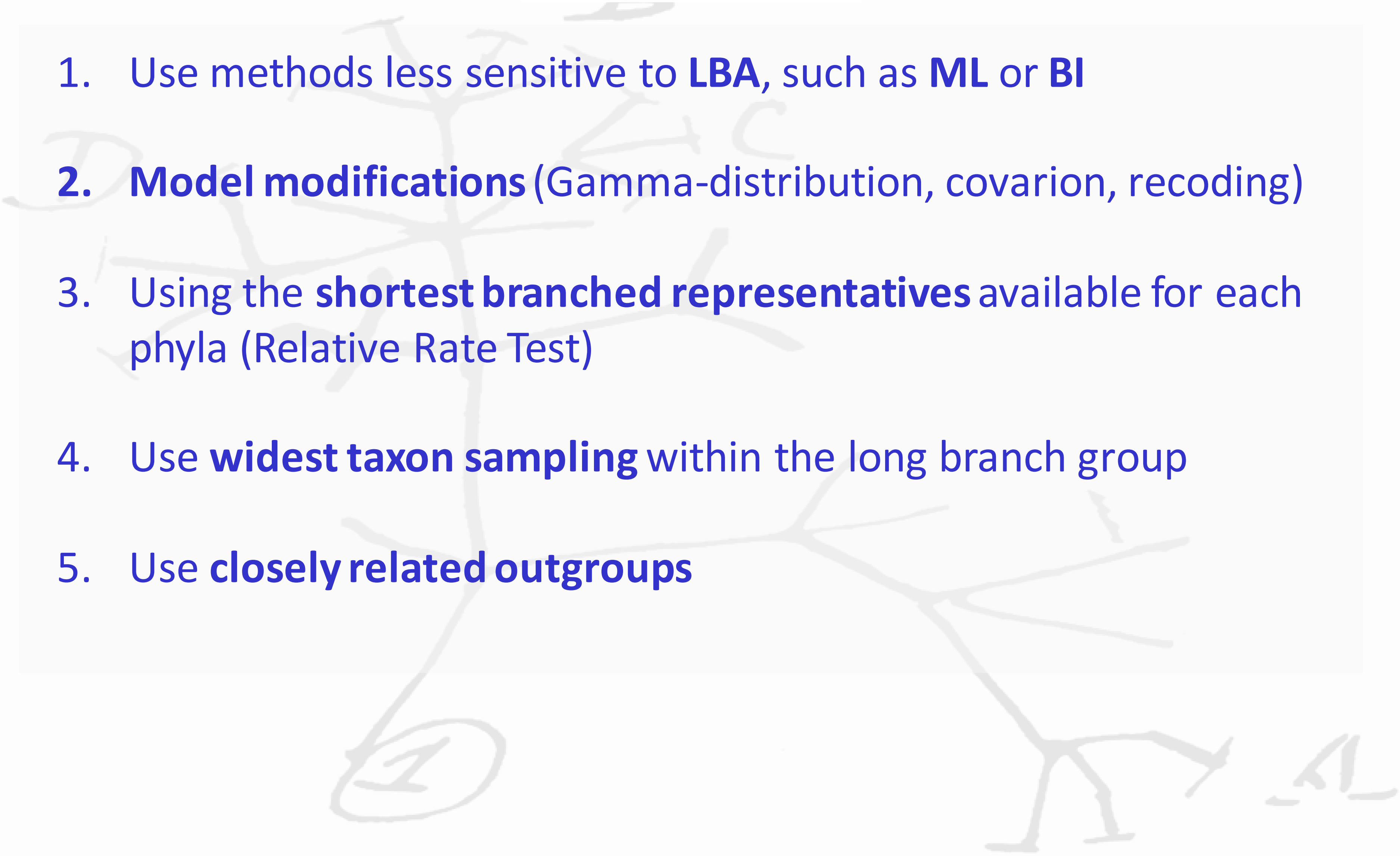


Long Branch Atraction



Carranza et al, 1997

Avoiding LBA



1. Use methods less sensitive to **LBA**, such as **ML** or **BI**

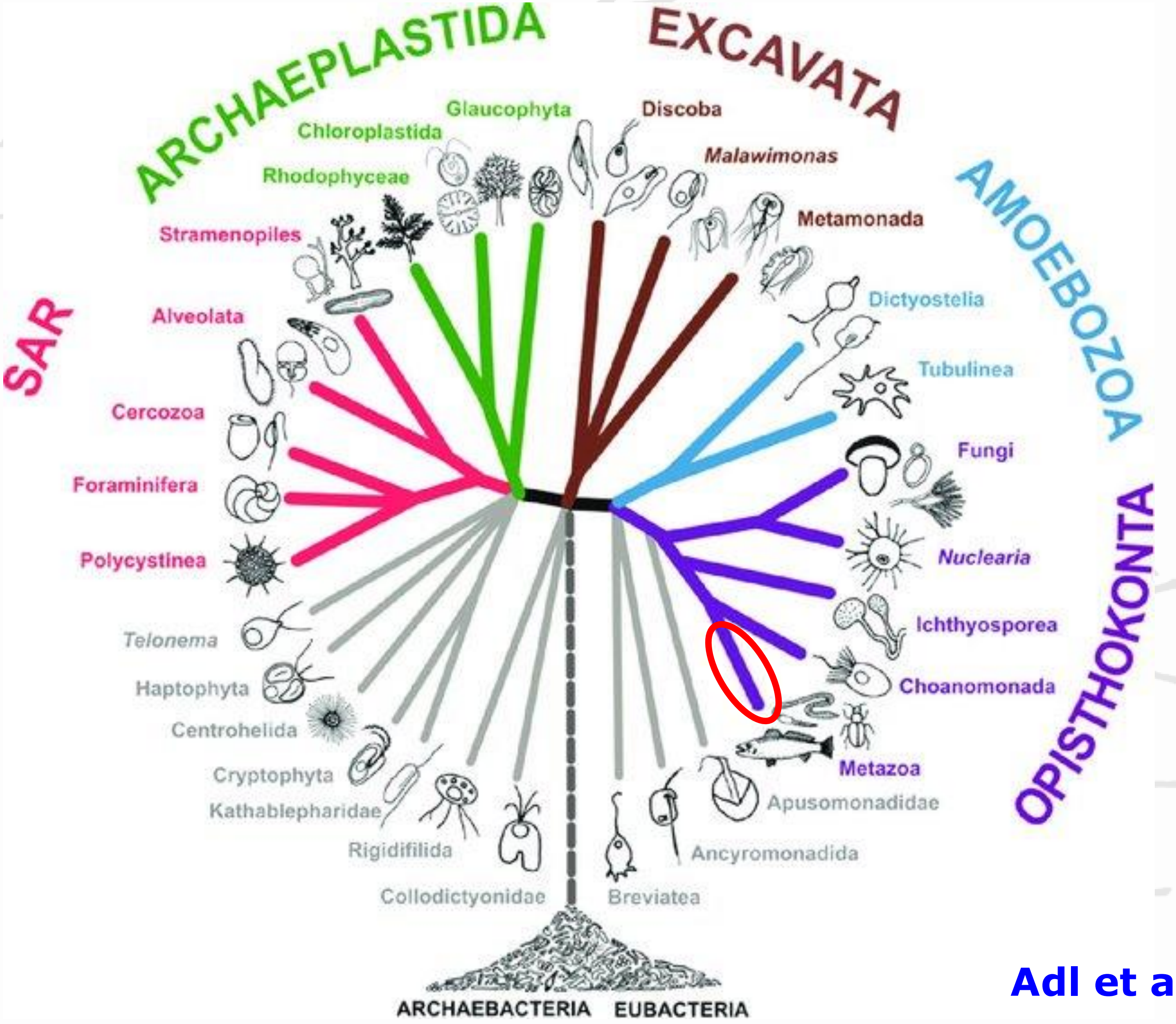
2. Model modifications (Gamma-distribution, covarion, recoding)

3. Using the **shortest branched representatives** available for each phyla (Relative Rate Test)

4. Use **widest taxon sampling** within the long branch group

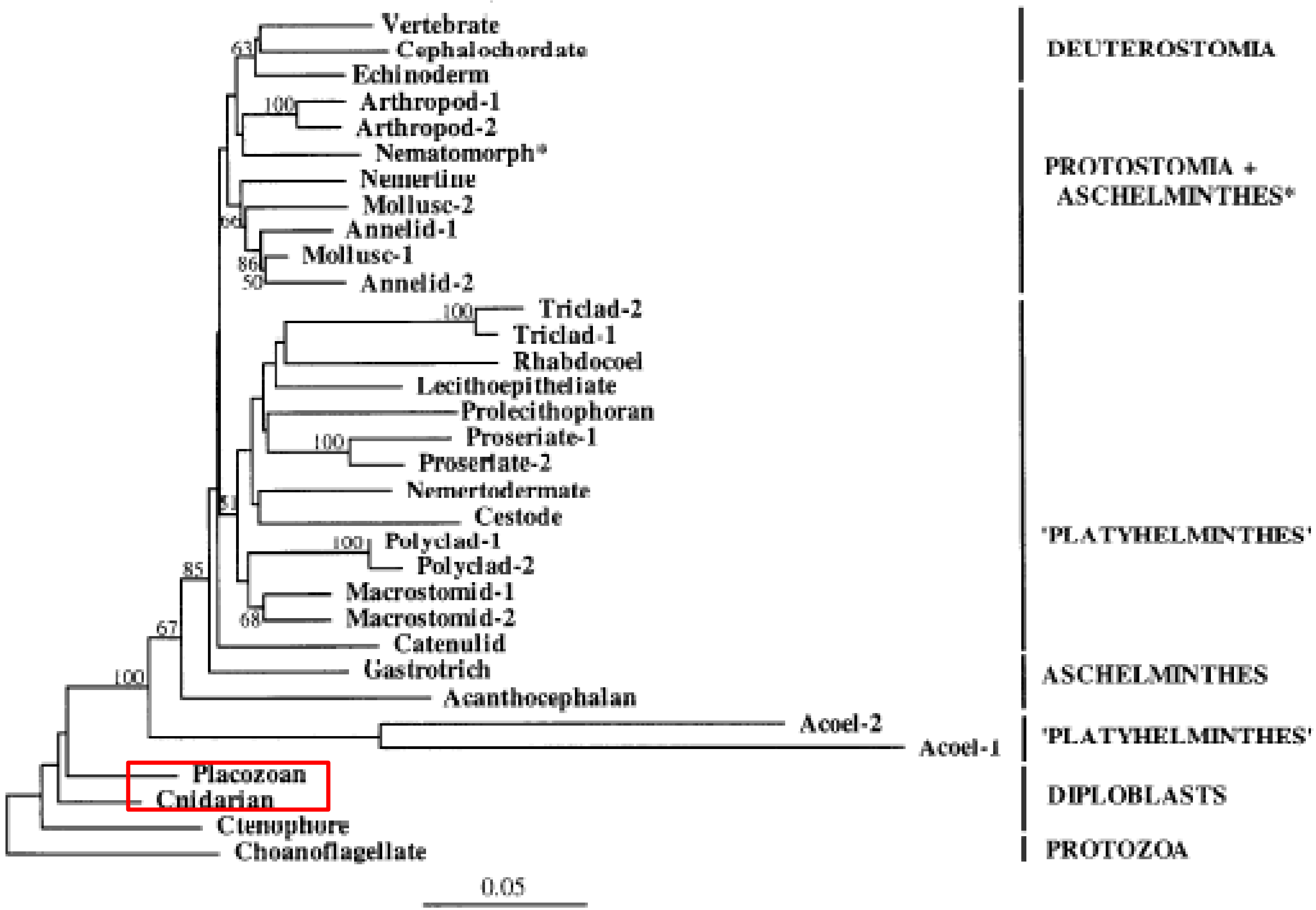
5. Use **closely related outgroups**

Long Branch Atraction



Adl et al, 2012

Long Branch Atraction

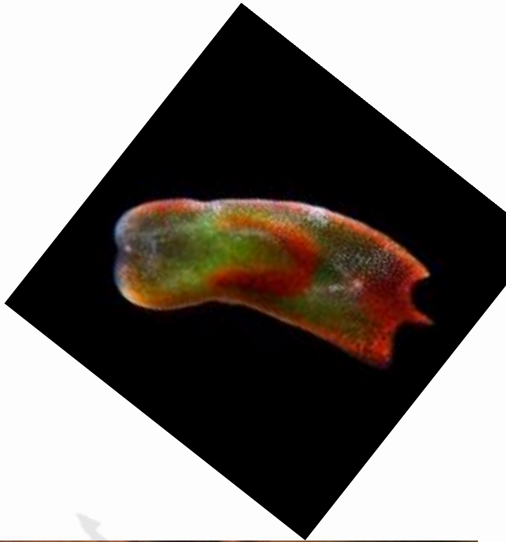


Carranza et al, 1997

Molecular: 18S

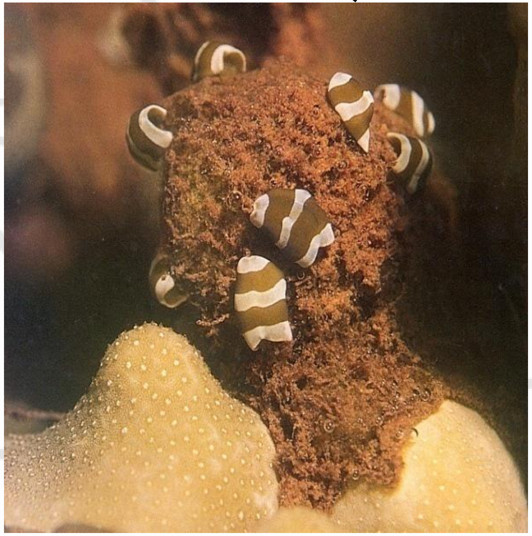
Acoel Flatworms: Earliest Extant Bilaterian Metazoans, Not Members of Platyhelminthes

Iñaki Ruiz-Trillo,[1] Marta Riutort,[1] D. Timothy J. Littlewood,[2]
Elisabeth A. Herniou,[2] Jaume Baguñà[1]*

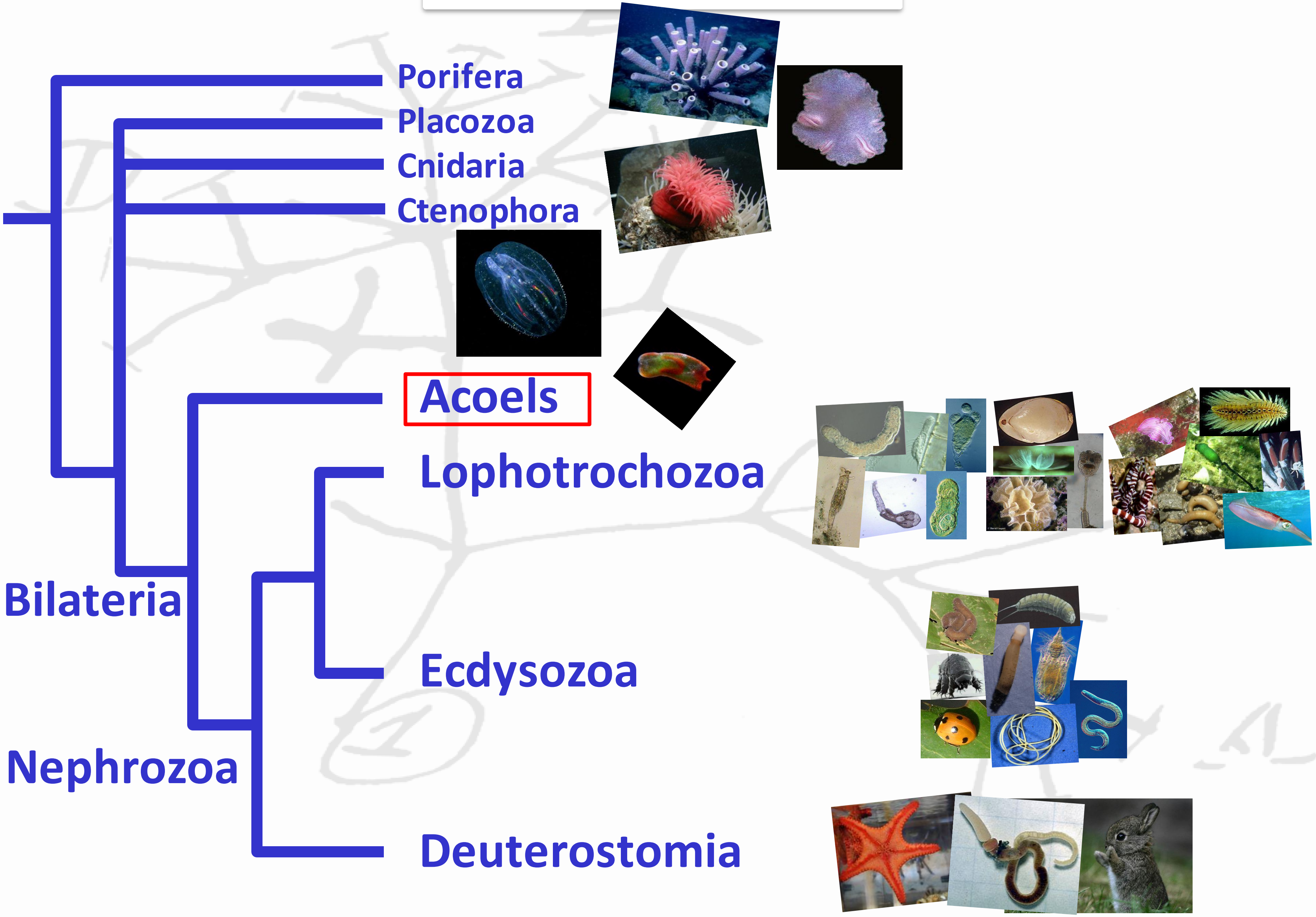


Marta Riutort

Ruiz-Trillo et al, 1999



Nephrozoa



Porifera

Placozoa

Cnidaria

Ctenophora

Acoels

Lophotrochozoa

Ecdysozoa

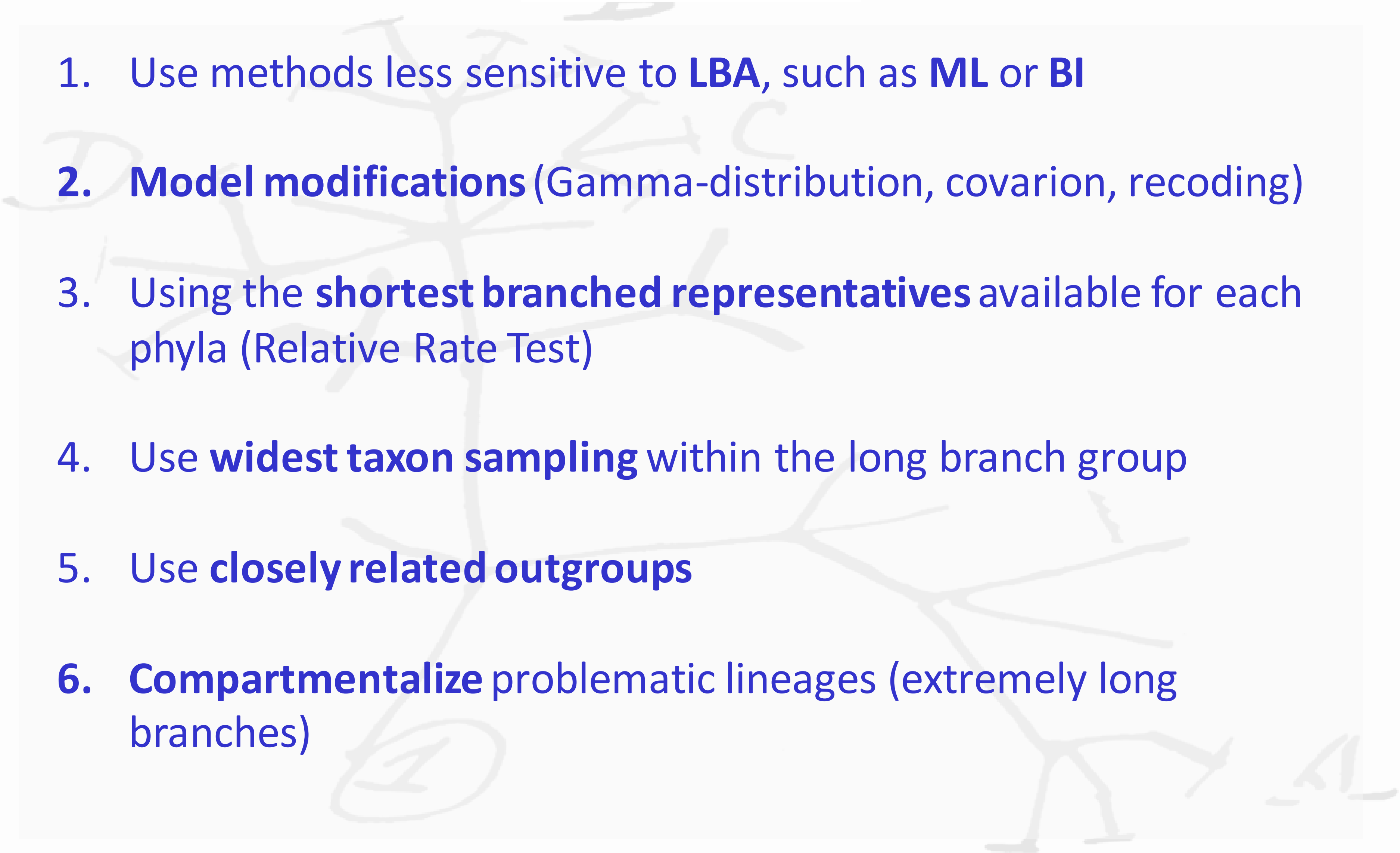
Deuterostomia

Bilateria

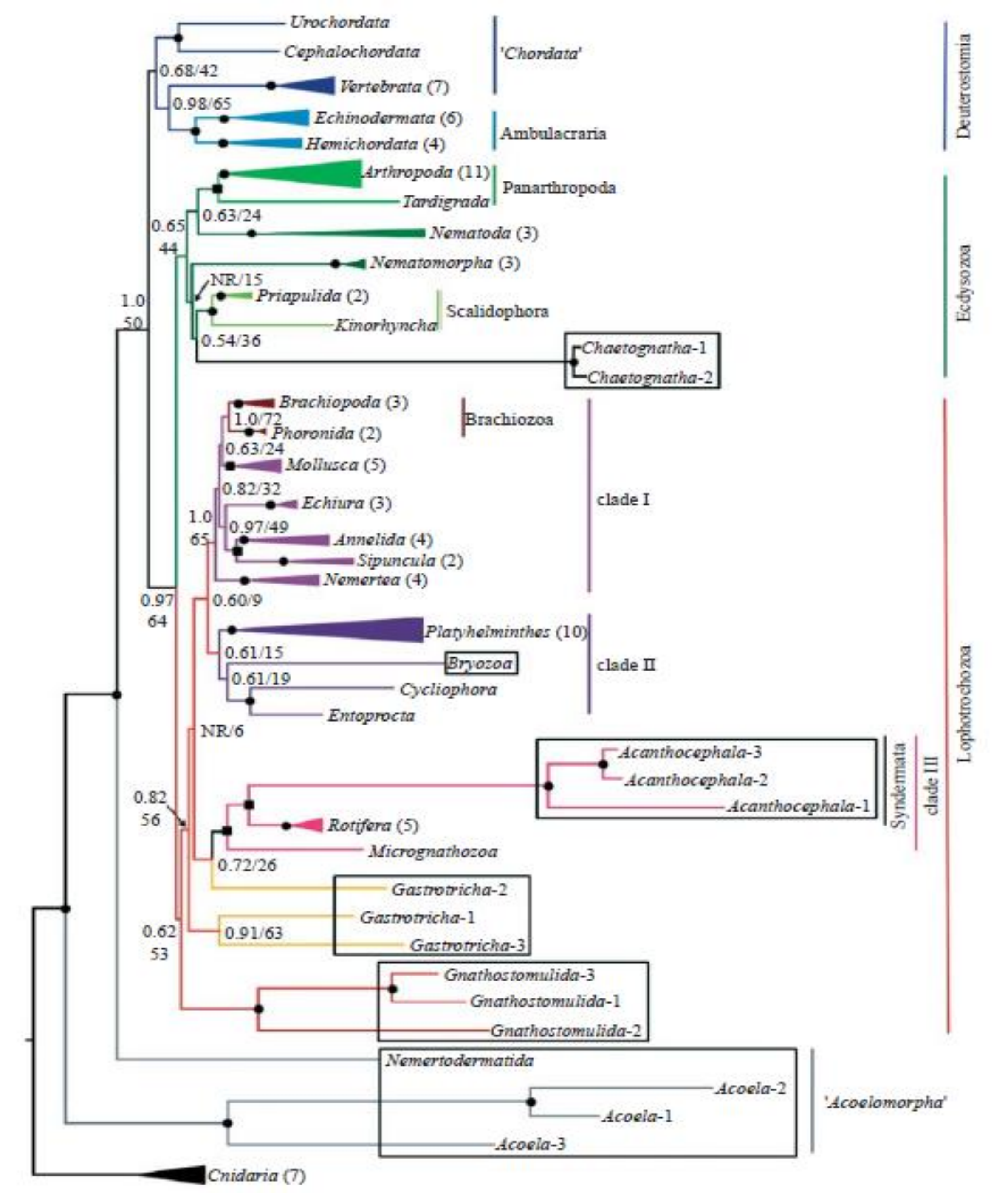
Nephrozoa


GENETIC EXPERIMENTS

Avoiding LBA

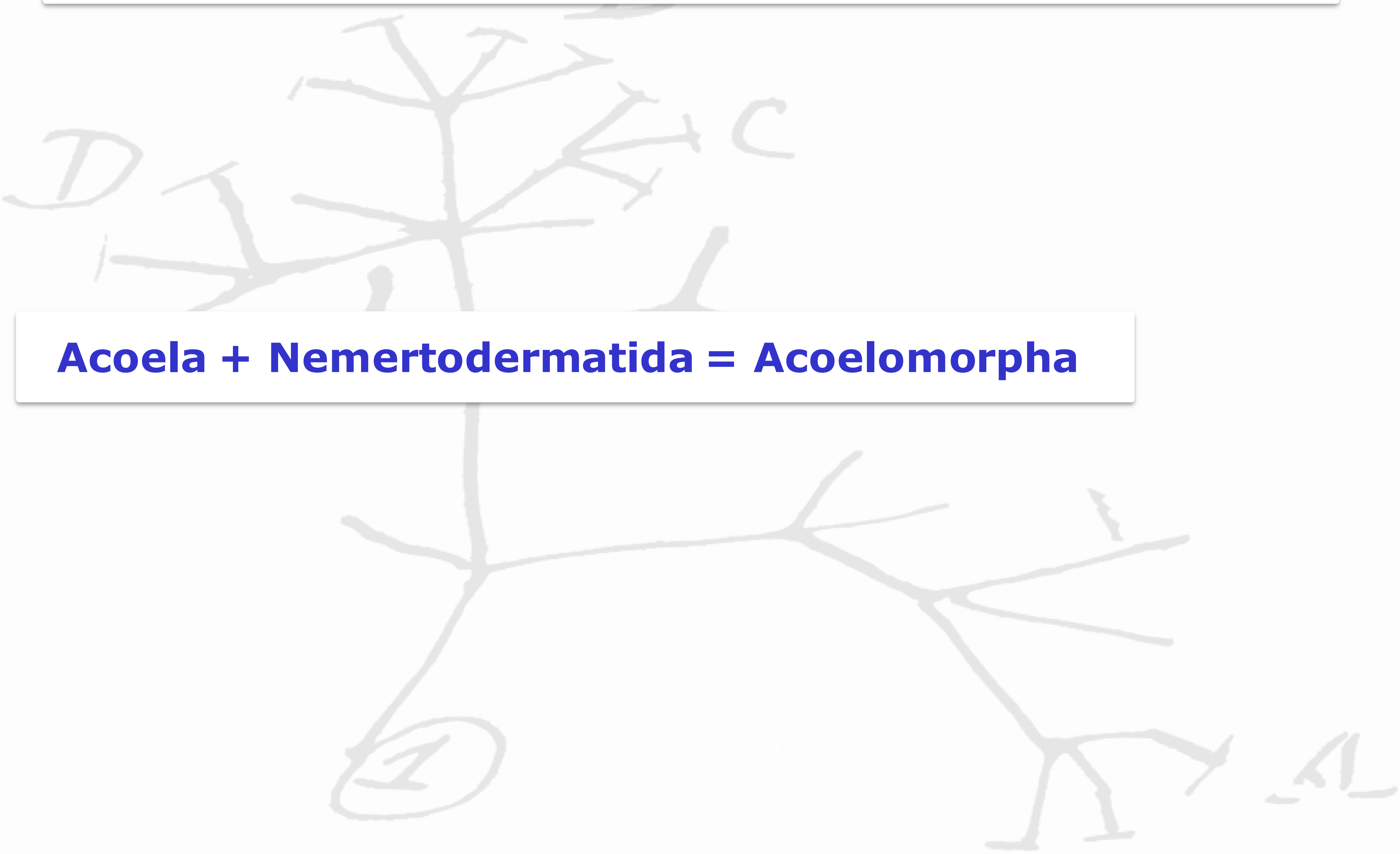


1. Use methods less sensitive to **LBA**, such as **ML** or **BI**
2. **Model modifications** (Gamma-distribution, covarion, recoding)
3. Using the **shortest branched representatives** available for each phyla (Relative Rate Test)
4. Use **widest taxon sampling** within the long branch group
5. Use **closely related outgroups**
6. **Compartmentalize** problematic lineages (extremely long branches)



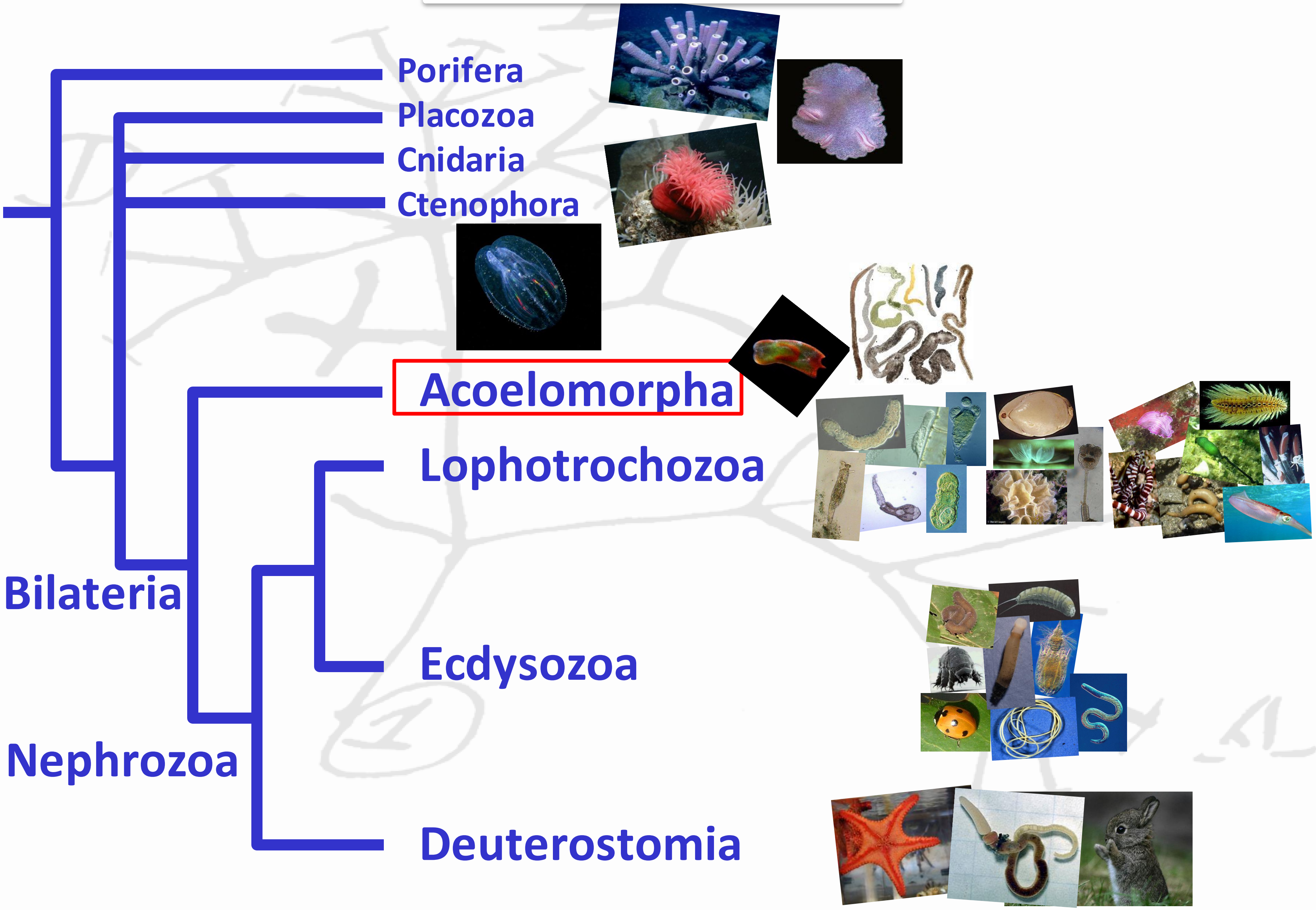
Paps et al, 2009a

Acoels, nemertodermatids, and xenoturbellids



Acoela + Nemertodermatida = Acoelomorpha

Nephrozoa



Porifera

Placozoa

Cnidaria

Ctenophora

Acoelomorpha

Lophotrochozoa

Ecdysozoa

Deuterostomia

Bilateria

Nephrozoa

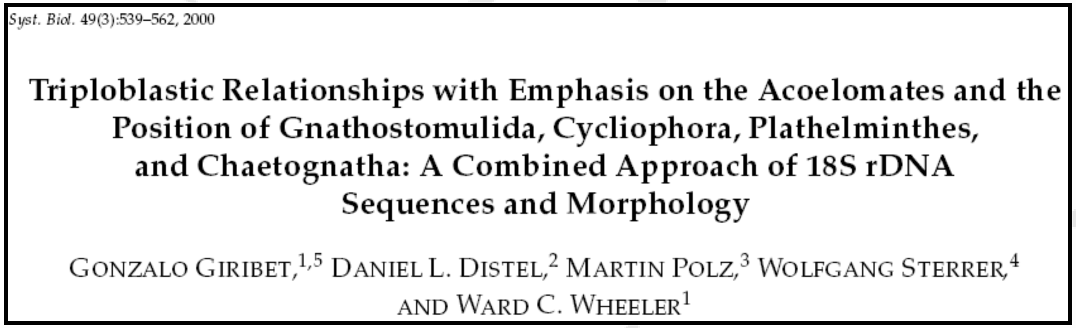
Taxon sampling in 18S

Cladistics **14**, 249–285 (1998)
WWW http://www.apnet.com
Article No. cl980070

Phylogeny of the Metazoa Based on Morphological and 18S Ribosomal DNA Evidence

Jan Zrzavý[1,2], Stanislav Mihulka[1], Pavel Kepka[1], Aleš Bezděk[1,2], and David Tietz[1]

Zrzavý et al, 1998
MP: 18S + Morphology
151 seqs
28 phyla bilaterians
Consensus sequences

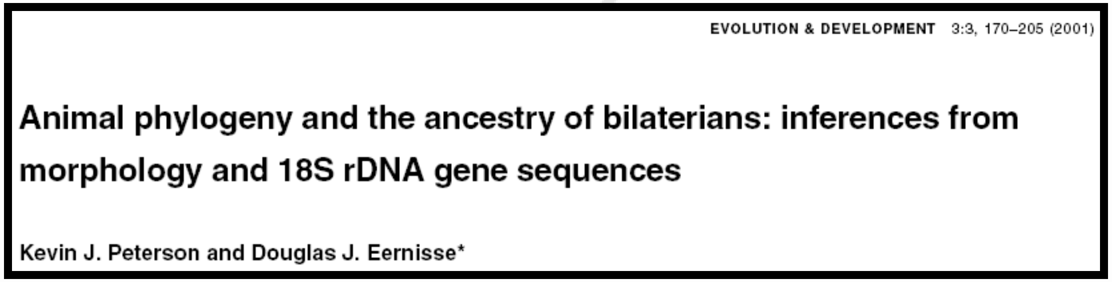


Syst. Biol. 49(3):539–562, 2000

Triploblastic Relationships with Emphasis on the Acoelomates and the Position of Gnathostomulida, Cycliophora, Plathelminthes, and Chaetognatha: A Combined Approach of 18S rDNA Sequences and Morphology

GONZALO GIRIBET,[1,5] DANIEL L. DISTEL,[2] MARTIN POLZ,[3] WOLFGANG STERRER,[4] AND WARD C. WHEELER[1]

Giribet et al, 2000
POY: 18S + Morphology
147 spp
28 phyla bilaterians

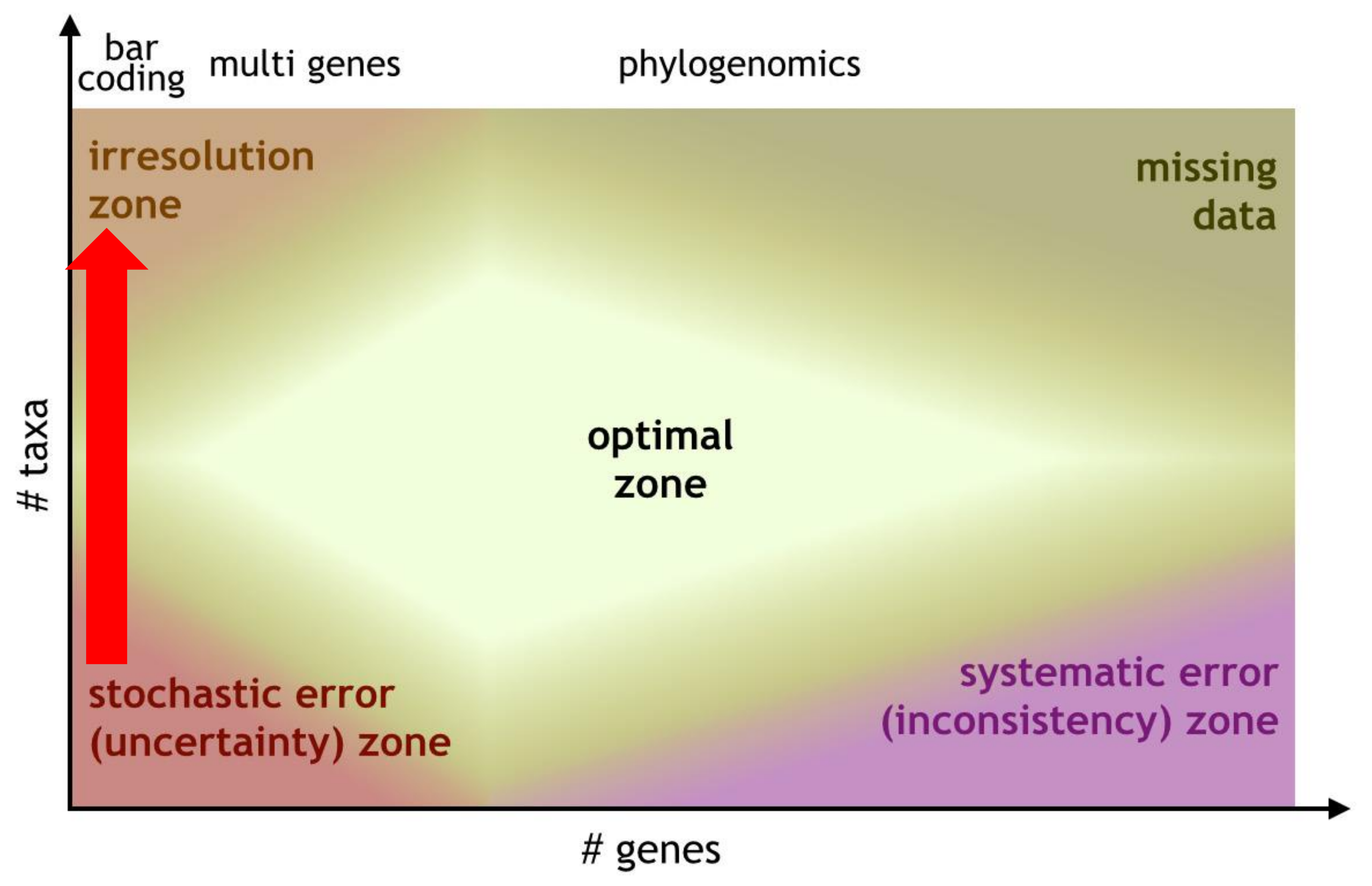


EVOLUTION & DEVELOPMENT 3:3, 170–205 (2001)

Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences

Kevin J. Peterson and Douglas J. Eernisse*

Peterson & Eernisse, 2001
MP: 18S + Morphology
304 seqs
28 phyla bilaterians
"AGG" removed



Philipe et al, 2017

Avoiding LBA

1. Use methods less sensitive to **LBA**, such as **ML** or **BI**

2. **Model modifications** (Gamma-distribution, covarion, recoding)

3. Using the **shortest branched representatives** available for each phyla (Relative Rate Test)

4. Use **widest taxon sampling** within the long branch group

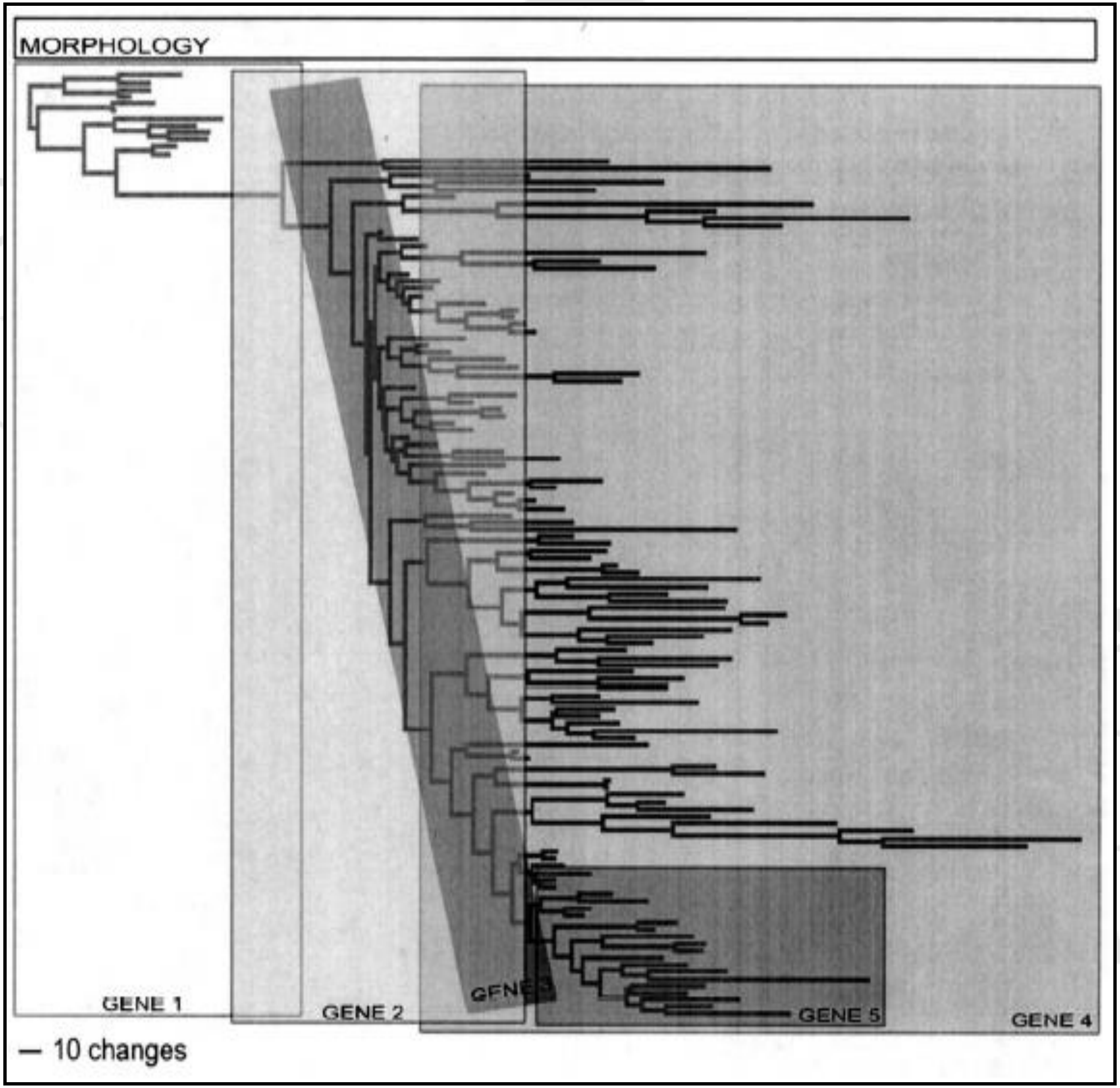
5. Use **closely related outgroups**

6. **Compartmentalize** problematic lineages (extremely long branches)

7. **More data!!!**

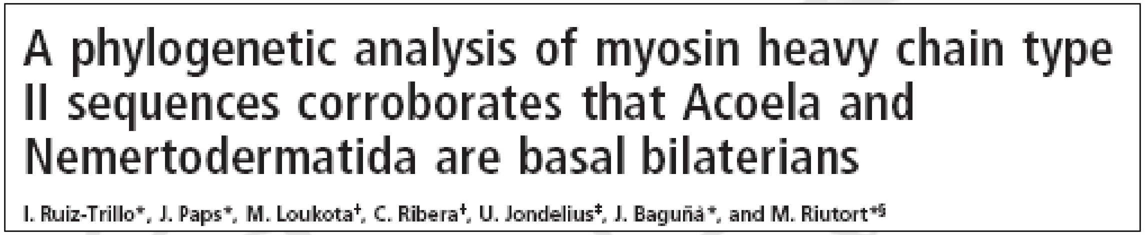
The Molecular Era:
More markers…

New markers



Giribet, 2000

New markers



A phylogenetic analysis of myosin heavy chain type II sequences corroborates that Acoela and Nemertodermatida are basal bilaterians

I. Ruiz-Trillo*, J. Paps*, M. Loukota†, C. Ribera†, U. Jondelius‡, J. Baguñà*, and M. Riutort*§

Ruiz-Trillo et al, 2002

18S + **Myosin**

THE ROYAL SOCIETY

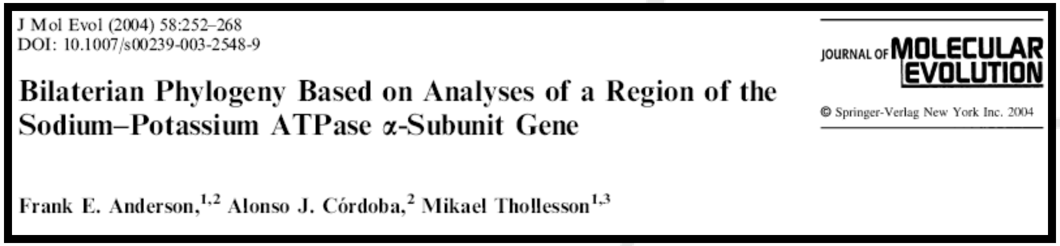
Received 26 September 2002
Accepted 15 January 2003

Combined large and small subunit ribosomal RNA phylogenies support a basal position of the acoelomorph flatworms

Maximilian J. Telford[1*], Anne E. Lockyer[2], Chloë Cartwright-Finch[1]
and D. Timothy J. Littlewood[2]

Telford et al, 2003

18S + **28S**



J Mol Evol (2004) 58:252–268
DOI: 10.1007/s00239-003-2548-9

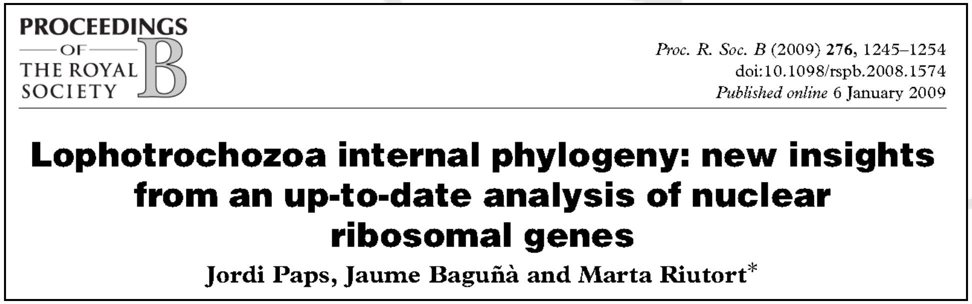
JOURNAL OF MOLECULAR EVOLUTION

Bilaterian Phylogeny Based on Analyses of a Region of the Sodium–Potassium ATPase α-Subunit Gene

Frank E. Anderson,[1,2] Alonso J. Córdoba,[2] Mikael Thollesson[1,3]

Anderson et al, 2004

ATPase alpha



PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2009) **276**, 1245–1254
doi:10.1098/rspb.2008.1574

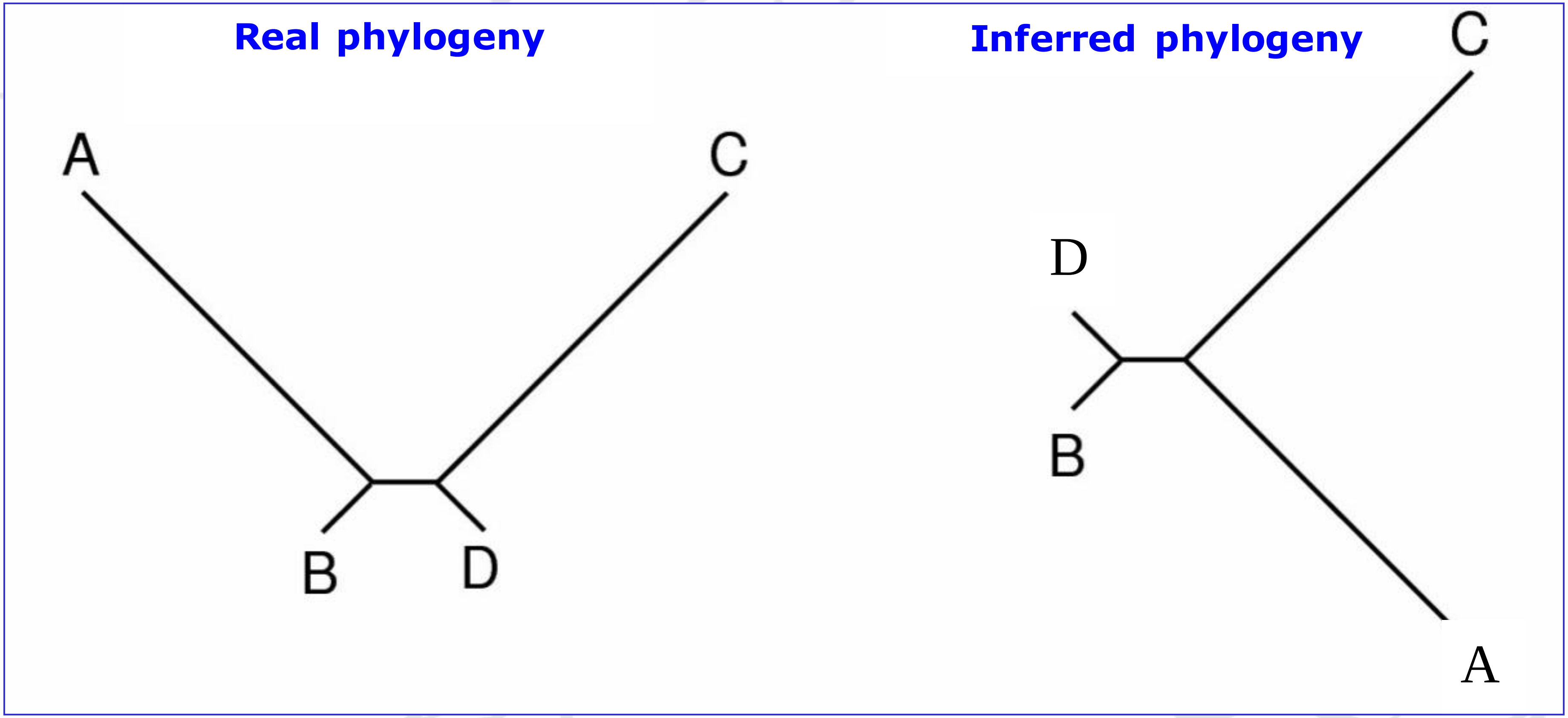
Lophotrochozoa internal phylogeny: new insights from an up-to-date analysis of nuclear ribosomal genes

Jordi Paps, Jaume Baguñà and Marta Riutort*

Paps et al, 2009

18S + **28S**

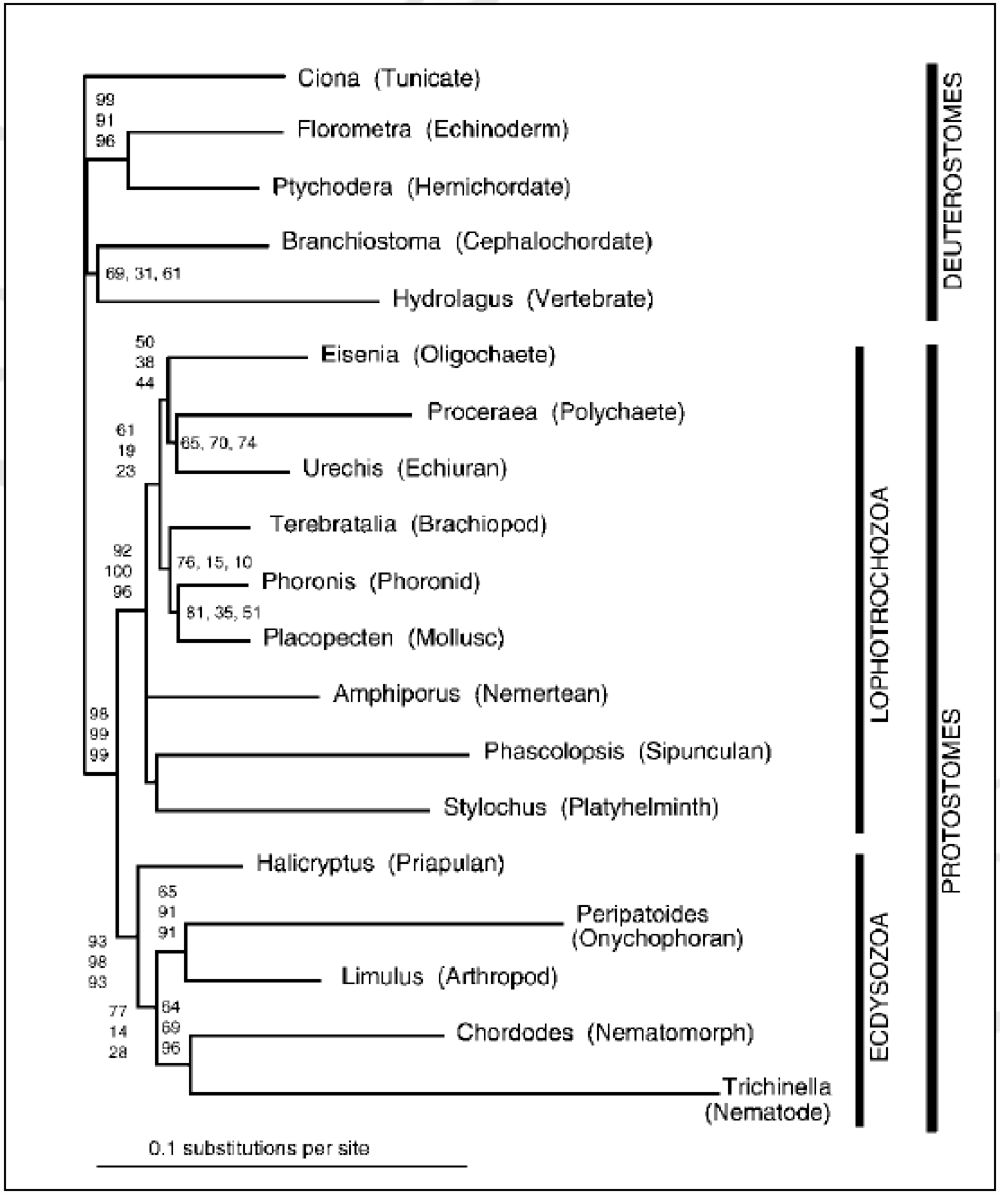
Long Branch Atraction



Real phylogeny

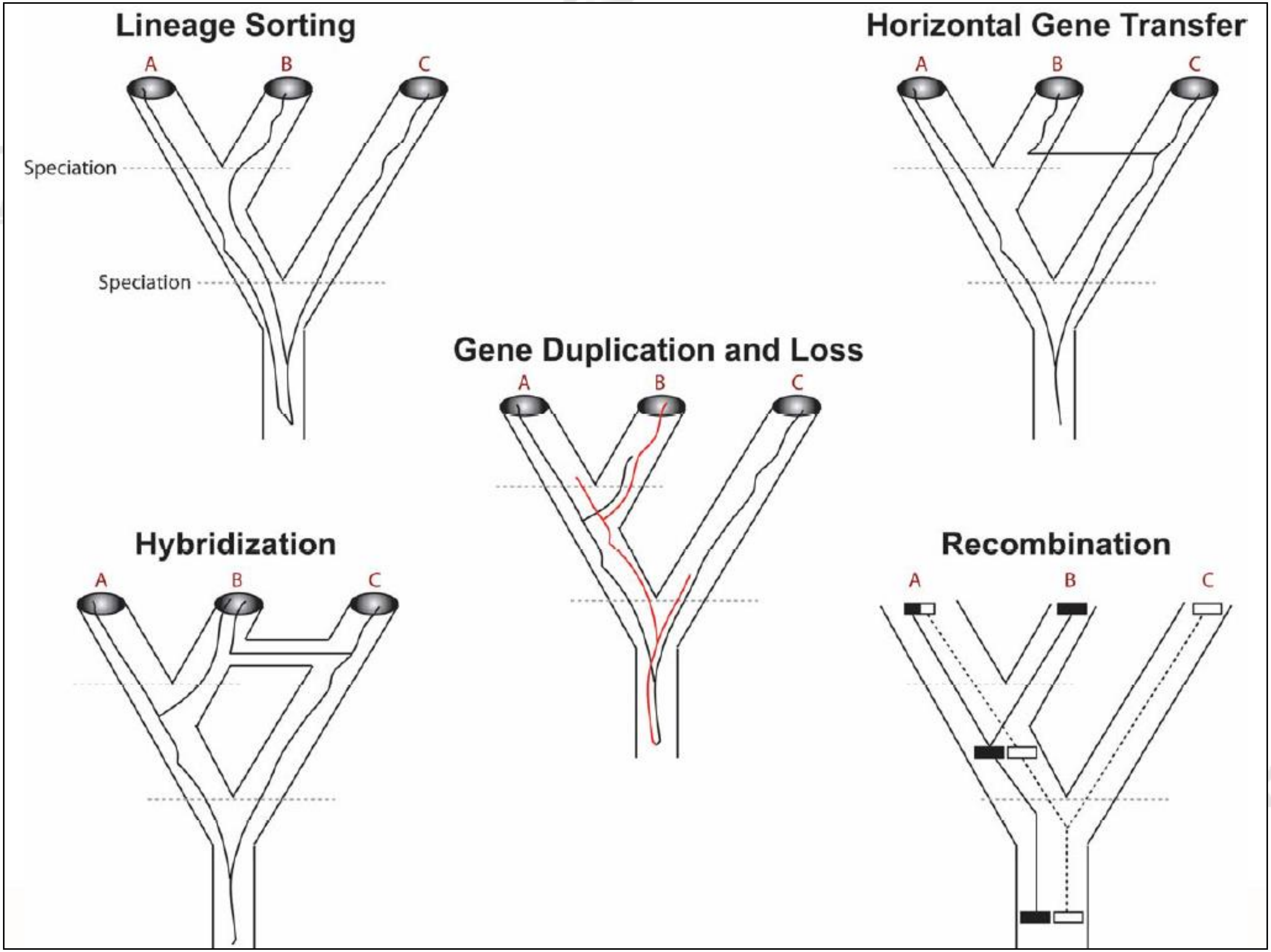
Inferred phylogeny

Poor taxon sampling



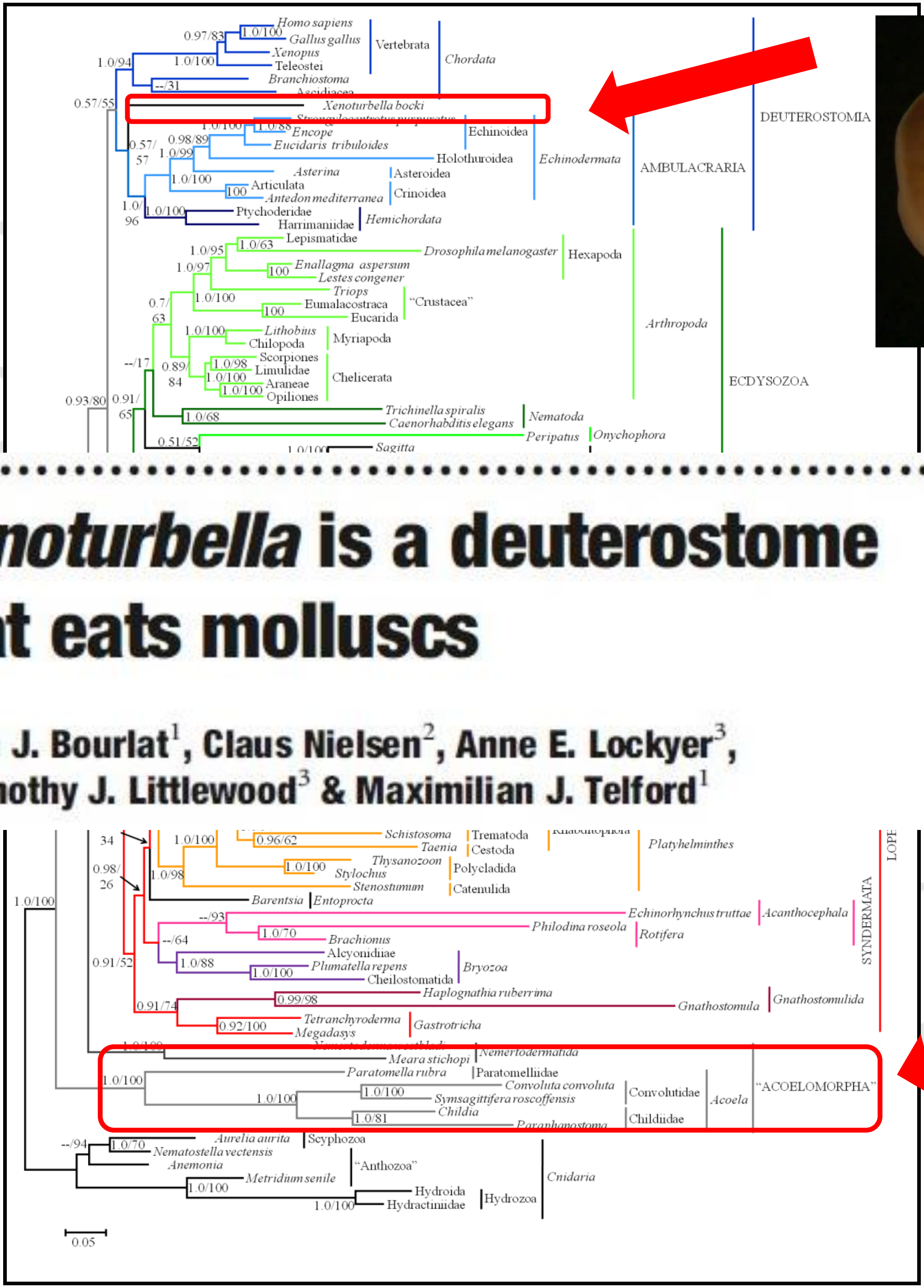
28S, Mallat & Whinchell 2002

Orthology concerns



13 Genes

- 27 Phyla
- 90 OTU
- 8.800 bp
- ML & Mr Bayes

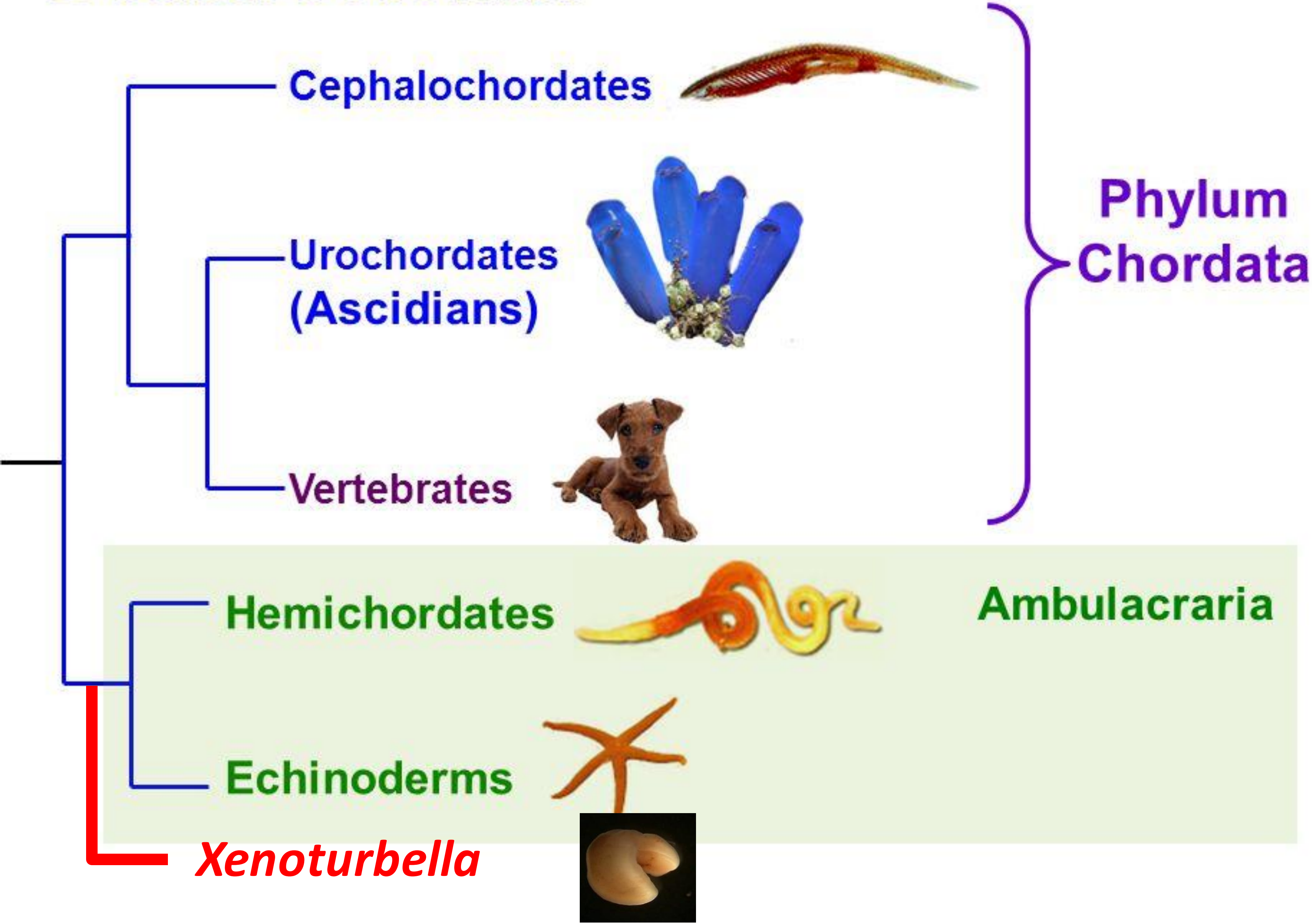


***Xenoturbella* is a deuterostome that eats molluscs**

Sarah J. Bourlat[1], Claus Nielsen[2], Anne E. Lockyer[3], D. Timothy J. Littlewood[3] & Maximilian J. Telford[1]

Paps et al 2009b

Deuterostomia



Cephalochordates

Urochordates (Ascidians)

Vertebrates

Phylum Chordata

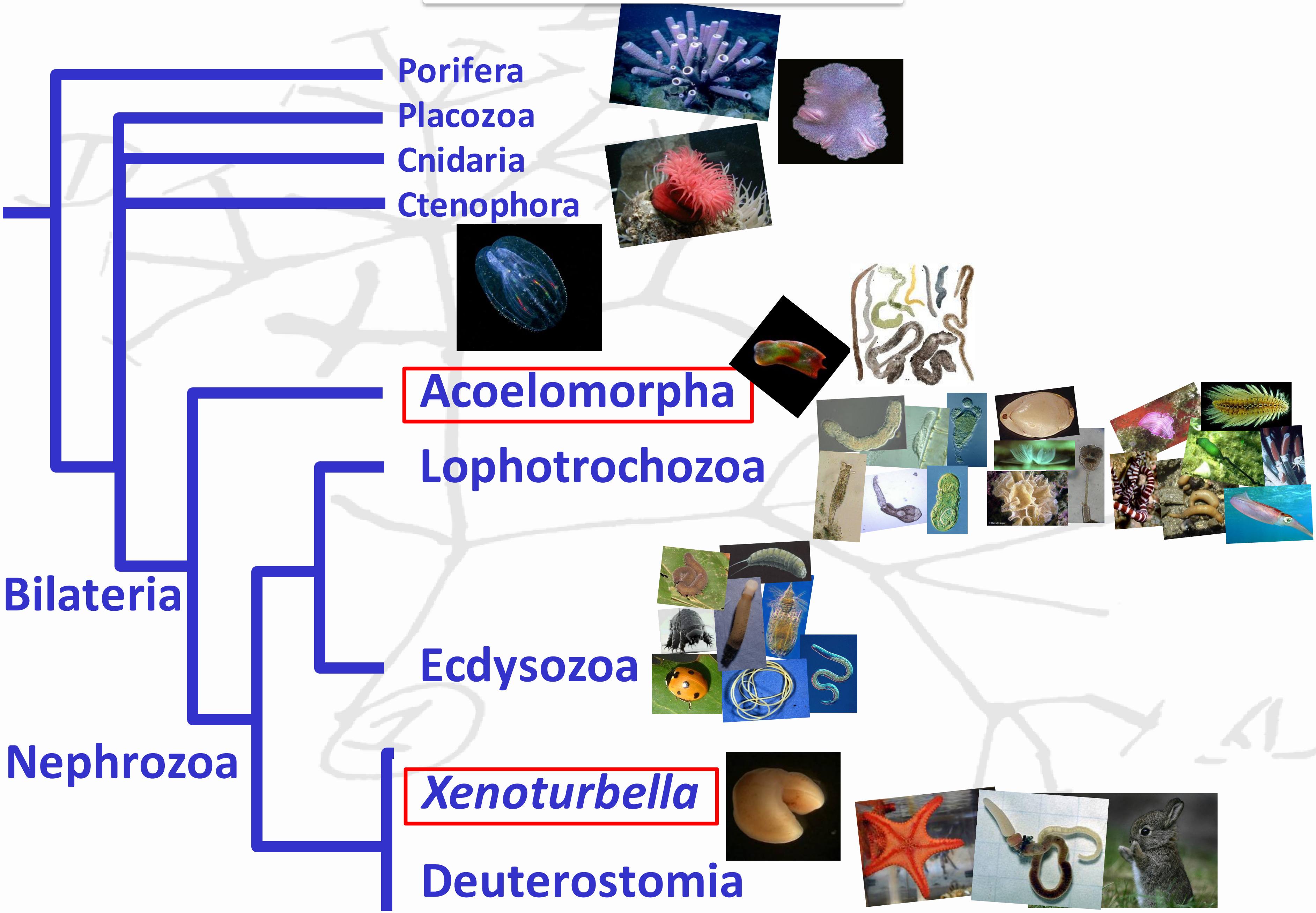
Hemichordates

Echinoderms

Ambulacraria

Xenoturbella

Nephrozoa



Porifera

Placozoa

Cnidaria

Ctenophora

Acoelomorpha

Lophotrochozoa

Ecdysozoa

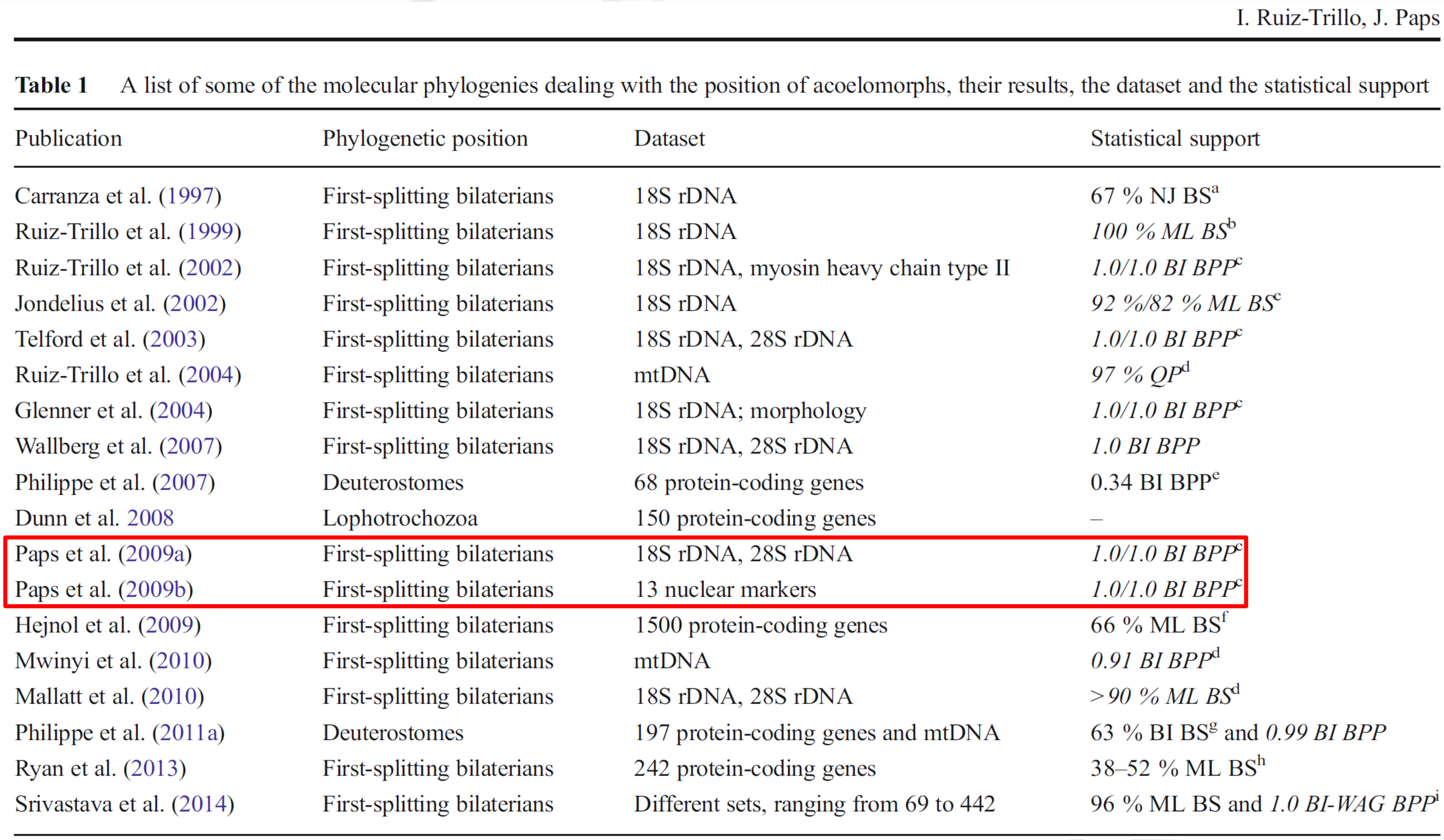
Xenoturbella

Deuterostomia

Bilateria

Nephrozoa

The last march of the PCR...



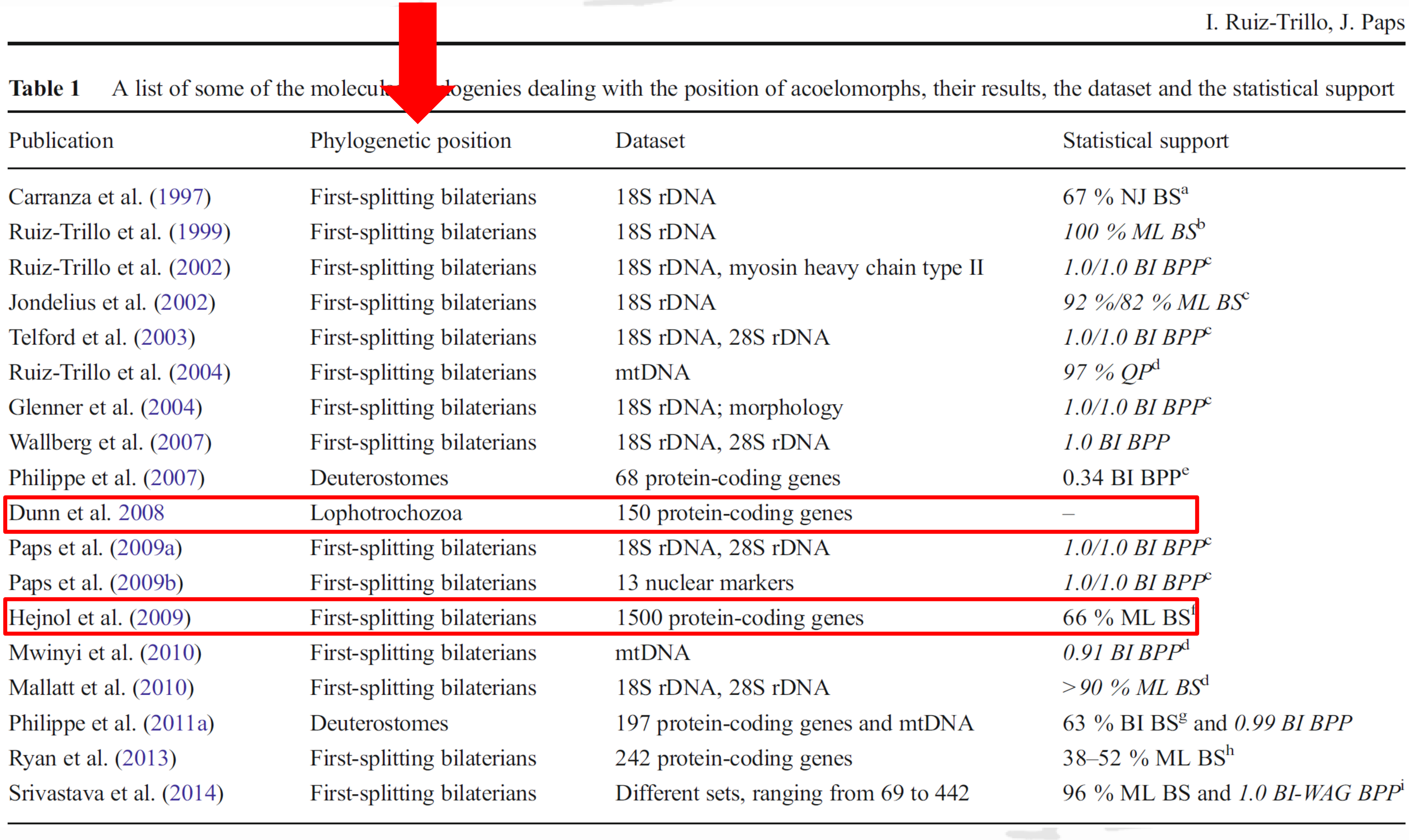
I. Ruiz-Trillo, J. Paps

Table 1 A list of some of the molecular phylogenies dealing with the position of acoelomorphs, their results, the dataset and the statistical support

Publication	Phylogenetic position	Dataset	Statistical support
Carranza et al. (1997)	First-splitting bilaterians	18S rDNA	67 % NJ BS[a]
Ruiz-Trillo et al. (1999)	First-splitting bilaterians	18S rDNA	*100 % ML BS*[b]
Ruiz-Trillo et al. (2002)	First-splitting bilaterians	18S rDNA, myosin heavy chain type II	*1.0/1.0 BI BPP*[c]
Jondelius et al. (2002)	First-splitting bilaterians	18S rDNA	*92 %/82 % ML BS*[c]
Telford et al. (2003)	First-splitting bilaterians	18S rDNA, 28S rDNA	*1.0/1.0 BI BPP*[c]
Ruiz-Trillo et al. (2004)	First-splitting bilaterians	mtDNA	*97 % QP*[d]
Glenner et al. (2004)	First-splitting bilaterians	18S rDNA; morphology	*1.0/1.0 BI BPP*[c]
Wallberg et al. (2007)	First-splitting bilaterians	18S rDNA, 28S rDNA	*1.0 BI BPP*
Philippe et al. (2007)	Deuterostomes	68 protein-coding genes	0.34 BI BPP[e]
Dunn et al. 2008	Lophotrochozoa	150 protein-coding genes	–
Paps et al. (2009a)	First-splitting bilaterians	18S rDNA, 28S rDNA	*1.0/1.0 BI BPP*[c]
Paps et al. (2009b)	First-splitting bilaterians	13 nuclear markers	*1.0/1.0 BI BPP*[c]
Hejnol et al. (2009)	First-splitting bilaterians	1500 protein-coding genes	66 % ML BS[f]
Mwinyi et al. (2010)	First-splitting bilaterians	mtDNA	*0.91 BI BPP*[d]
Mallatt et al. (2010)	First-splitting bilaterians	18S rDNA, 28S rDNA	*>90 % ML BS*[d]
Philippe et al. (2011a)	Deuterostomes	197 protein-coding genes and mtDNA	63 % BI BS[g] and *0.99 BI BPP*
Ryan et al. (2013)	First-splitting bilaterians	242 protein-coding genes	38–52 % ML BS[h]
Srivastava et al. (2014)	First-splitting bilaterians	Different sets, ranging from 69 to 442	96 % ML BS and *1.0 BI-WAG BPP*[i]

Ruiz-Trillo and Paps 2015

…the start of the phylotranscriptomics age



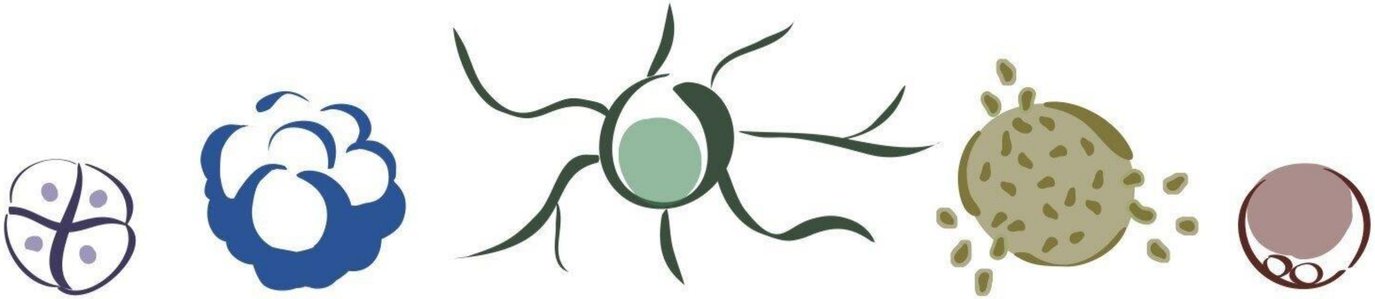
I. Ruiz-Trillo, J. Paps

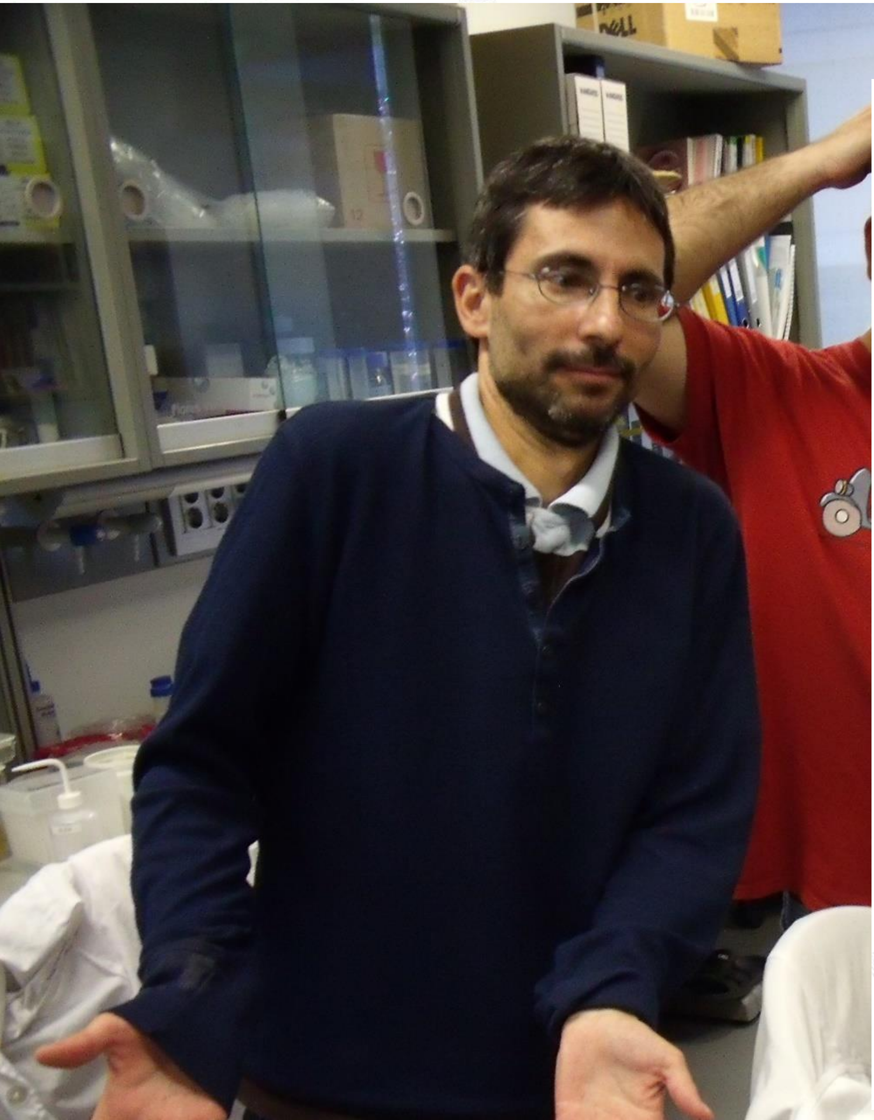
Table 1 A list of some of the molecu logenies dealing with the position of acoelomorphs, their results, the dataset and the statistical support

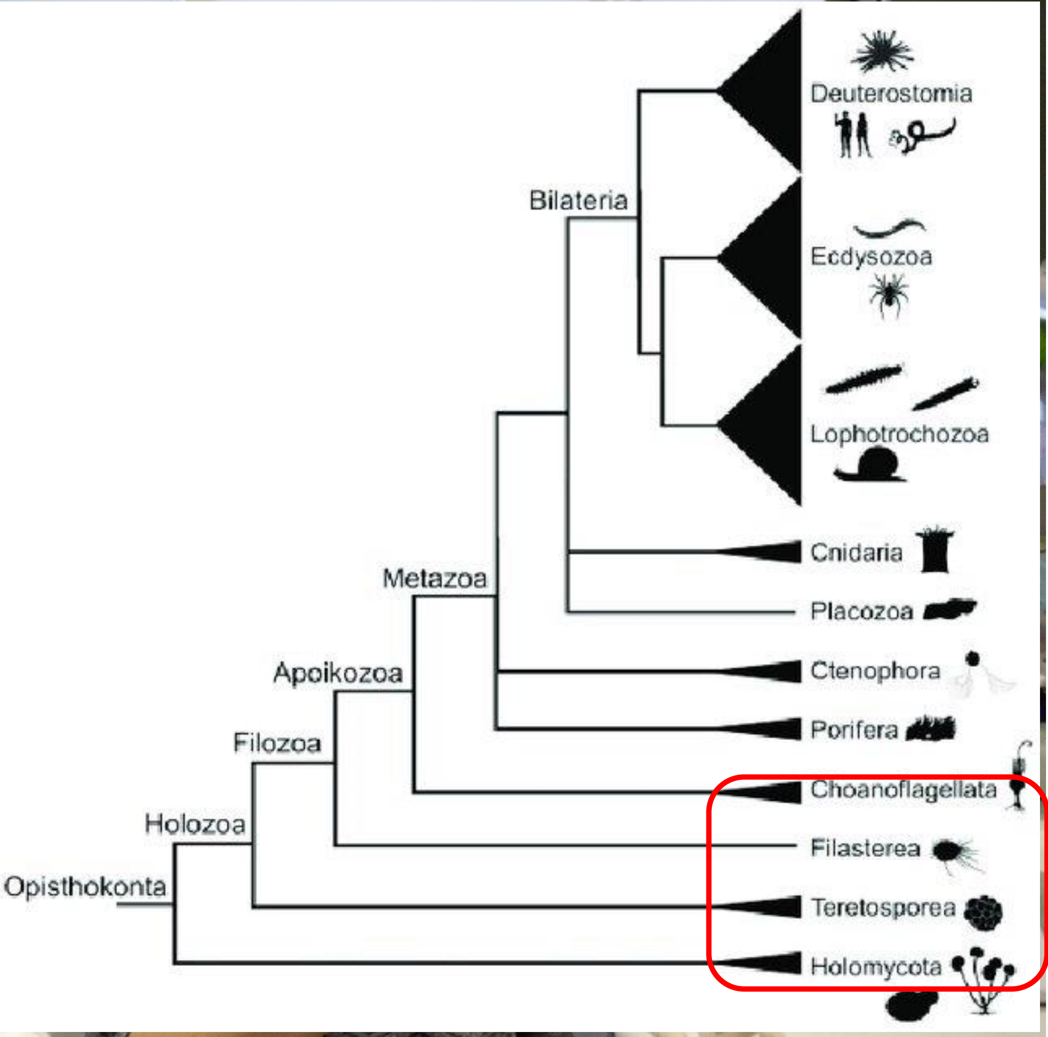
Publication	Phylogenetic position	Dataset	Statistical support
Carranza et al. (1997)	First-splitting bilaterians	18S rDNA	67 % NJ BS^{a}
Ruiz-Trillo et al. (1999)	First-splitting bilaterians	18S rDNA	*100 % ML BS*^{b}
Ruiz-Trillo et al. (2002)	First-splitting bilaterians	18S rDNA, myosin heavy chain type II	*1.0/1.0 BI BPP*^{c}
Jondelius et al. (2002)	First-splitting bilaterians	18S rDNA	*92 %/82 % ML BS*^{c}
Telford et al. (2003)	First-splitting bilaterians	18S rDNA, 28S rDNA	*1.0/1.0 BI BPP*^{c}
Ruiz-Trillo et al. (2004)	First-splitting bilaterians	mtDNA	*97 % QP*^{d}
Glenner et al. (2004)	First-splitting bilaterians	18S rDNA; morphology	*1.0/1.0 BI BPP*^{c}
Wallberg et al. (2007)	First-splitting bilaterians	18S rDNA, 28S rDNA	*1.0 BI BPP*
Philippe et al. (2007)	Deuterostomes	68 protein-coding genes	0.34 BI BPP^{e}
Dunn et al. 2008	Lophotrochozoa	150 protein-coding genes	–
Paps et al. (2009a)	First-splitting bilaterians	18S rDNA, 28S rDNA	*1.0/1.0 BI BPP*^{c}
Paps et al. (2009b)	First-splitting bilaterians	13 nuclear markers	*1.0/1.0 BI BPP*^{c}
Hejnol et al. (2009)	First-splitting bilaterians	1500 protein-coding genes	66 % ML BS^{f}
Mwinyi et al. (2010)	First-splitting bilaterians	mtDNA	*0.91 BI BPP*^{d}
Mallatt et al. (2010)	First-splitting bilaterians	18S rDNA, 28S rDNA	*>90 % ML BS*^{d}
Philippe et al. (2011a)	Deuterostomes	197 protein-coding genes and mtDNA	63 % BI BS^{g} and *0.99 BI BPP*
Ryan et al. (2013)	First-splitting bilaterians	242 protein-coding genes	38–52 % ML BS^{h}
Srivastava et al. (2014)	First-splitting bilaterians	Different sets, ranging from 69 to 442	96 % ML BS and *1.0 BI-WAG BPP*^{i}

Ruiz-Trillo and Paps 2015

UB Parc Científic
Barcelona

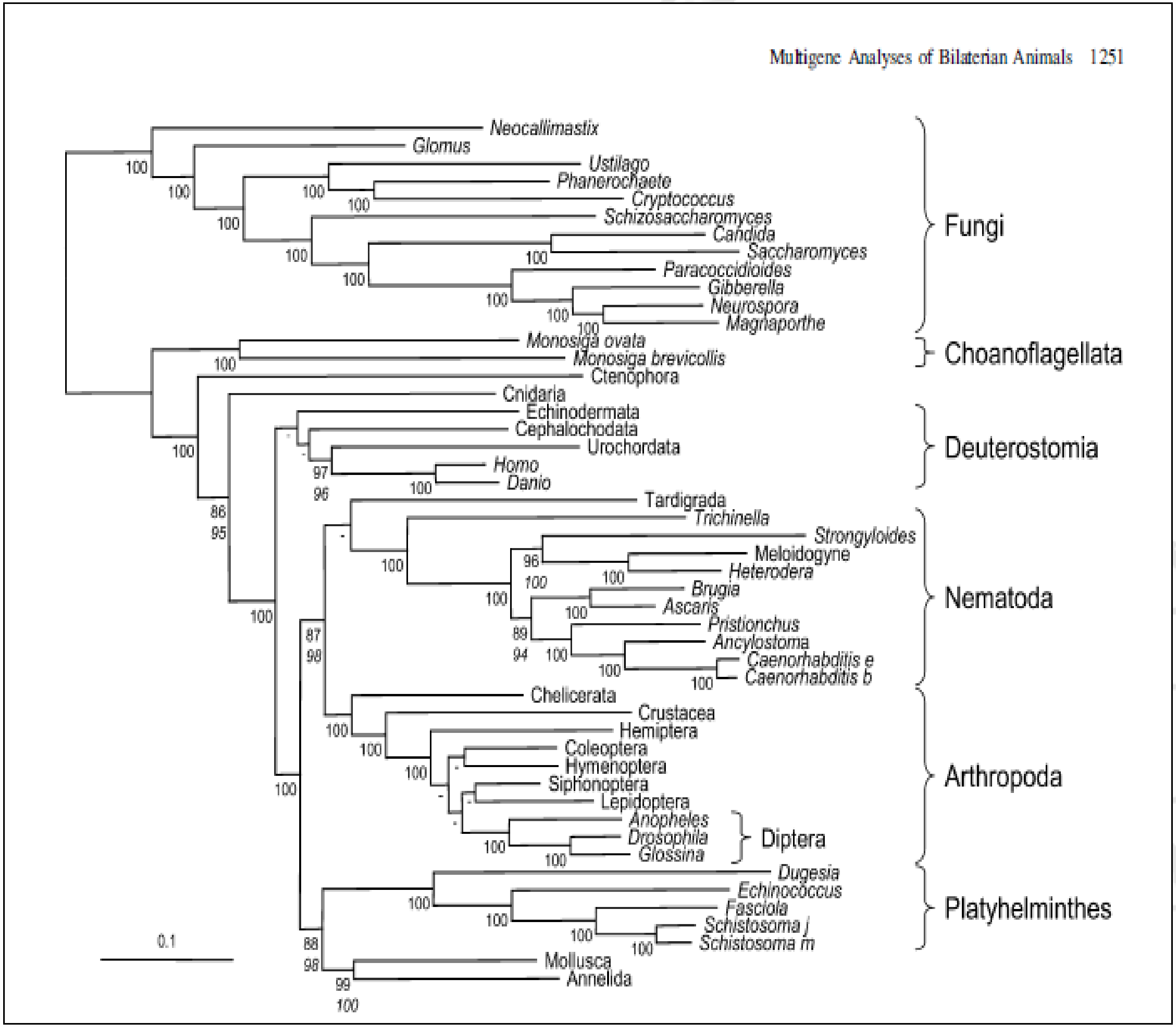






PHYLOTRANSCRIPTOMICS

Phylogenomics


Multigene Analyses of Bilaterian Animals 1251

Philippe 2005
146 genes
35,000 chars
10 phyla

Phylogenomics 2008

Broad phylogenomic sampling improves resolution of the animal tree of life

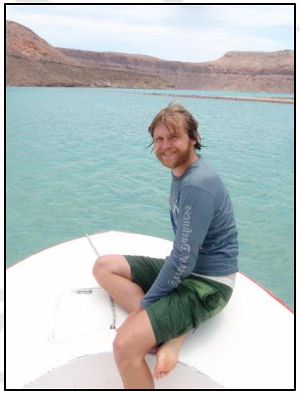
Casey W. Dunn[1]†, Andreas Hejnol[1], David Q. Matus[1], Kevin Pang[1], William E. Browne[1], Stephen A. Smith[2], Elaine Seaver[1], Greg W. Rouse[3], Matthias Obst[4], Gregory D. Edgecombe[5], Martin V. Sørensen[6], Steven H. D. Haddock[7], Andreas Schmidt-Rhaesa[8], Akiko Okusu[9], Reinhardt Møbjerg Kristensen[10], Ward C. Wheeler[11], Mark Q. Martindale[1] & Gonzalo Giribet[12,13]

Dunn et al. (2008)

150 genes or 21,000 chars

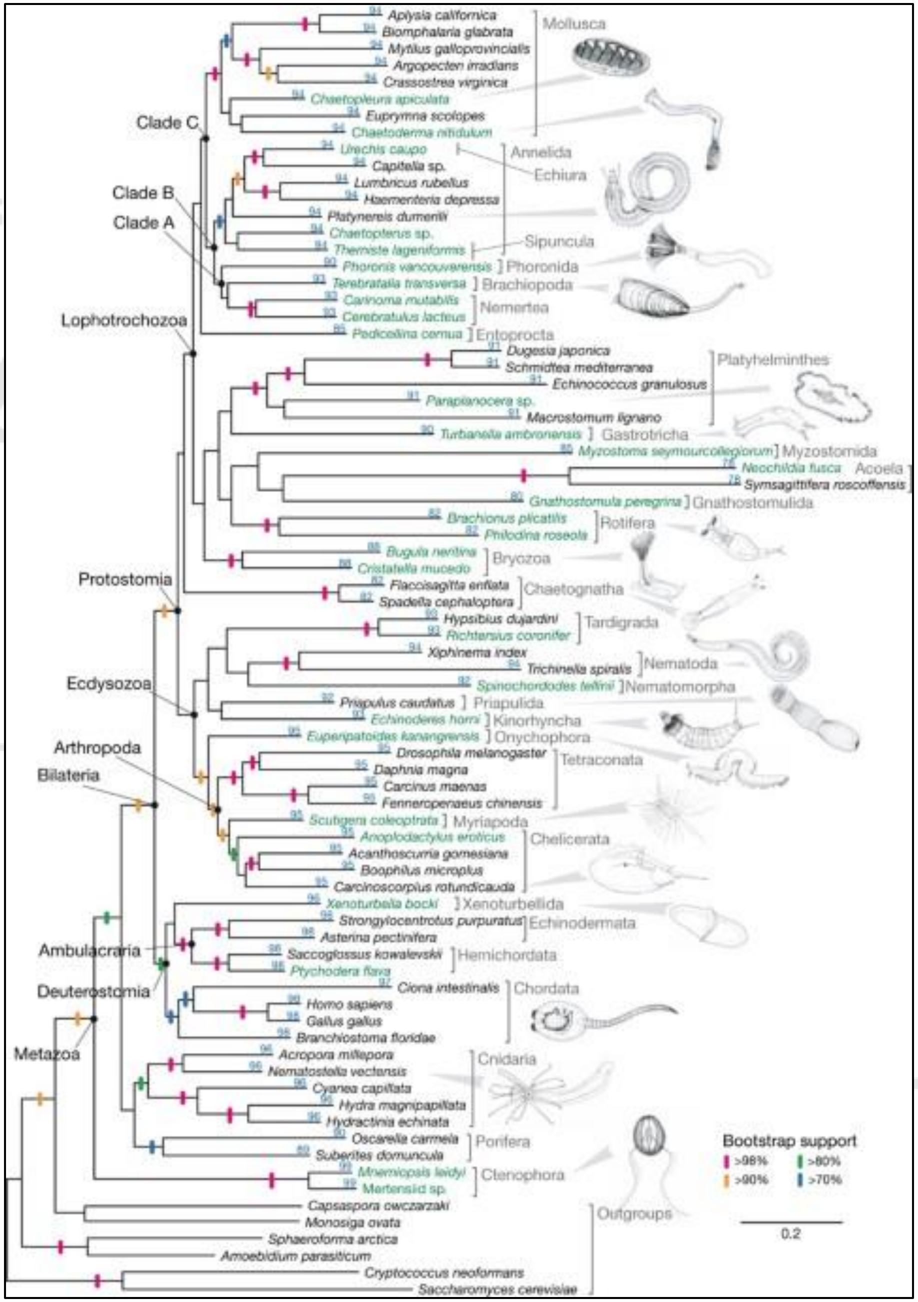
26 phyla

Gonzalo Giribet

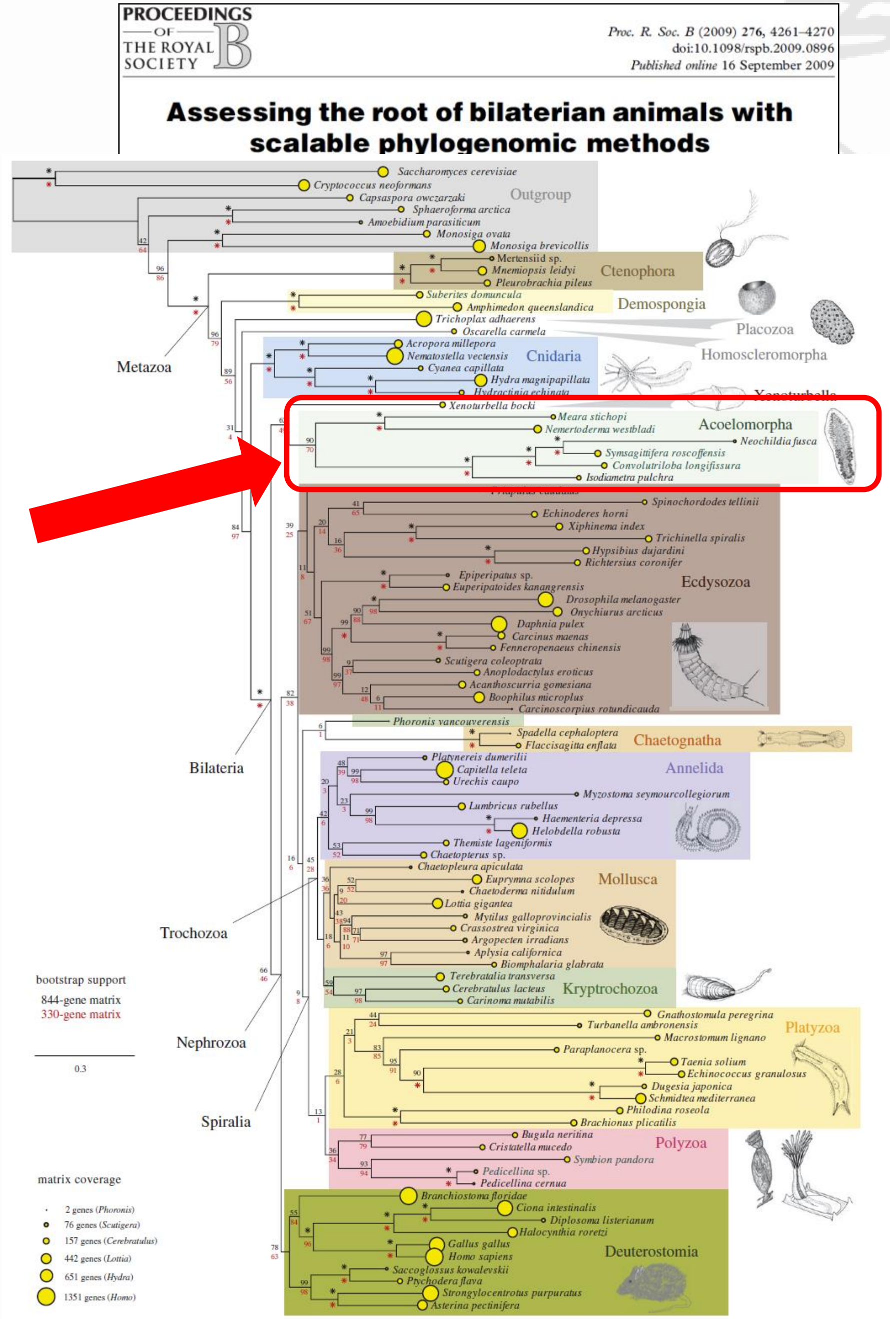


Casey W. Dunn

Andreas Hejnol



Phylogenomics: Acoelomorpha



PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2009) 276, 4261–4270
doi:10.1098/rspb.2009.0896
Published online 16 September 2009

Assessing the root of bilaterian animals with scalable phylogenomic methods

Hejnol et al. (2009)

1,500 genes or 270,000 chars

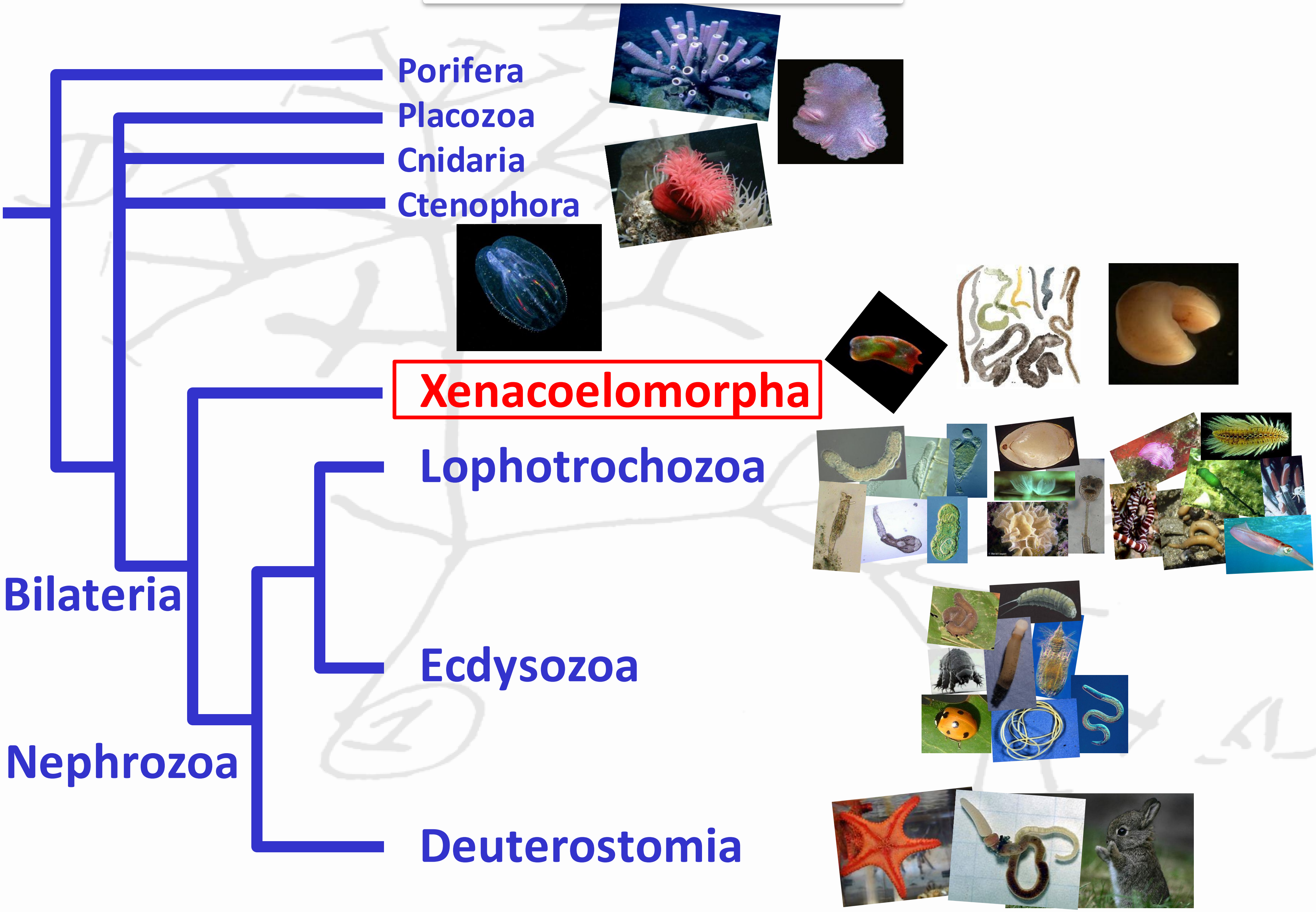
27 phyla

Acoels, nemertodermatids, and xenoturbellids

Acoela + Nemertodermatida = Acoelomorpha

Xenoturbellida +Acoelomorpha = Xenacelomorpha

Nephrozoa



Porifera

Placozoa

Cnidaria

Ctenophora

Xenacoelomorpha

Lophotrochozoa

Ecdysozoa

Deuterostomia

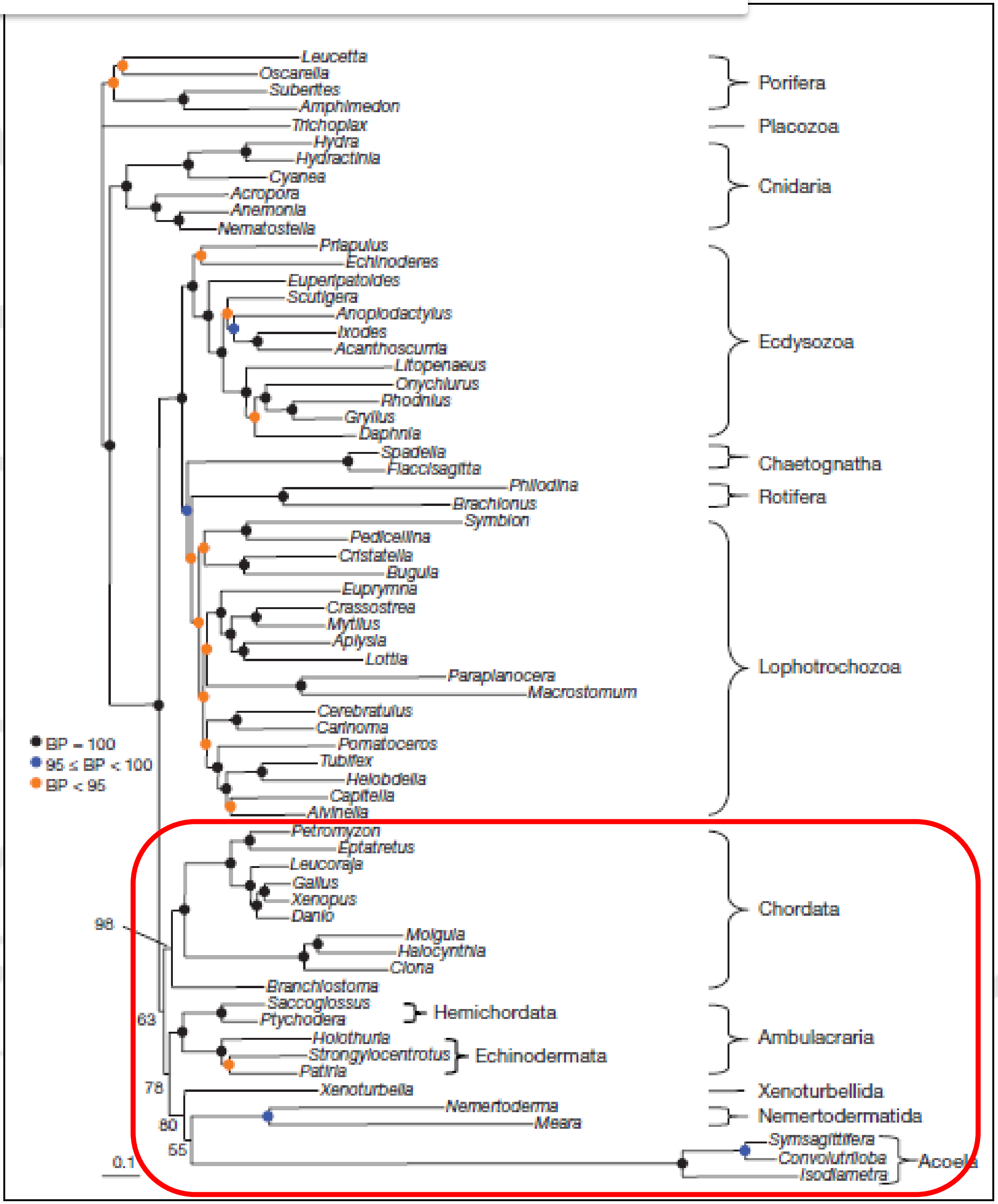
Bilateria

Nephrozoa

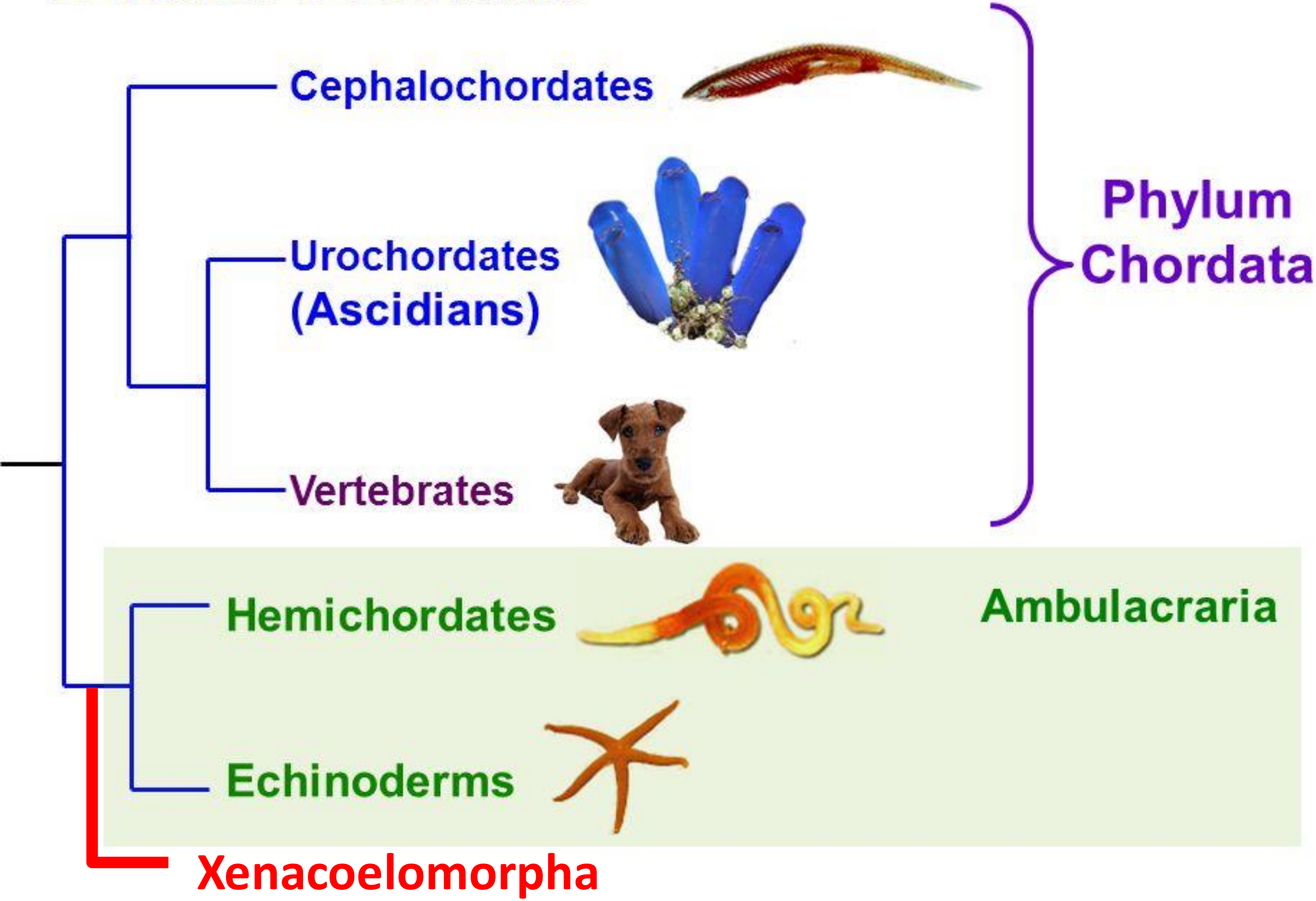
Xenambulacraria 2011

Hejnol et al, 2009
1,500 genes
270,000 chars

Philippe et al, 2011
197 markers
38,000 chars
CAT model



Deuterostomia



Cephalochordates

Urochordates (Ascidians)

Vertebrates

Phylum Chordata

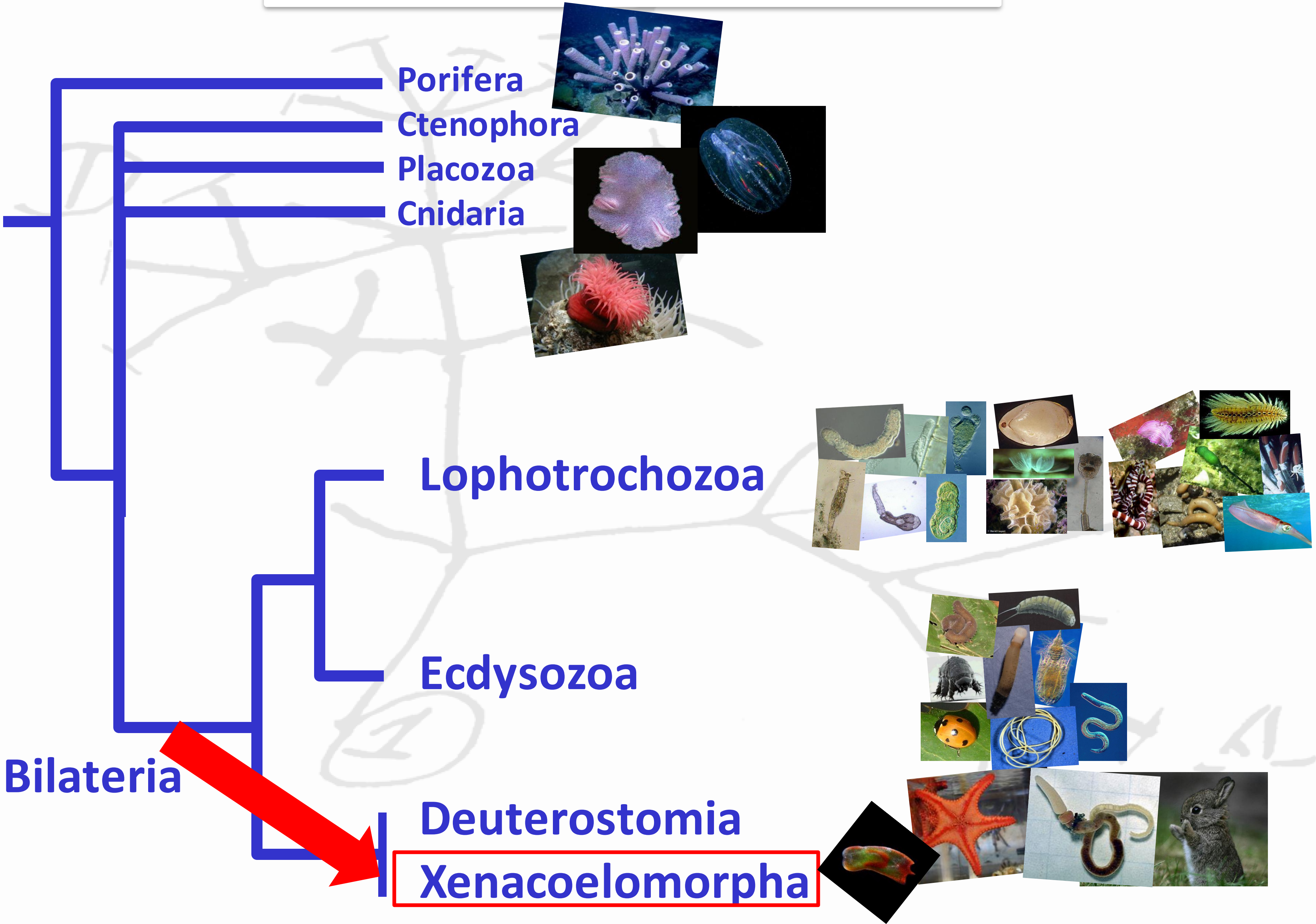
Hemichordates

Echinoderms

Ambulacraria

Xenacoelomorpha

Xenambulacraria 2011



Porifera

Ctenophora

Placozoa

Cnidaria

Lophotrochozoa

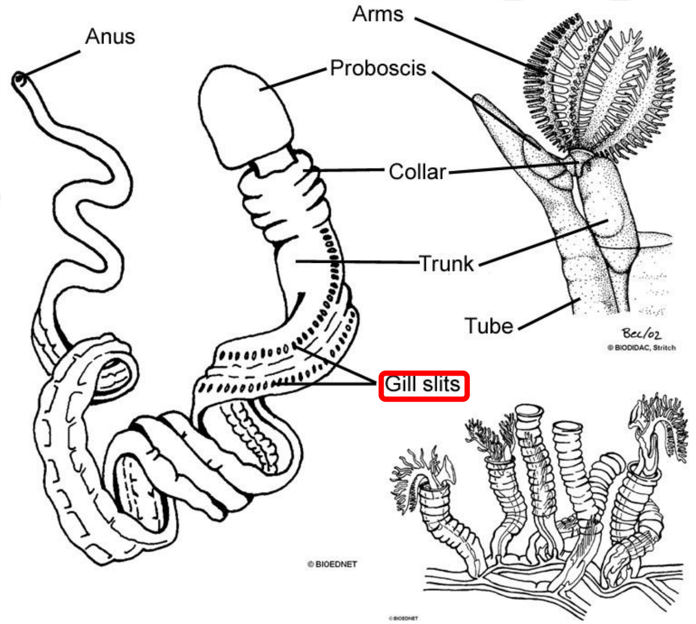
Ecdysozoa

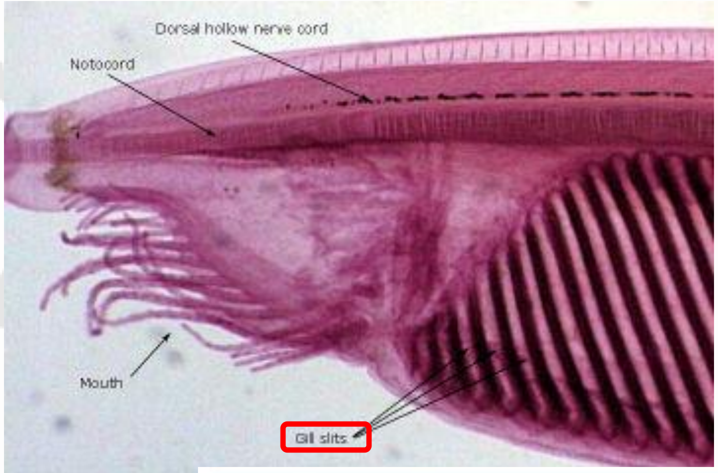
Bilateria

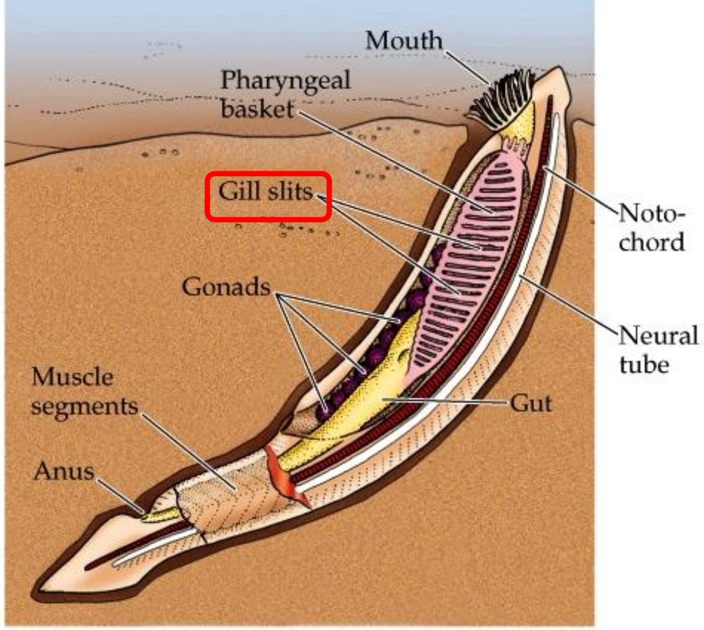
Deuterostomia

Xenacoelomorpha

Deuterostome synapomorphy: Gill slits







© 2001 Sinauer Associates, Inc.

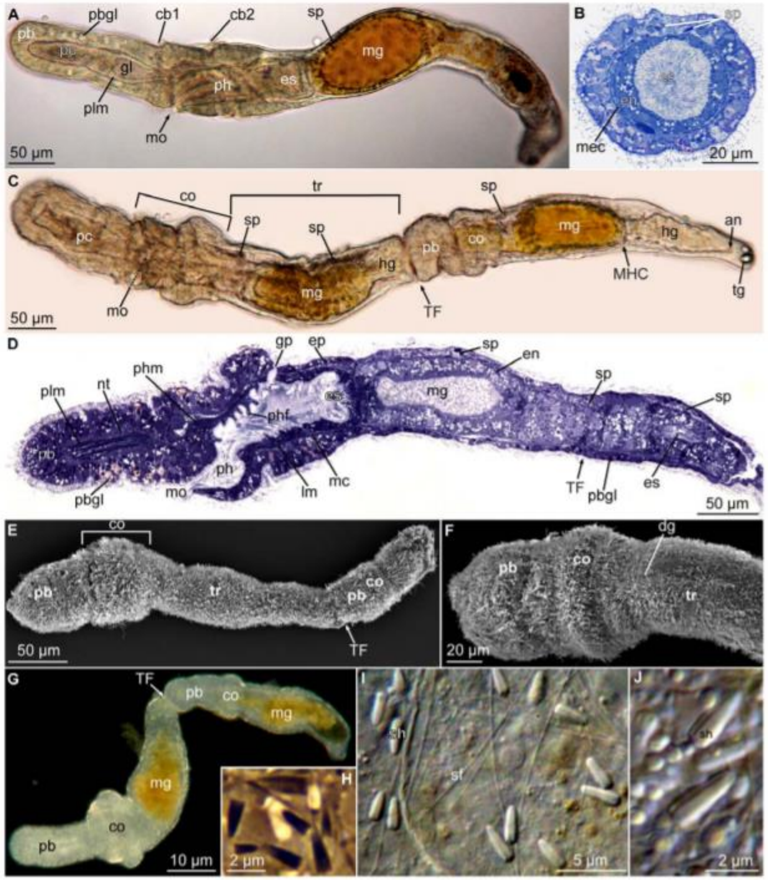
Miniature deuterostomes still hace coelom


OPEN ACCESS Freely available online

PLOS | ONE

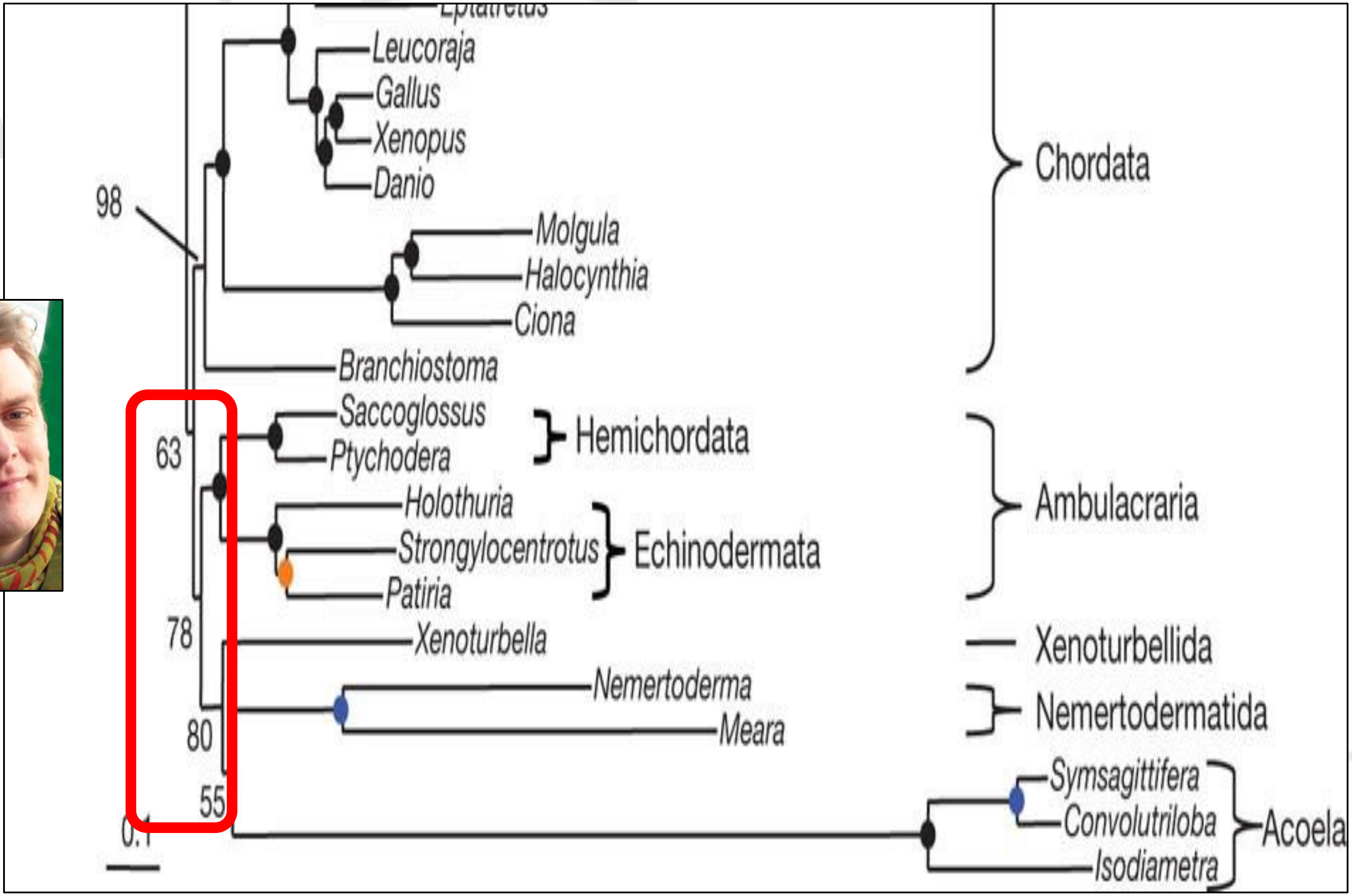
An Anatomical Description of a Miniaturized Acorn Worm (Hemichordata, Enteropneusta) with Asexual Reproduction by Paratomy

Katrine Worsaae[1]*, Wolfgang Sterrer[2], Sabrina Kaul-Strehlow[3], Anders Hay-Schmidt[4], Gonzalo Giribet[5]



Xenambulacraria 2011

Philippe et al, 2011
197 markers
38,000 chars
30% missing data
CAT model



Xenacoelomorpha 2016

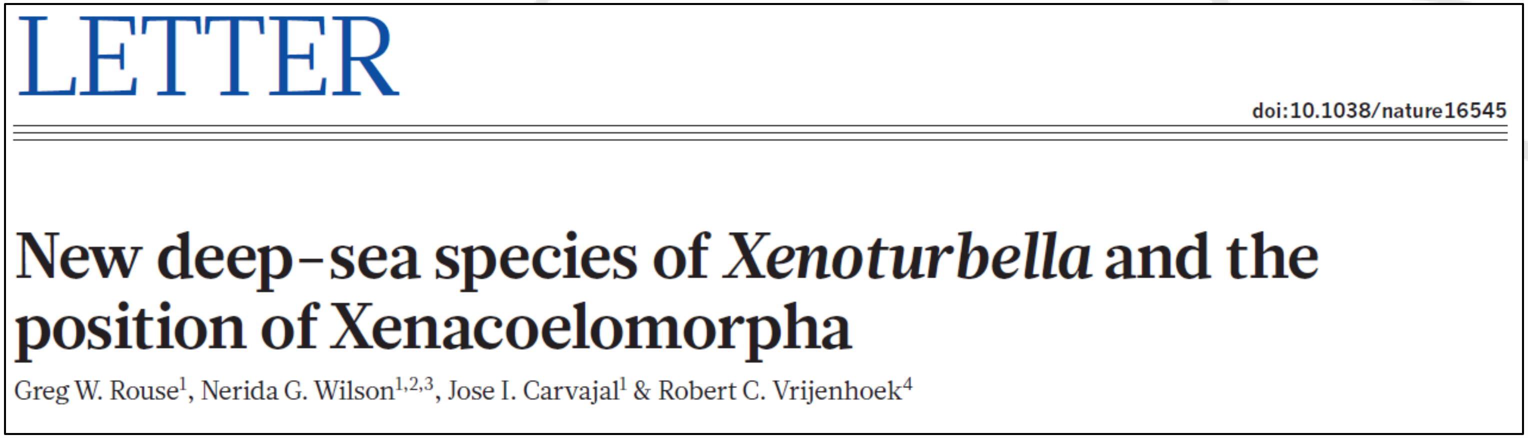
Cannon et al, 2016
217 markers
45,000 chars
CAT-GTR model

LETTER

doi:10.1038/nature16520

Xenacoelomorpha is the sister group to Nephrozoa

Johanna Taylor Cannon[1], Bruno Cossermelli Vellutini[2], Julian Smith III[3], Fredrik Ronquist[1], Ulf Jondelius[1] & Andreas Hejnol[2]



LETTER

doi:10.1038/nature16545

New deep-sea species of *Xenoturbella* and the position of Xenacoelomorpha

Greg W. Rouse[1], Nerida G. Wilson[1,2,3], Jose I. Carvajal[1] & Robert C. Vrijenhoek[4]


a
10 cm
b
Es
Ad
Ad
5 cm
c
anterior
st
rf
m
1 cm
d
1 cm
e
anterior
rf
st
st
v
oo
5 mm
f
m
rf
vgn
oo
posterior
5 mm
g
vgn
3 mm
h
rf
m
vgn
3 mm


Science & Environment

Mystery of deep-sea 'purple sock' solved

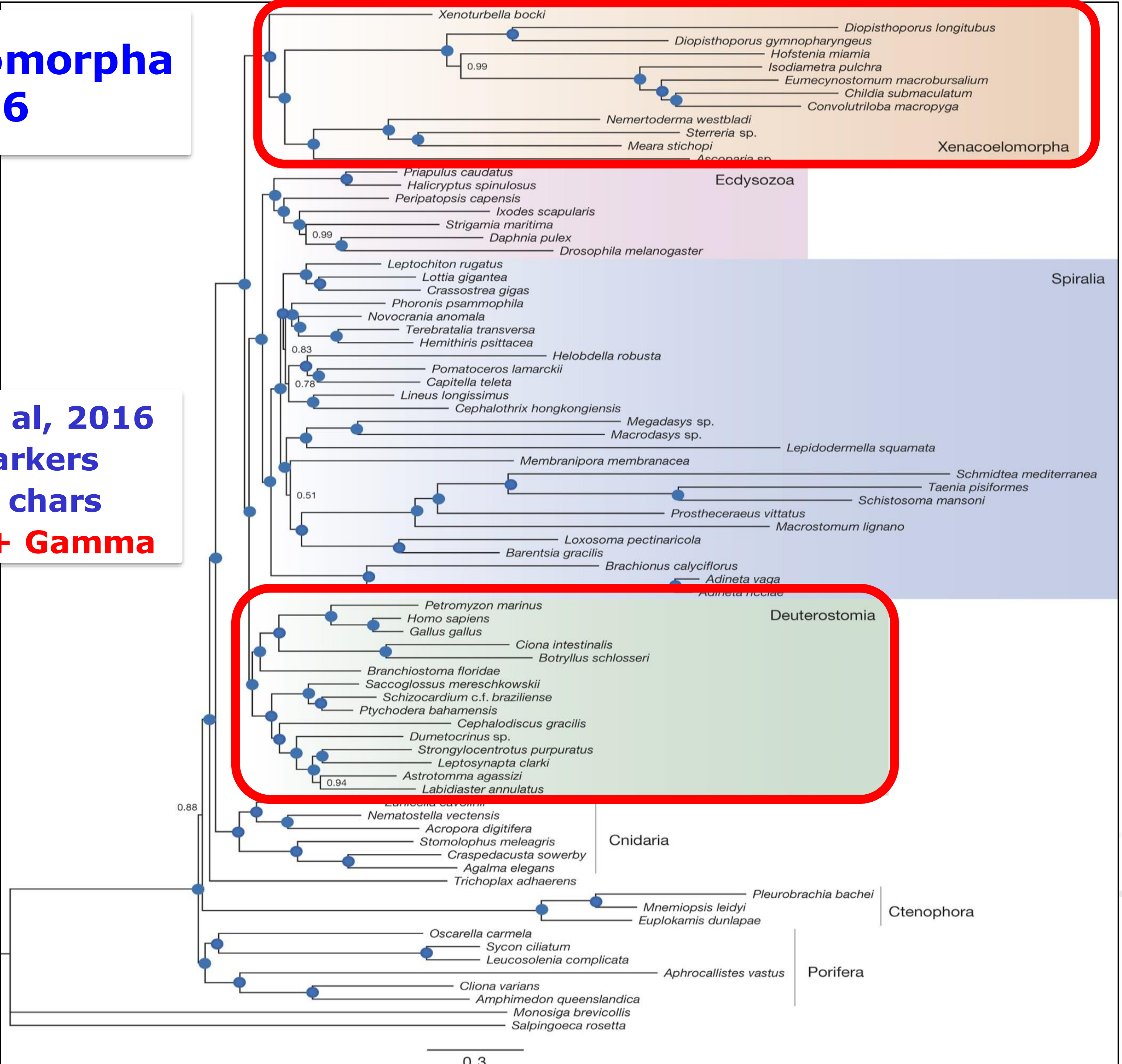
By Rebecca Morelle
Science Correspondent, BBC News

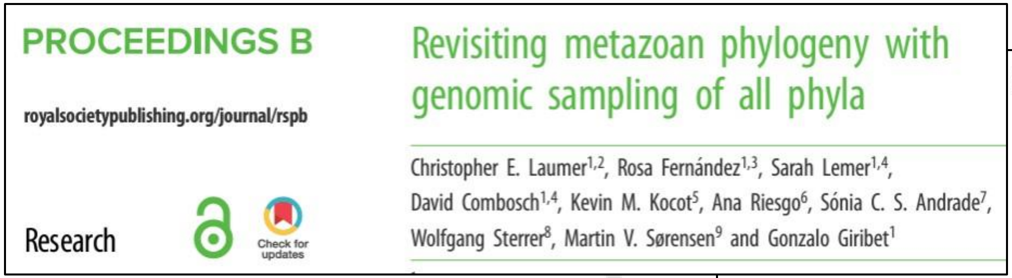
3 February 2016 | Science & Environment


Research Highlight: Deep-Sea Churro Finds its Evolutionary Home

Xenacoelomorpha 2016

Cannon et al, 2016
217 markers
45,000 chars
CAT-GTR + Gamma





PROCEEDINGS B

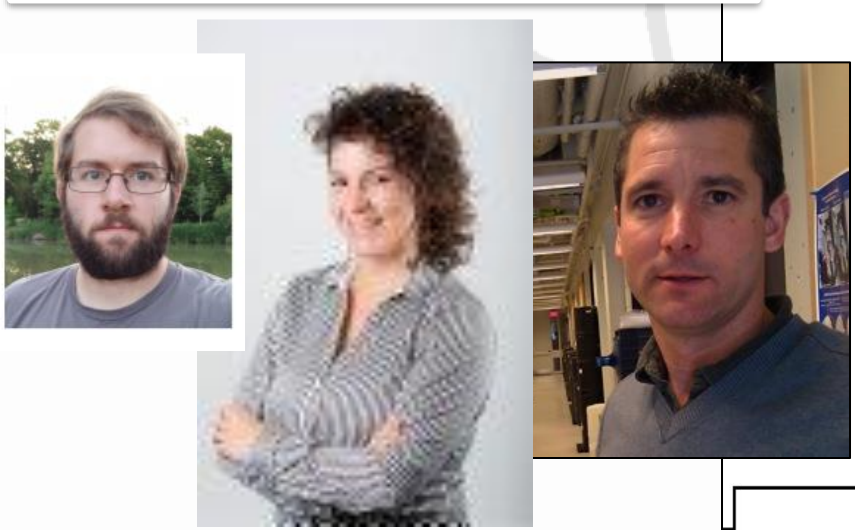
royalsocietypublishing.org/journal/rspb

Research

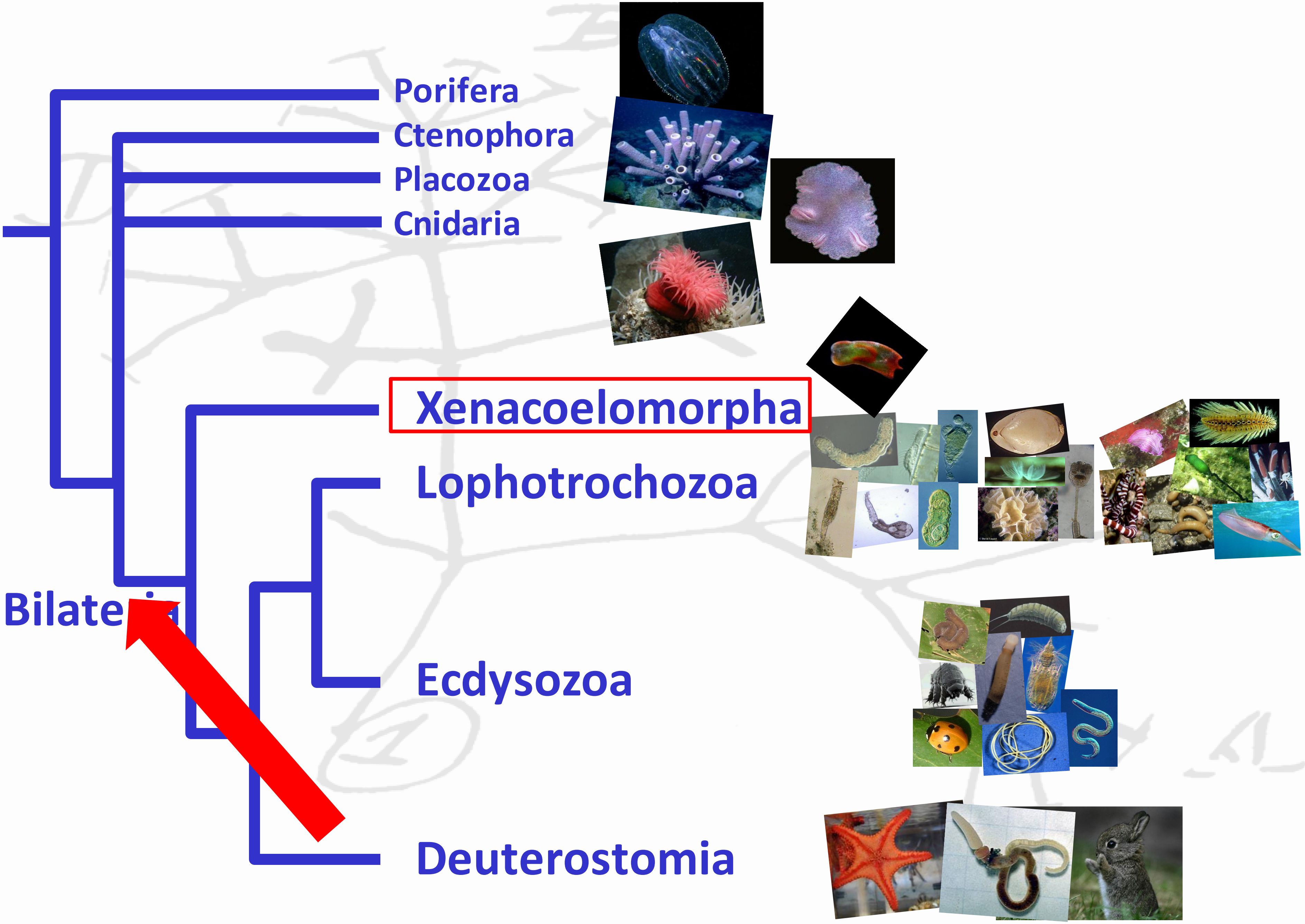
Revisiting metazoan phylogeny with genomic sampling of all phyla

Christopher E. Laumer[1,2], Rosa Fernández[1,3], Sarah Lemer[1,4], David Combosch[1,4], Kevin M. Kocot[5], Ana Riesgo[6], Sónia C. S. Andrade[7], Wolfgang Sterrer[8], Martin V. Sørensen[9] and Gonzalo Giribet[1]

Laumer et al, 2019
422 markers
53,000 chars
60% completeness
167 taxa
Almost all phyla
CAT-GTR + Gamma model



Xenacoelomorpha 2016-19



Porifera

Ctenophora

Placozoa

Cnidaria

Xenacoelomorpha

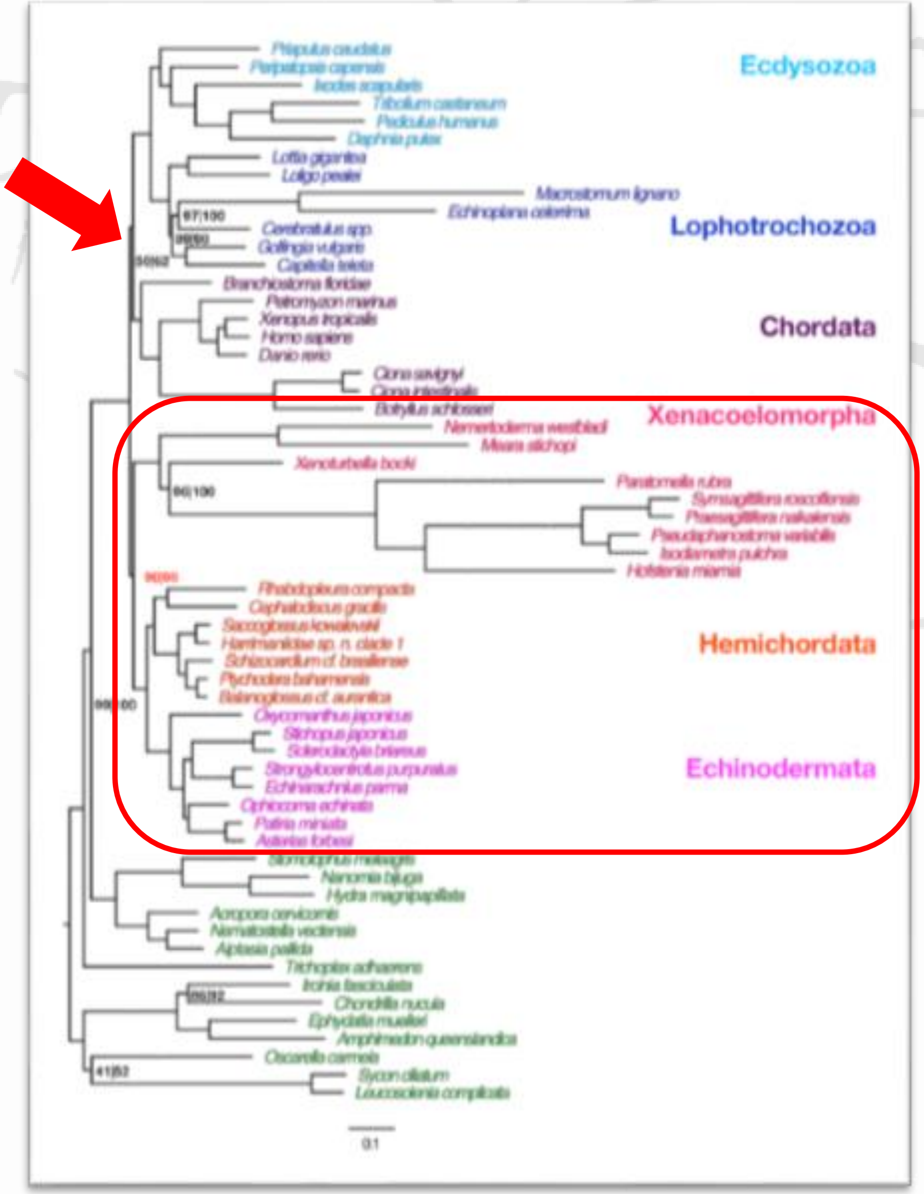
Lophotrochozoa

Ecdysozoa

Deuterostomia

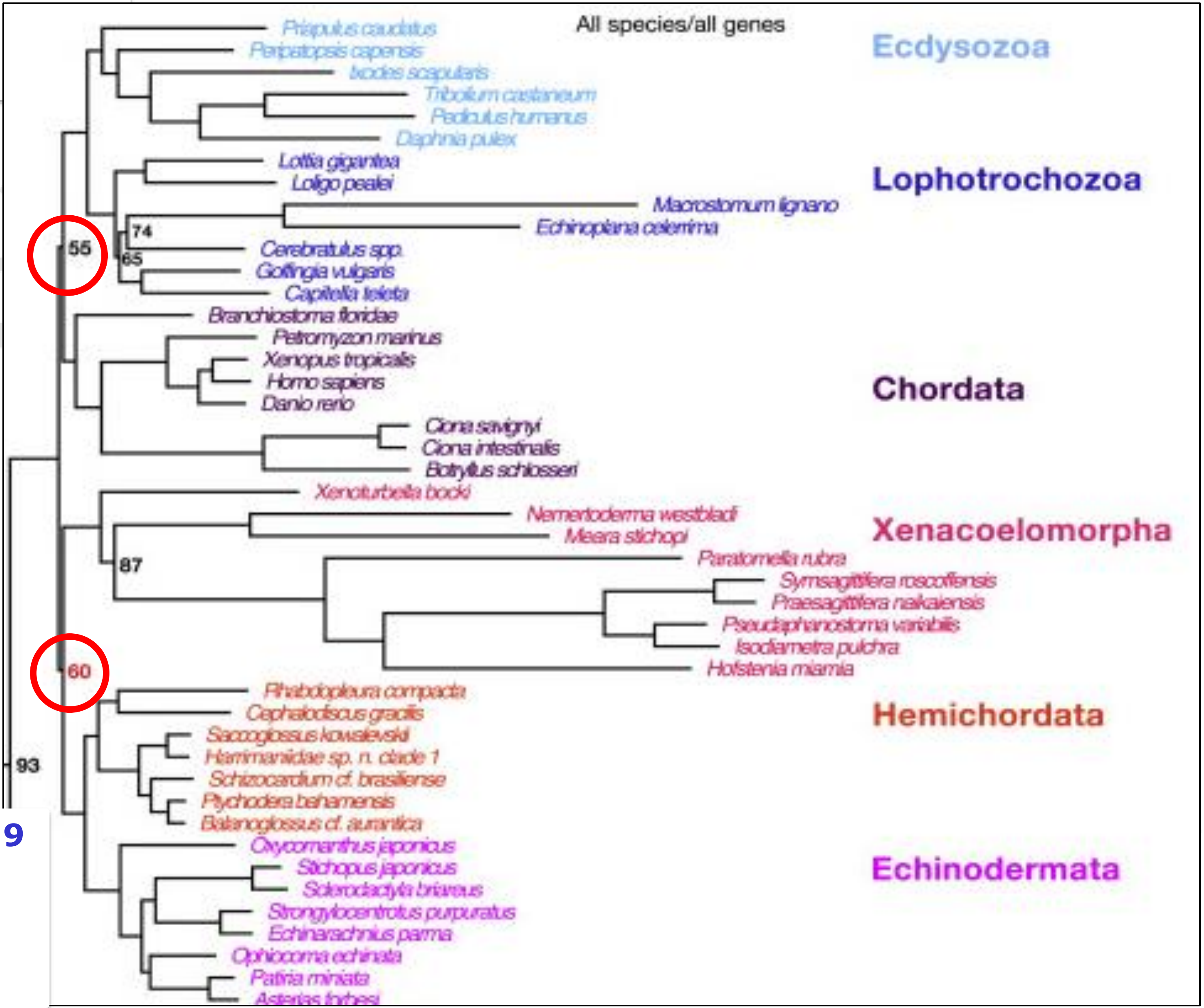
Bilate ia

Xenambulacraria 2019
Paraphyletic deuterostomes

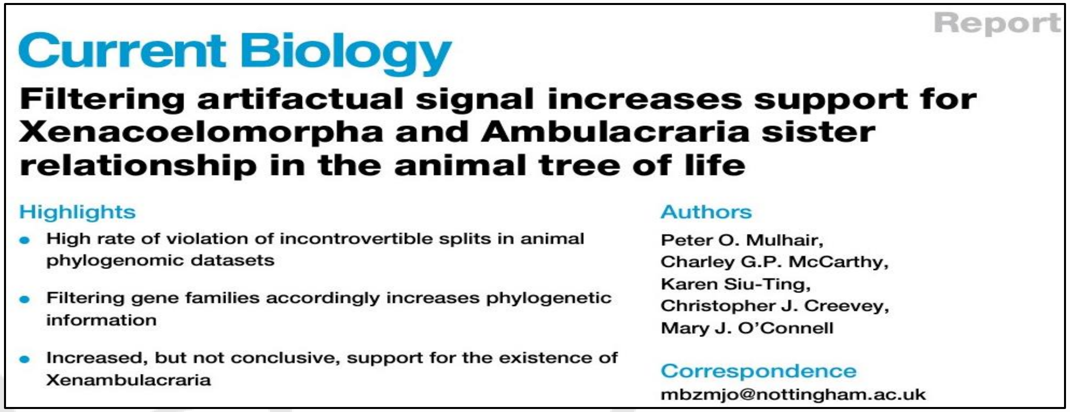


Philippe et al, 2019
1,173 markers
350,000 chars
CAT-GTR model

Xenambulacraria 2019



Philippe et al, 2019
1,173 markers
350,000 chars
CAT-GTR model



Report

Current Biology

Filtering artifactual signal increases support for Xenacoelomorpha and Ambulacraria sister relationship in the animal tree of life

Highlights

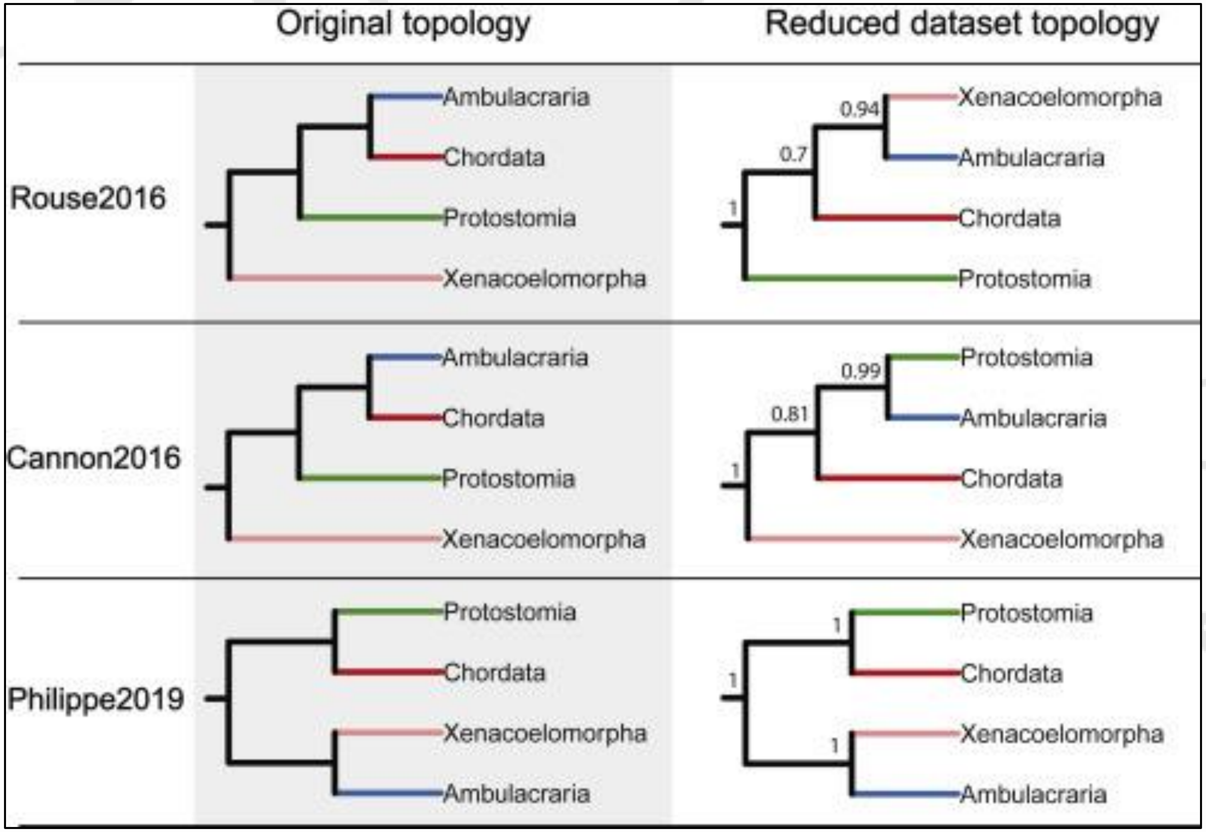
- High rate of violation of incontrovertible splits in animal phylogenomic datasets
- Filtering gene families accordingly increases phylogenetic information
- Increased, but not conclusive, support for the existence of Xenambulacraria

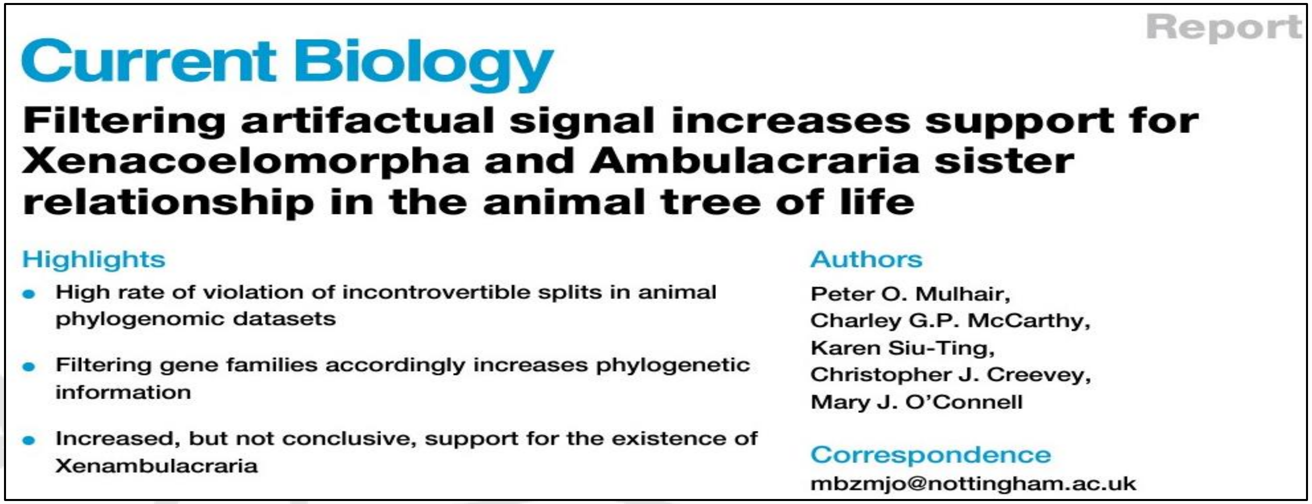
Authors

Peter O. Mulhair,
Charley G.P. McCarthy,
Karen Siu-Ting,
Christopher J. Creevey,
Mary J. O'Connell

Correspondence

mbzmjo@nottingham.ac.uk





Report

Current Biology

Filtering artifactual signal increases support for Xenacoelomorpha and Ambulacraria sister relationship in the animal tree of life

Highlights

- High rate of violation of incontrovertible splits in animal phylogenomic datasets
- Filtering gene families accordingly increases phylogenetic information
- Increased, but not conclusive, support for the existence of Xenambulacraria

Authors

Peter O. Mulhair,
Charley G.P. McCarthy,
Karen Siu-Ting,
Christopher J. Creevey,
Mary J. O'Connell

Correspondence

mbzmjo@nottingham.ac.uk

The slowly evolving genome of the xenacoelomorph worm *Xenoturbella bocki*

Philipp H. Schiffer[1,2,*], Paschalis Natsidis[1], Daniel J. Leite[1,3], Helen E. Robertson[1], François Lapraz[1,4], Ferdinand Marlétaz[1], Bastian Fromm[5], Liam Baudry[6], Fraser Simpson[1], Eirik Høye[7,8], Anne-C. Zakrzewski[1,9], Paschalia Kapli[1], Katharina J. Hoff[10,11], Steven Mueller[1,12], Martial Marbouty[13], Heather Marlow[14], Richard R. Copley[15], Romain Koszul[13], Peter Sarkies[16], Maximilian J. Telford[1,*]

*Corresponding authors: p.schiffer@uni-koeln.de, m.telfor

Peer-reviewed & recommended by PCI Genomics

SCIENCE ADVANCES | RESEARCH ARTICLE

EVOLUTIONARY BIOLOGY

Lack of support for Deuterostomia prompts reinterpretation of the first Bilateria

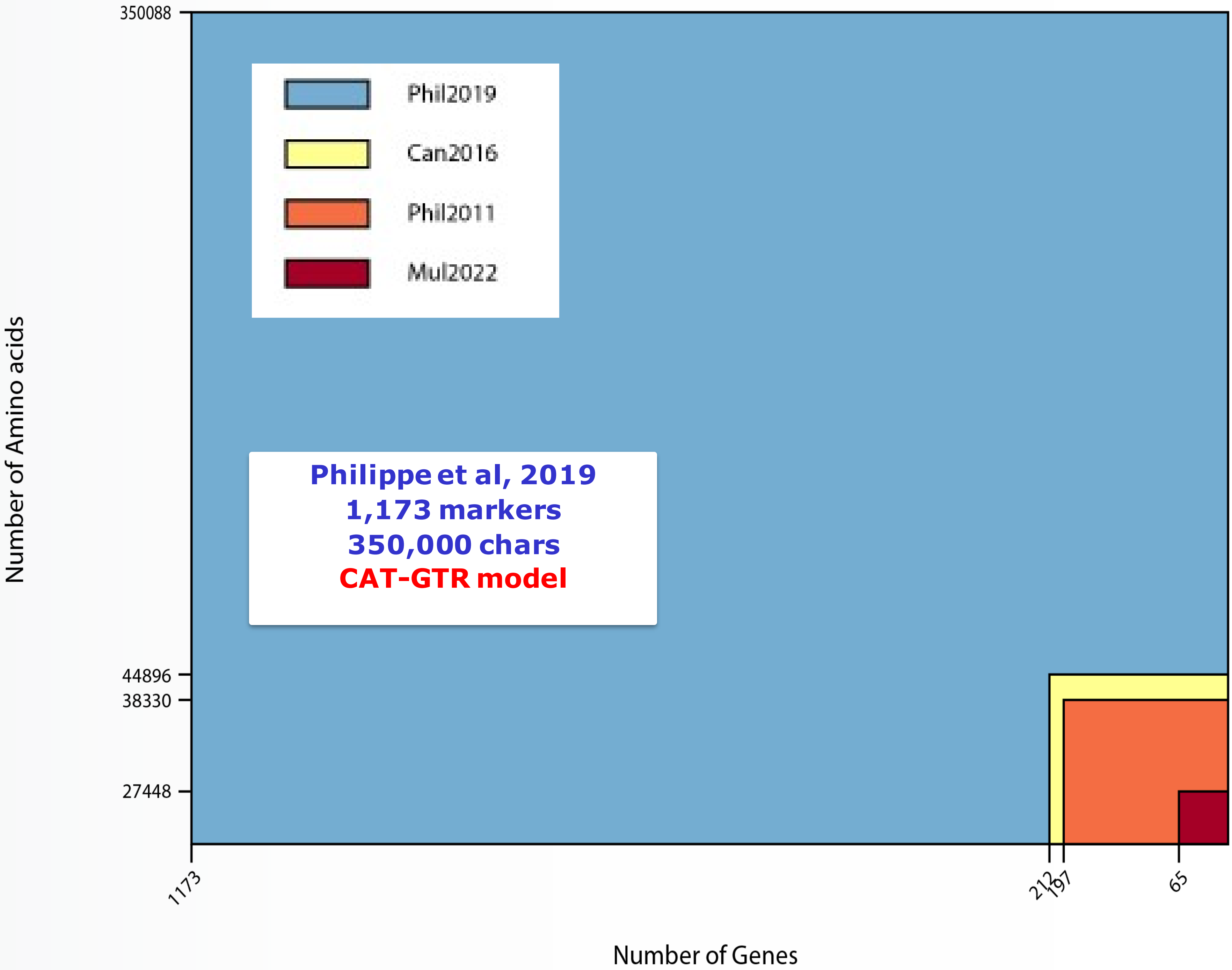
Paschalia Kapli[1], Paschalis Natsidis[1], Daniel J. Leite[1], Maximilian Fursman[1]*, Nadia Jeffrie[1], Imran A. Rahman[2], Hervé Philippe[3], Richard R. Copley[4], Maximilian J. Telford[1†]

SCIENCE ADVANCES | RESEARCH ARTICL

EVOLUTIONARY BIOLOGY

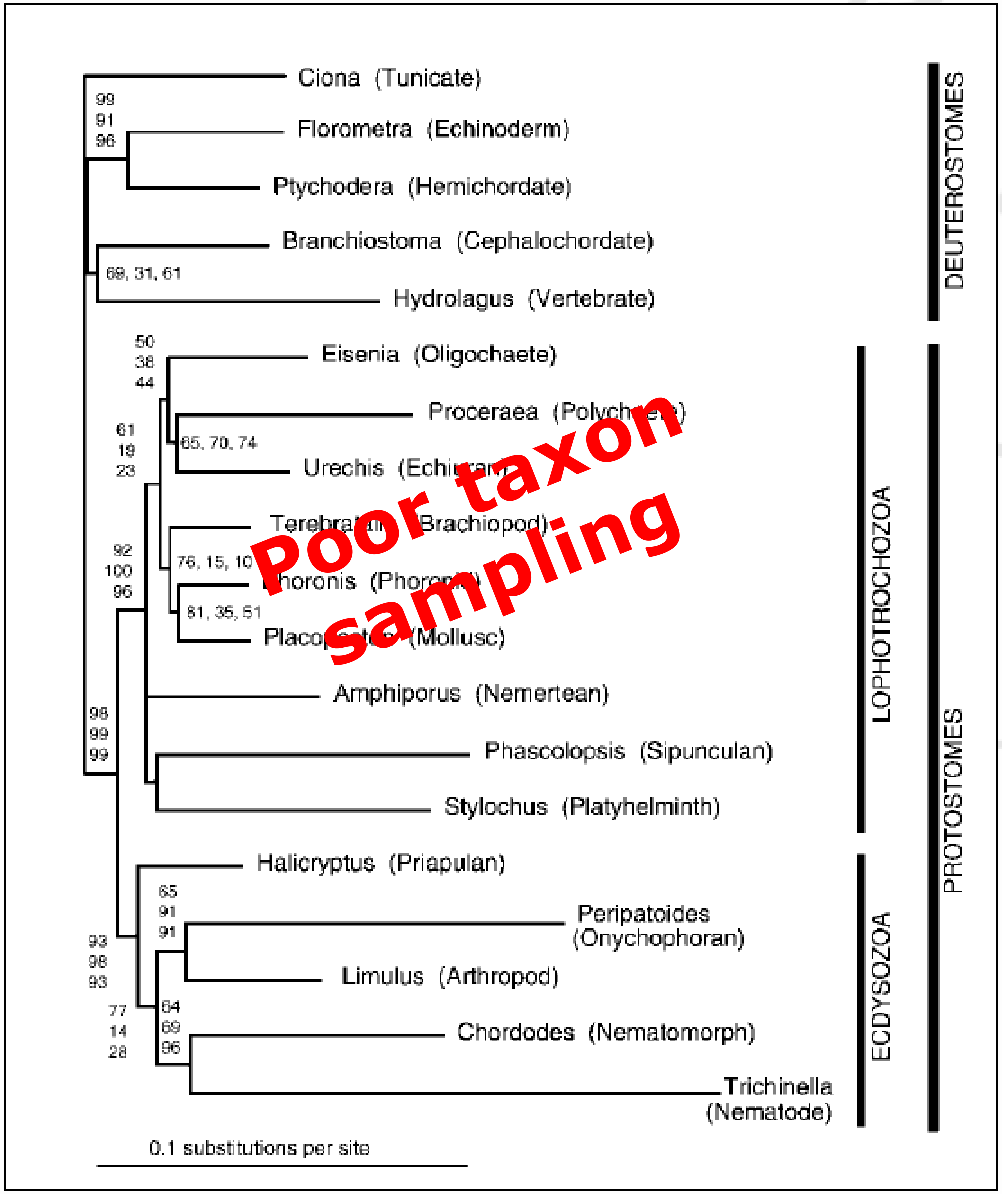
Topology-dependent asymmetry in systematic errors affects phylogenetic placement of Ctenophora and Xenacoelomorpha

Paschalia Kapli and Maximilian J. Telford*

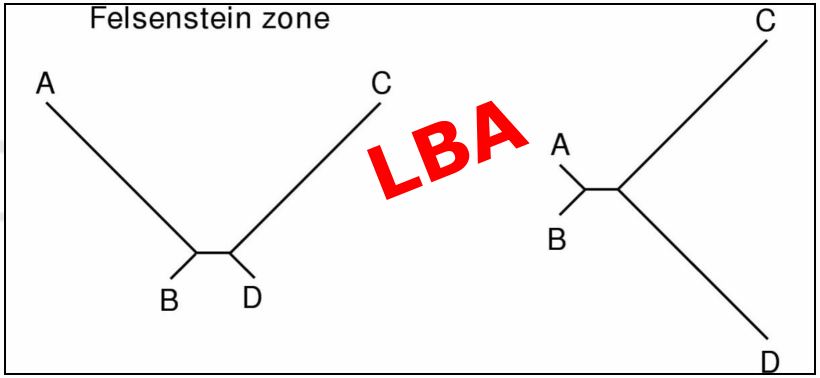


Philippe et al, 2019
1,173 markers
350,000 chars
CAT-GTR model

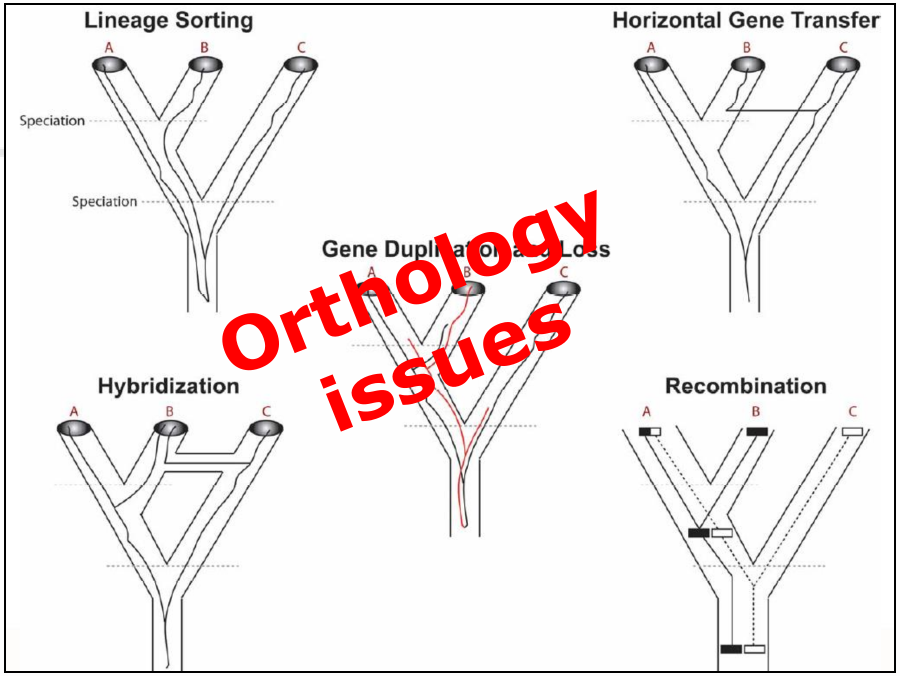
Problems



Poor taxon sampling

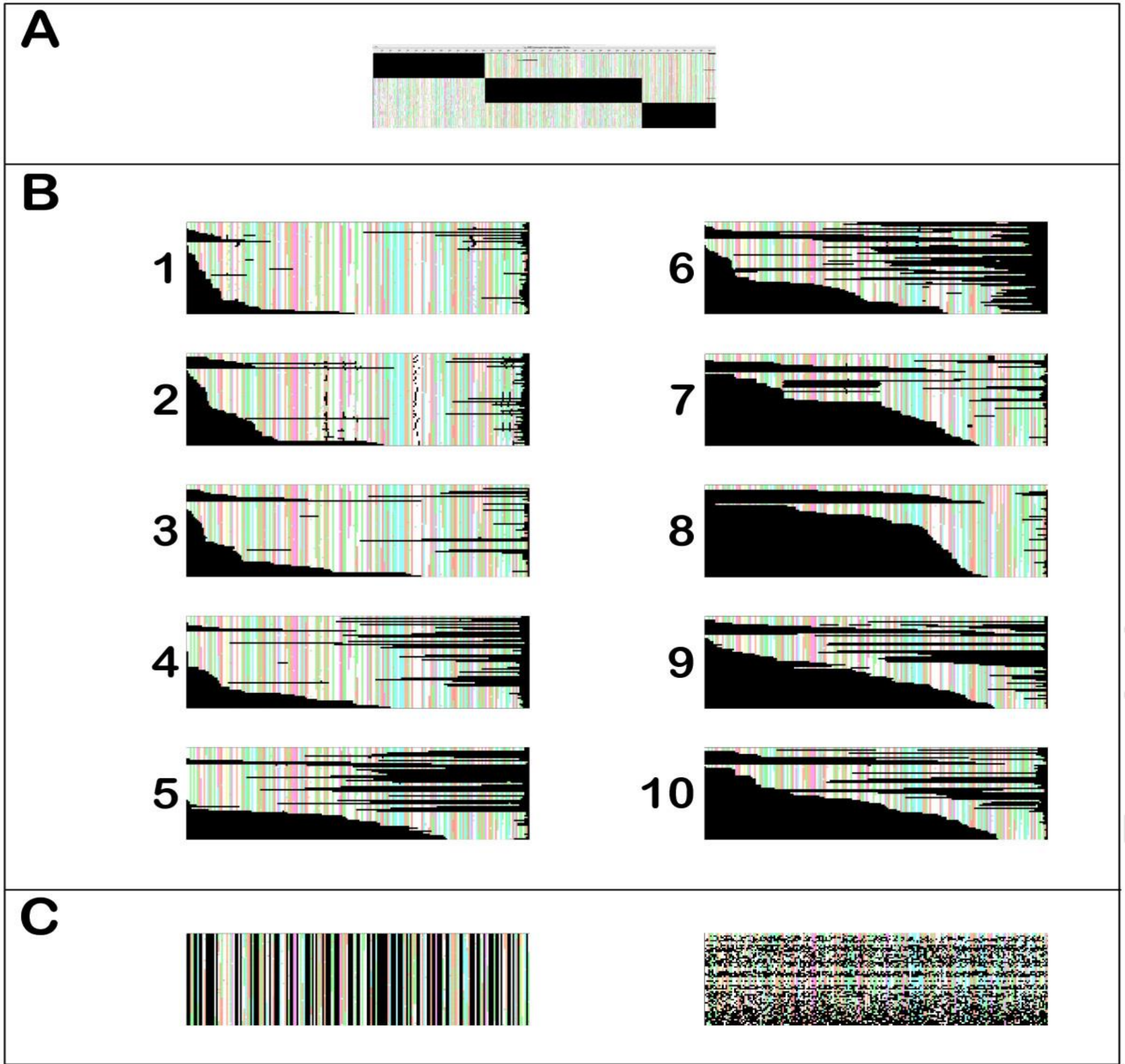


LBA



Orthology issues

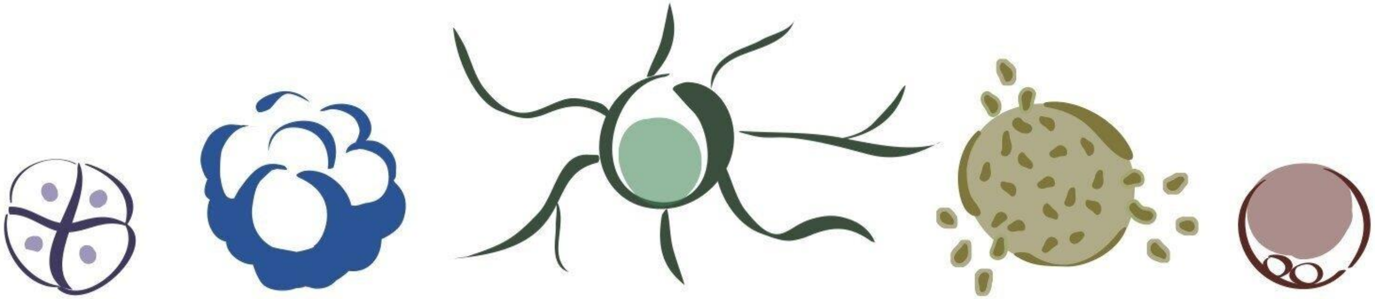
Patchy matrices

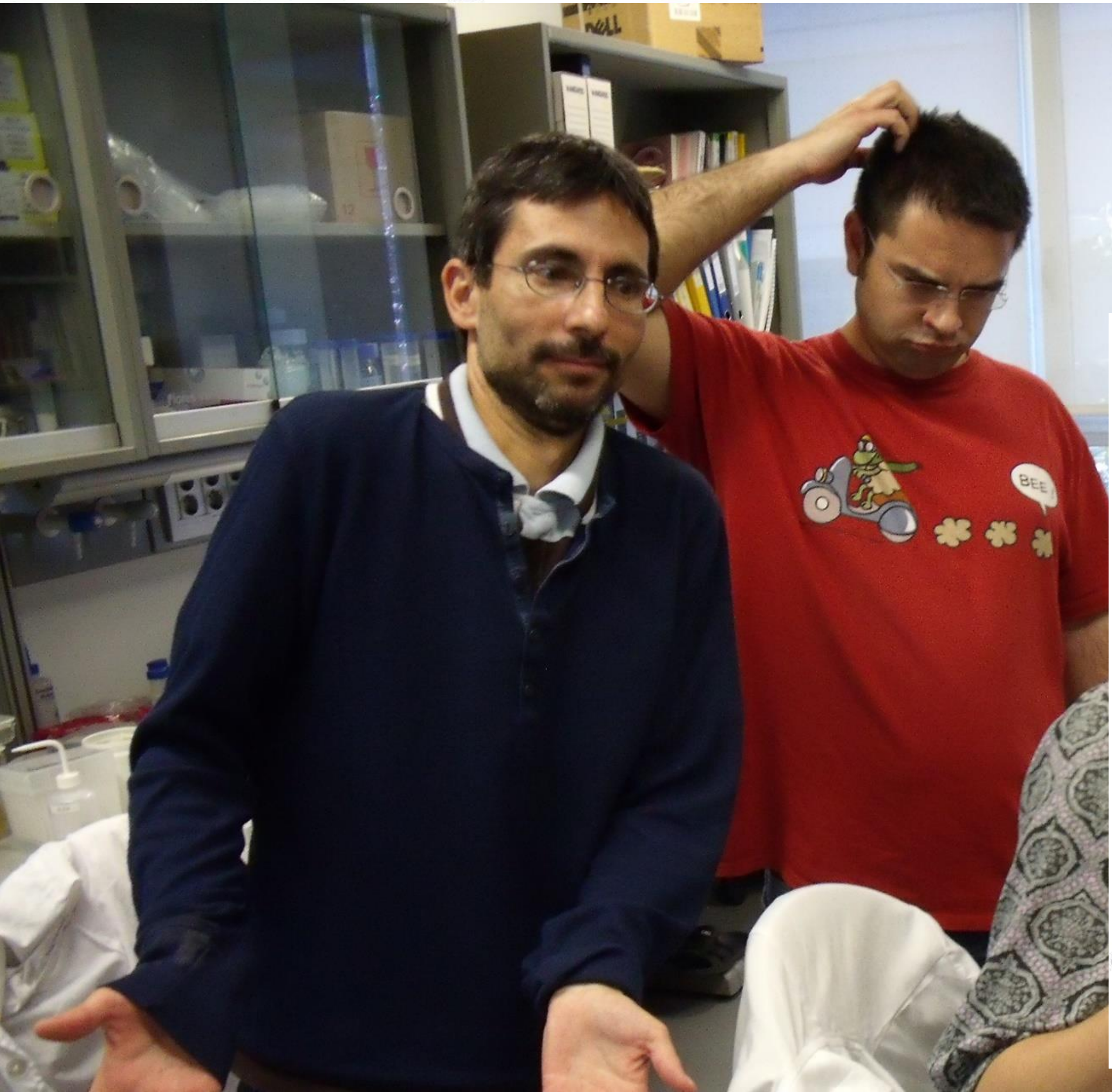


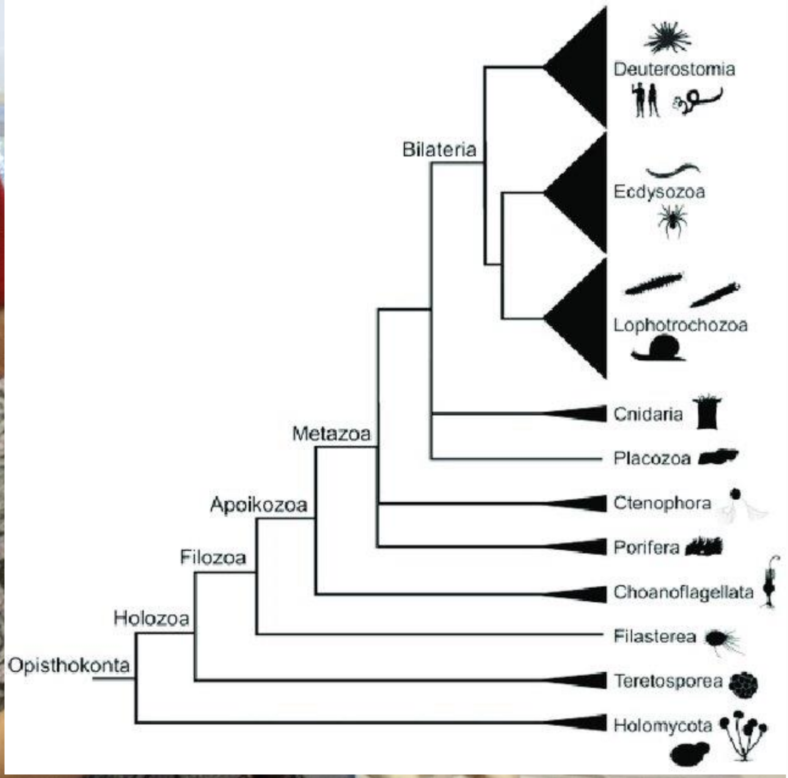
Hartman and Vision, 2008

COMPARATIVE GENOMICS

UB
Parc Científic
Barcelona

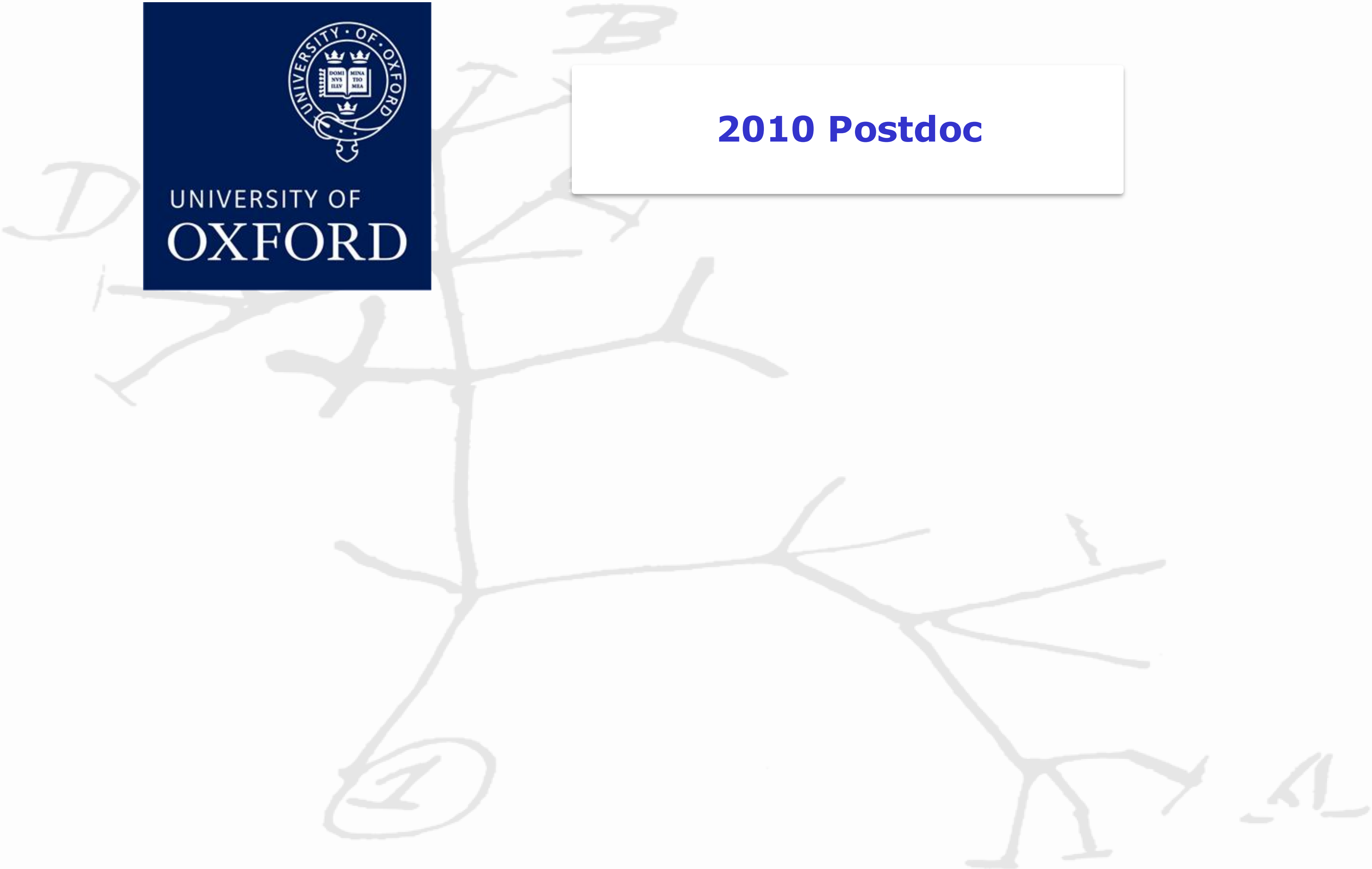






UNIVERSITY OF OXFORD

2010 Postdoc

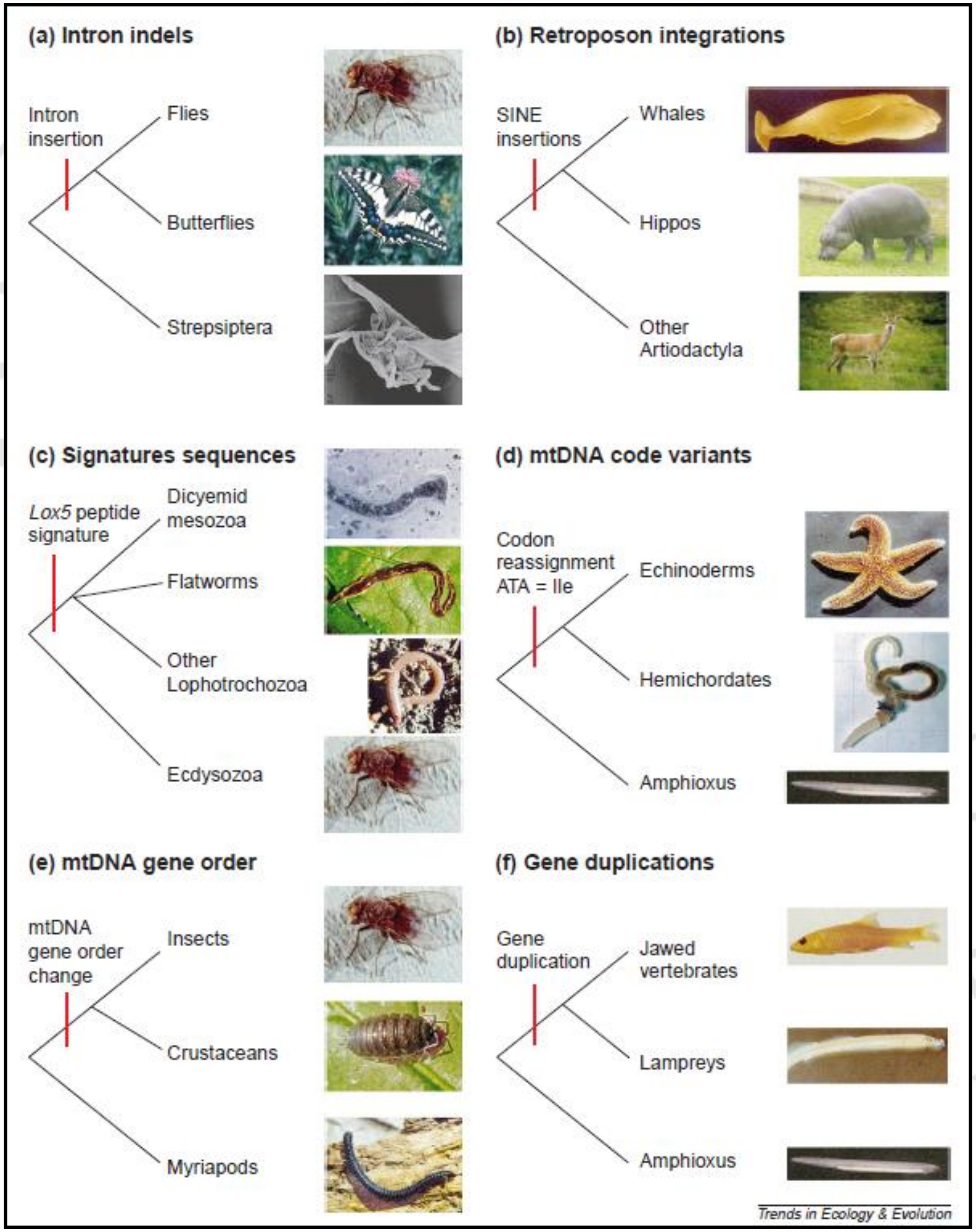


Genomic content as phylogenetic marker

Rare Genomic
Changes
(RGCs)

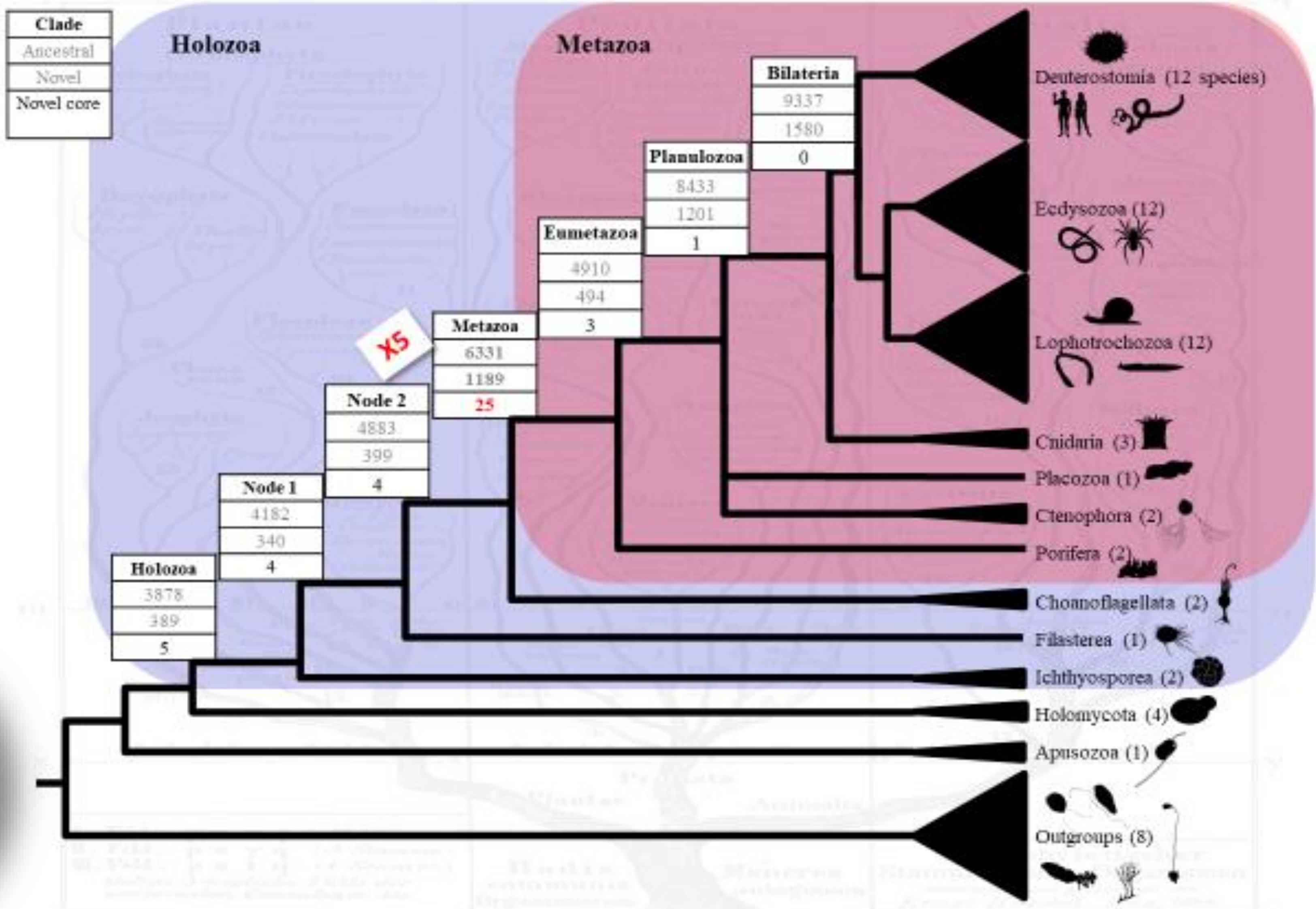
Antonis Rokas

Peter WH Holland



Rokas & Holland 2000
TREE

Gene novelty



Peter WH Holland

Paps & Holland 2018

UNIVERSITY OF OXFORD

2010 Postdoc

↓

University of Essex

2015 Lecturer

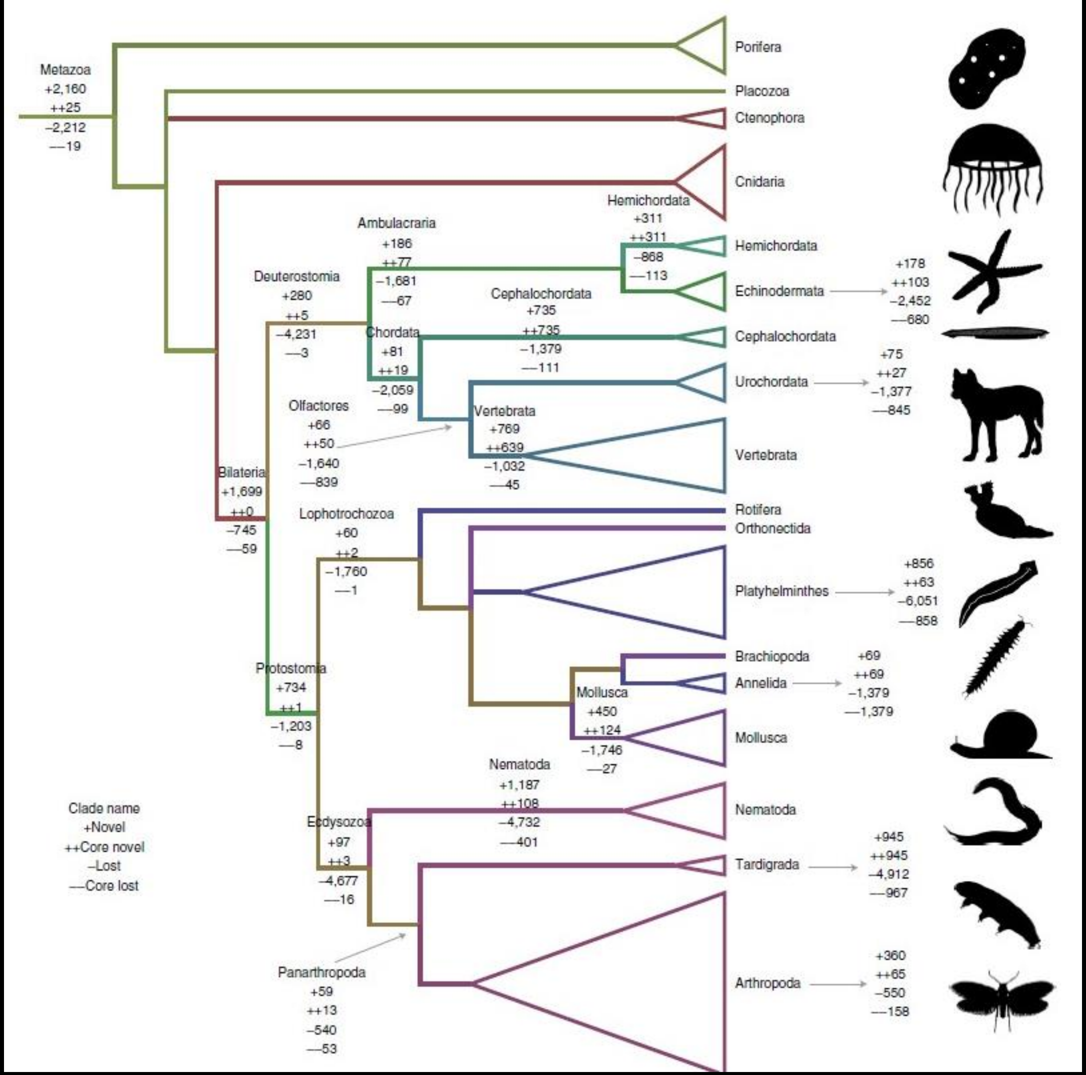
↓

University of BRISTOL

2019 Senior Lecturer

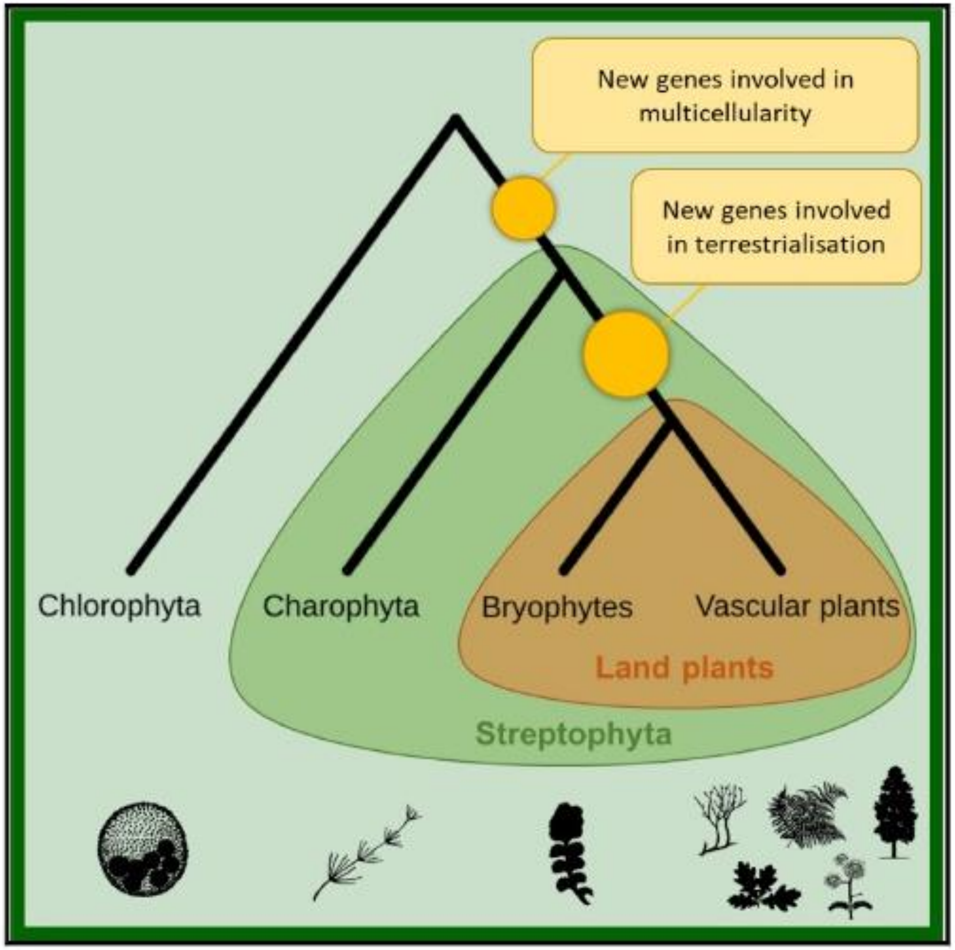
Widespread patterns of gene loss in the evolution of the animal kingdom

Cristina Guijarro-Clarke[1], Peter W. H. Holland[2] and Jordi Paps[1,2,3] ✉



The Origin of Land Plants Is Rooted in Two Bursts of Genomic Novelty

Graphical Abstract



Authors

Alexander M.C. Bowles,
Ulrike Bechtold, Jordi Paps

Correspondence

ubech@essex.ac.uk (U.B.),
jordi.paps@bristol.ac.uk (J.P.)

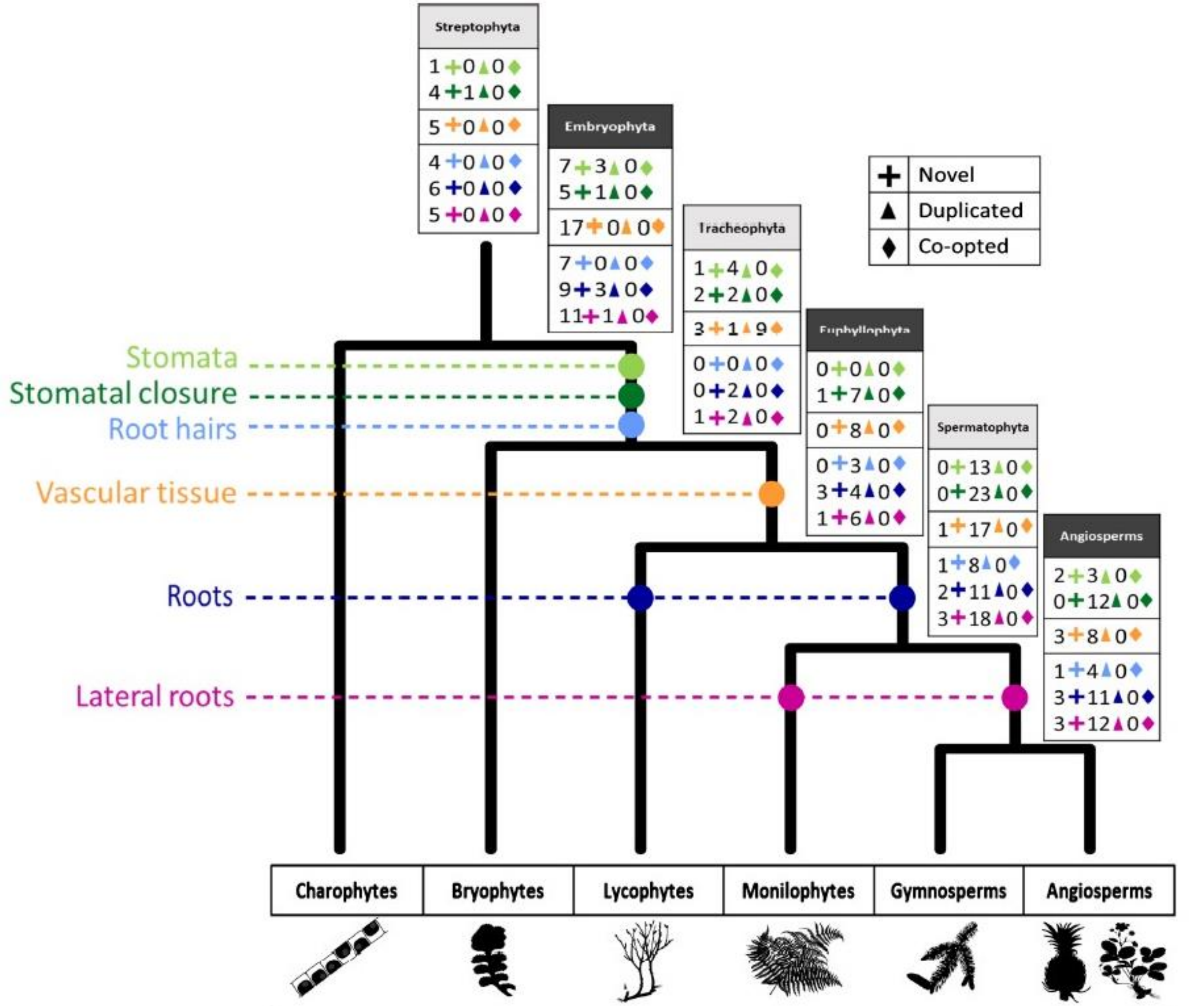
In Brief

Bowles et al. show that two consecutive bursts of genomic novelty predate the origin of land plants. Identifying these events provides insights into the evolution of flora that has defined modern ecosystems.

Water-related innovations in land plants evolved by different patterns of gene cooption and novelty

Alexander M. C. Bowles[1,2], **Jordi Paps**[1,3] and **Ulrike Bechtold**[1,4]

[1]School of Life Sciences, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK; [2]School of Geographical Sciences, University of Bristol, University Road, Bristol, BS8 1RL, UK; [3]School of Biological Sciences, University of Bristol, 24 Tyndall Avenue, Bristol, BS8 1TQ, UK; [4]Present address: Department of Biosciences, Durham University, South Road, Durham, DH1 3LE, UK



Marta Álvarez-Presas

USING GENOME-LEVEL PROCESSES TO SOLVE TRICKY NODES IN PHYLOGENETICS

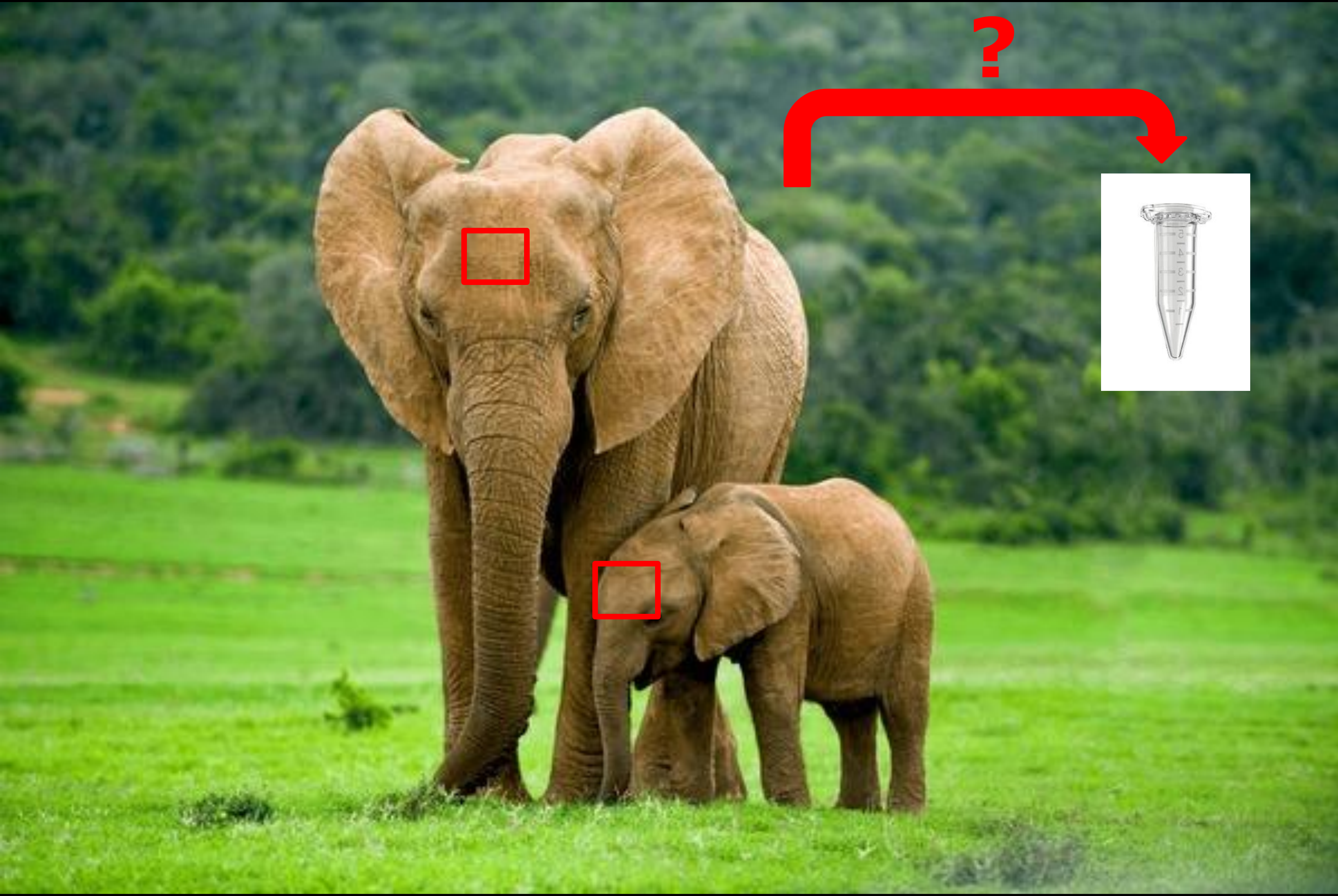
Iñaki Ruiz-Trillo's lab
(Barcelona, Spain)

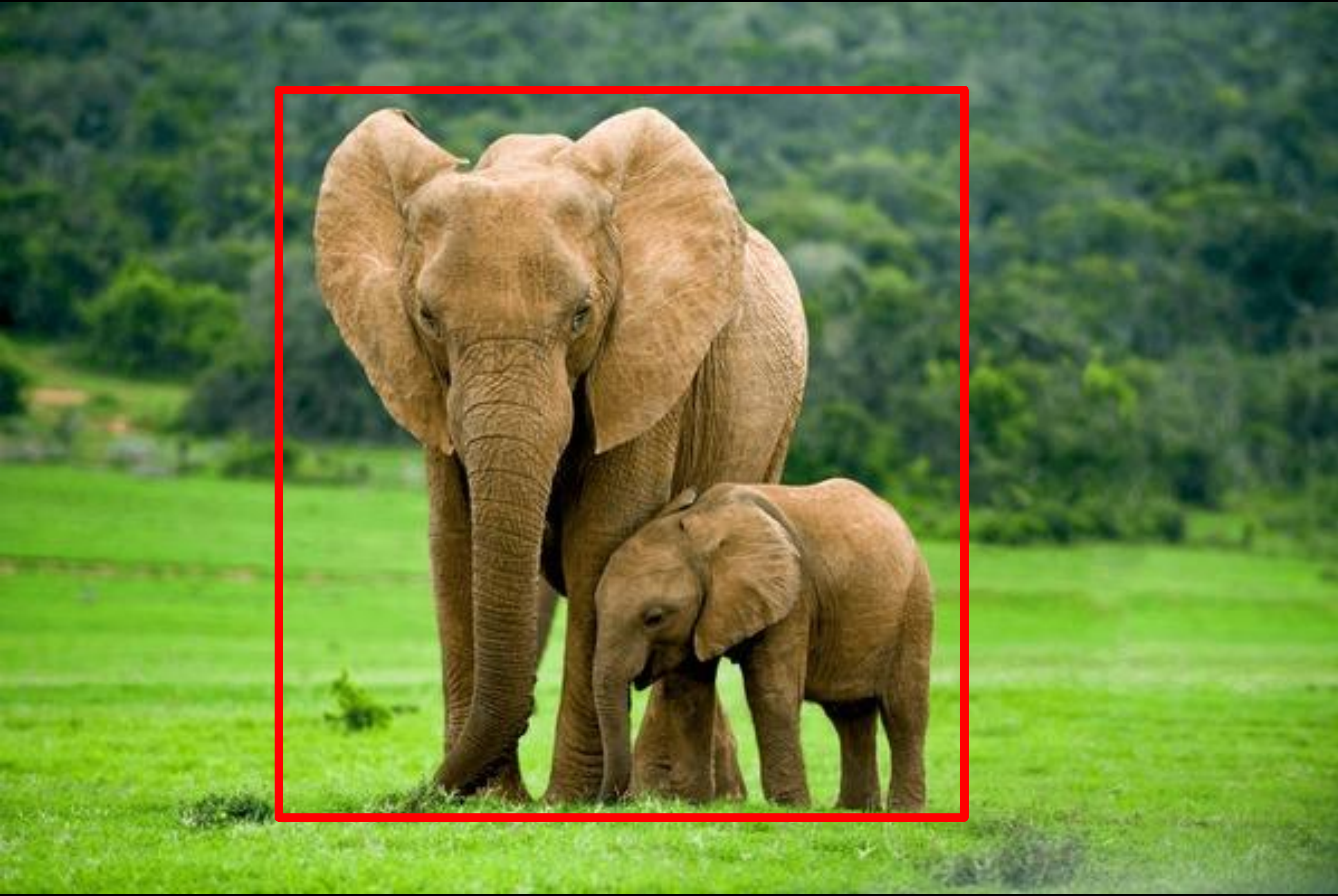
University of BRISTOL


THERE IS NO
PLANET B

Two major issues

- **Orthology issues derived from “phylotranscriptomics”**
    - ESTs/RNAseq only represent a fraction of the genes in a genome, limited to a tissue and/or developmental stage.





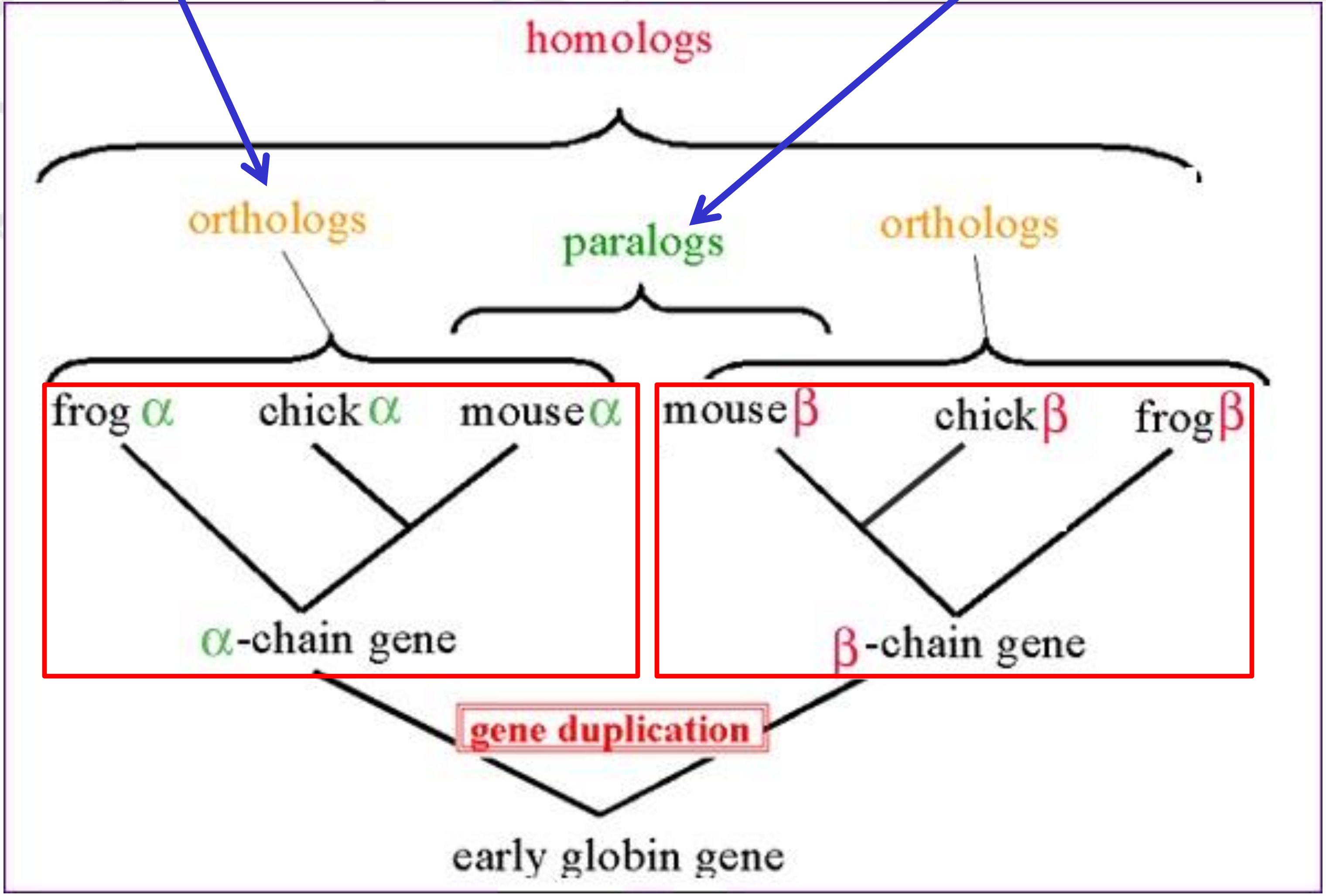
Two major issues

- **Orthology issues derived from “phylotranscriptomics”**
  - ESTs/RNAseq only represent a fraction of the genes in a genome, limited to a tissue and/or developmental stage.
  - Genes are often incomplete and fragmented
  - More prone to contamination than genomes

Homology and genes

Orthologous genes:
a product of speciation

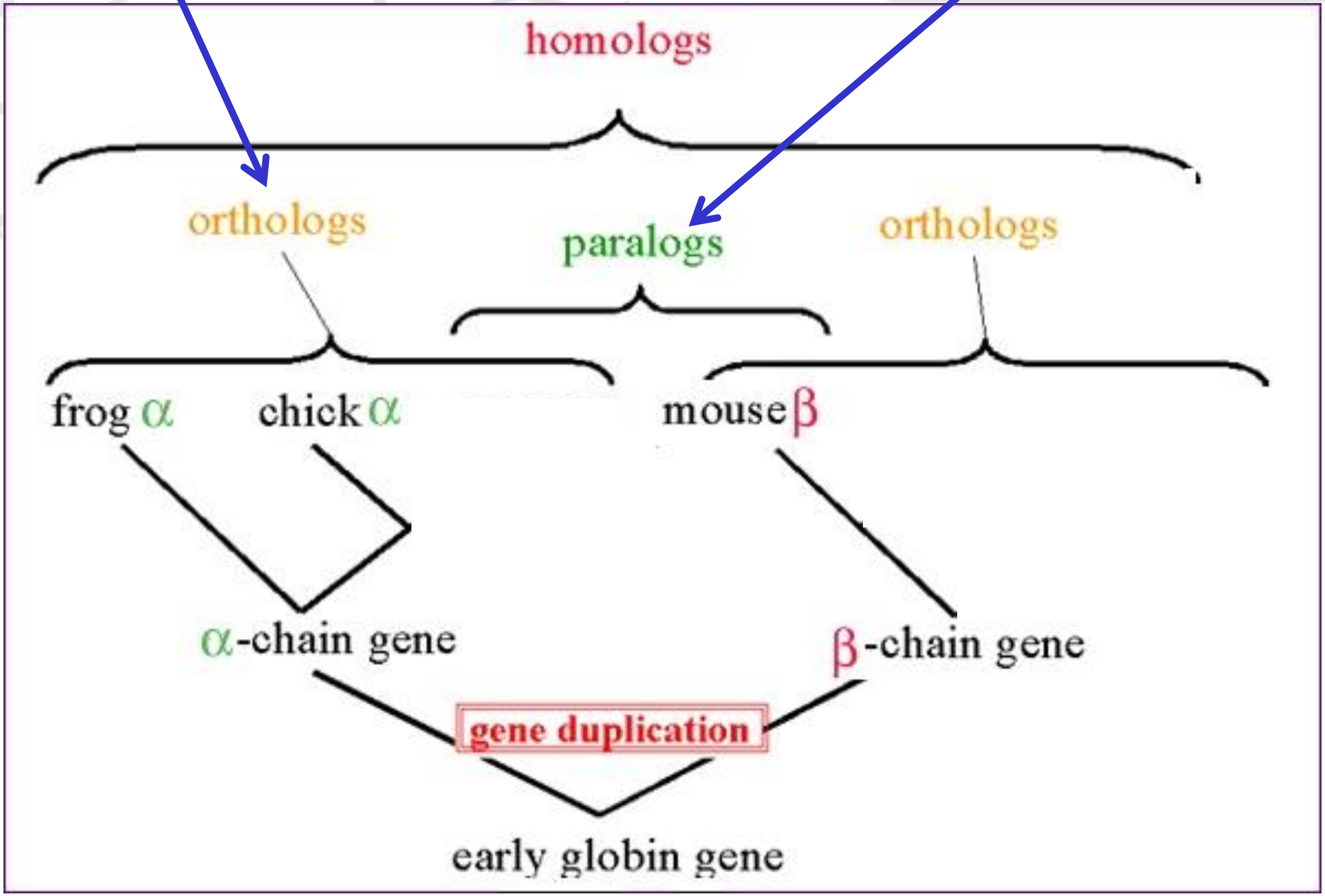
Paralogous genes:
duplication within a genome



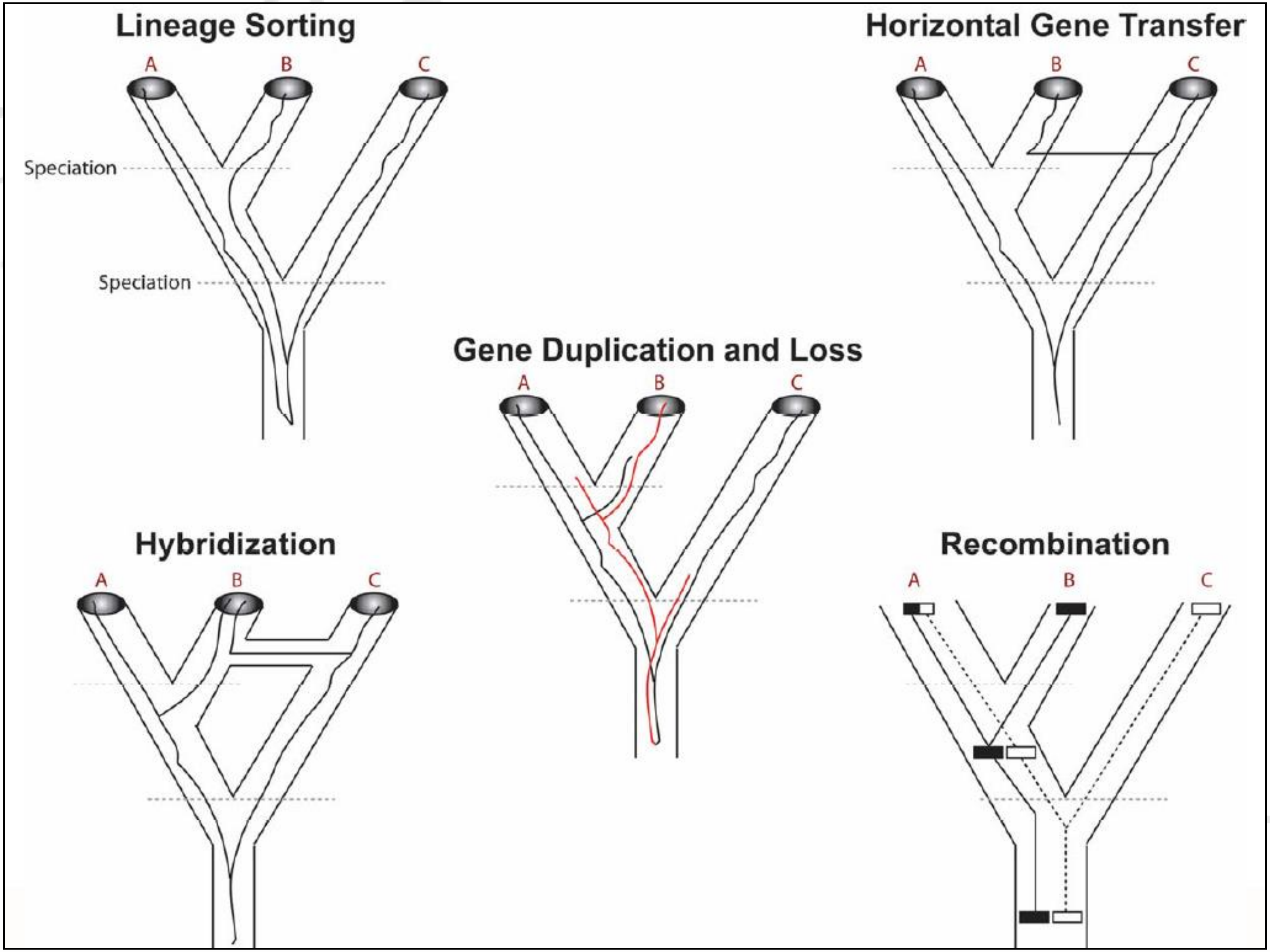
Hidden paralogy

Orthologous genes:
a product of speciation

Paralogous genes:
duplication within a genome



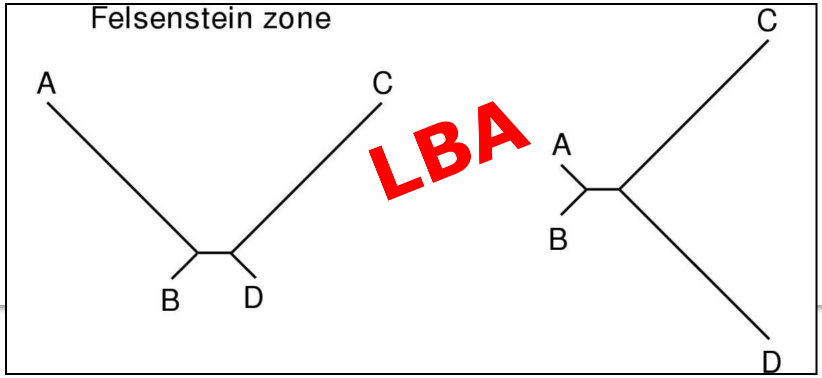
Orthology issues



Two major issues

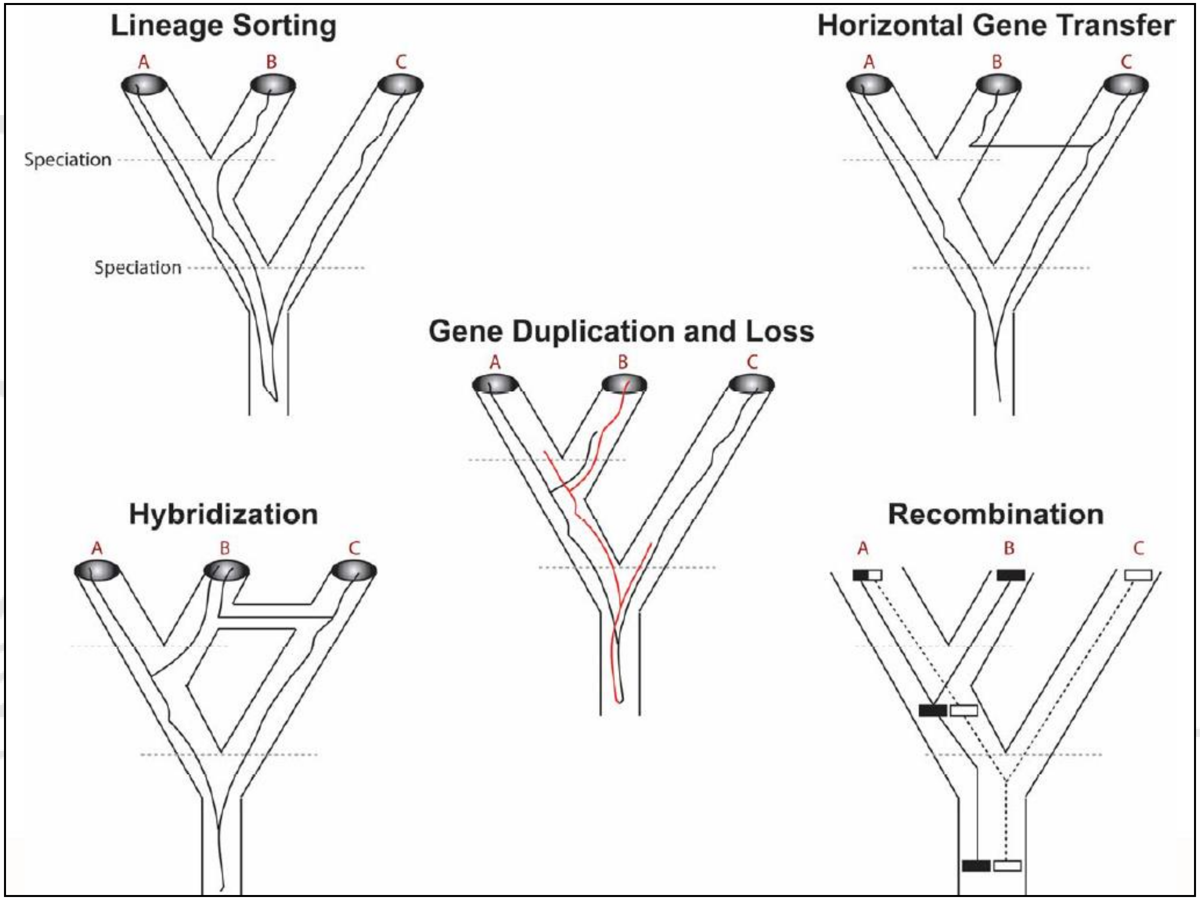
- **Orthology issues derived from "phylotranscriptomics"**
  - ESTs/RNAseq only represent a fraction of the genes in a genome, limited to a tissue and/or developmental stage.
  - Genes are often incomplete and fragmented
  - More prone to contamination than genomes

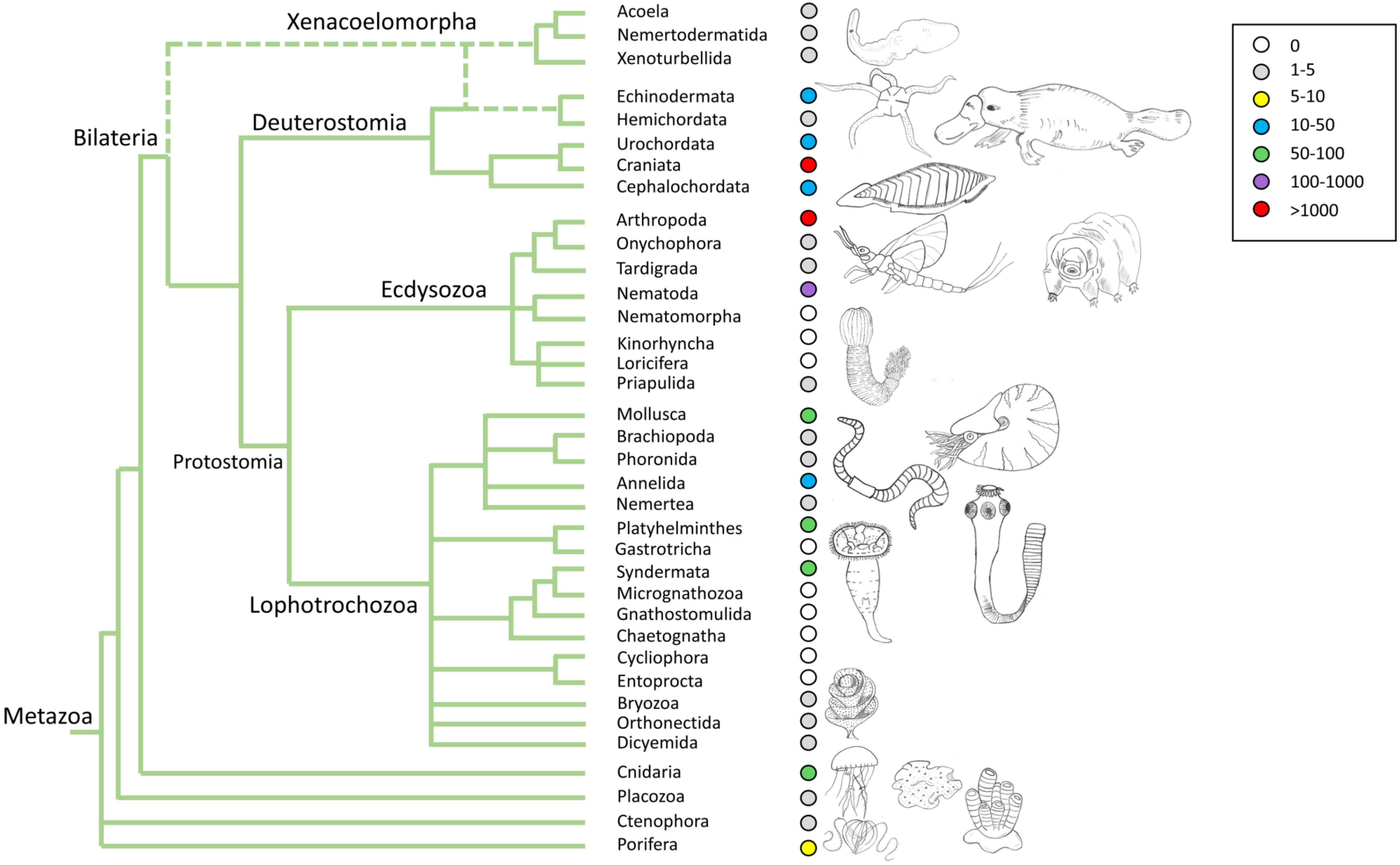
- **Systematic errors (e.g., LBA) cause by evolutionary model misspecification applied to an alignment**



The approaches

- Orthology issues -> use only complete genomes
- Systematic errors -> embrace genome-level processes!





Xenacoelomorpha

Bilateria

Deuterostomia

Ecdysozoa

Protostomia

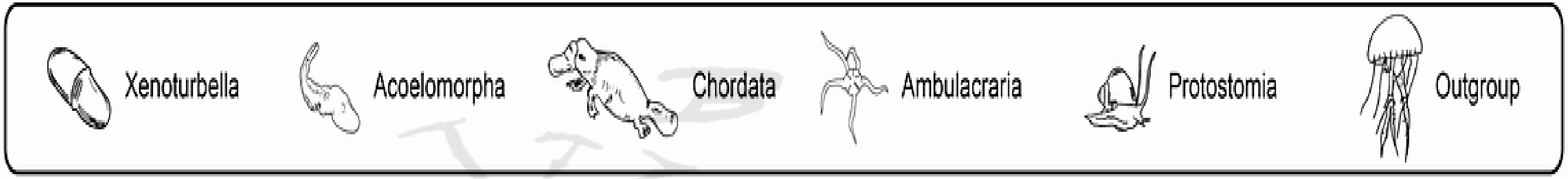
Lophotrochozoa

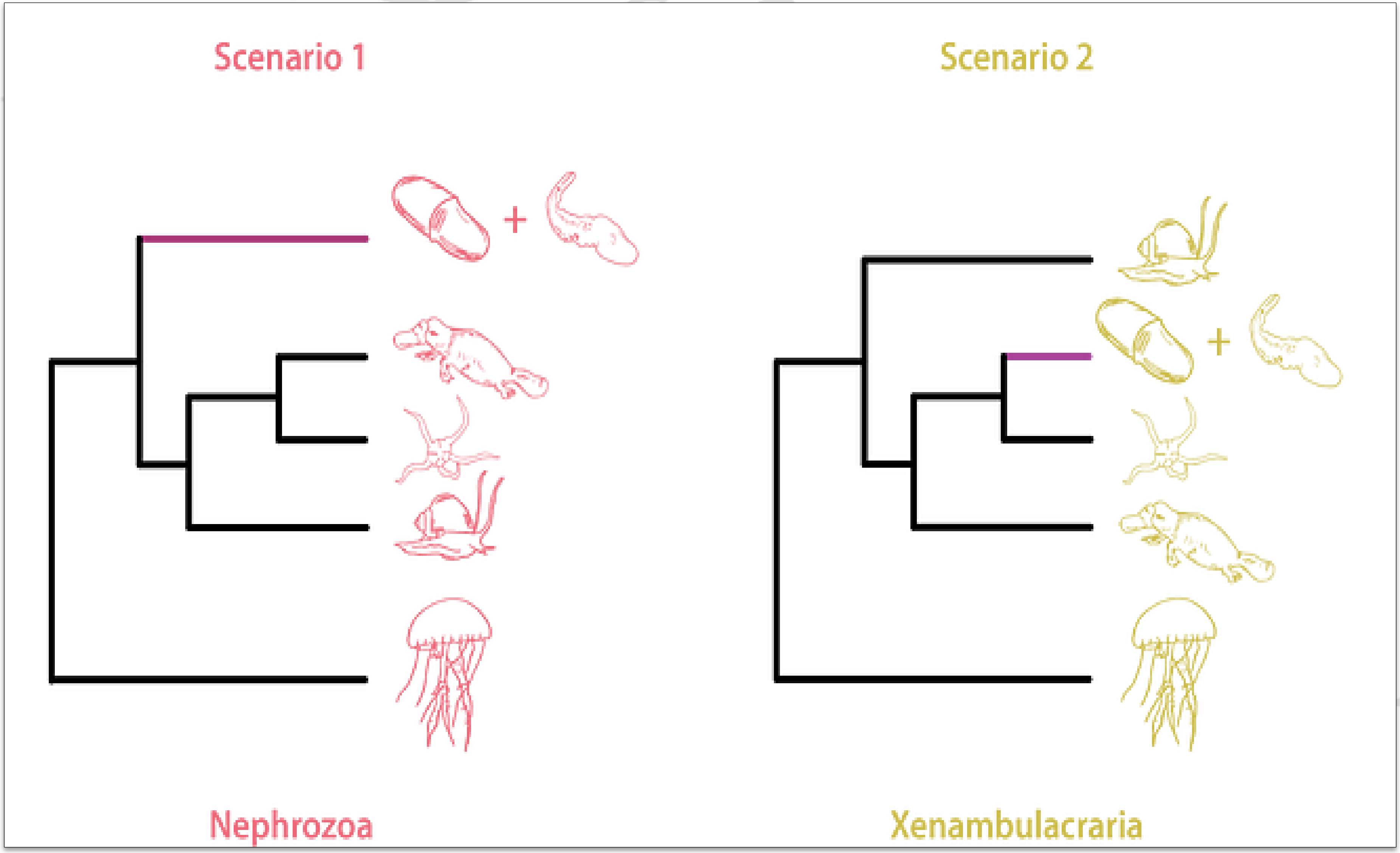
Metazoa

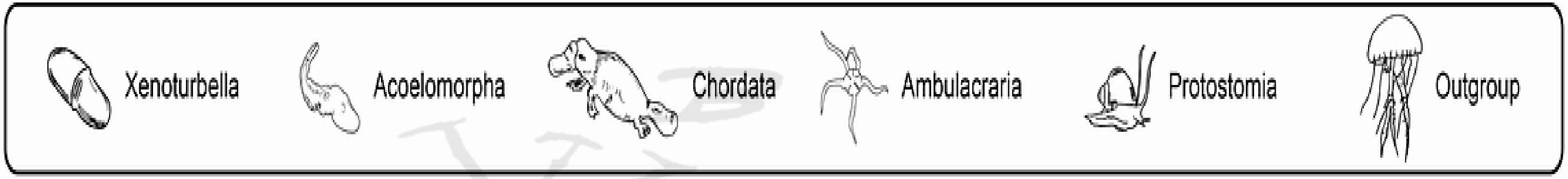
Acoela	◍
Nemertodermatida	◍
Xenoturbellida	◍
Echinodermata	🔵
Hemichordata	◍
Urochordata	🔵
Craniata	🔴
Cephalochordata	🔵
Arthropoda	🔴
Onychophora	◍
Tardigrada	◍
Nematoda	🟣
Nematomorpha	○
Kinorhyncha	○
Loricifera	○
Priapulida	◍
Mollusca	🟢
Brachiopoda	◍
Phoronida	◍
Annelida	🔵
Nemertea	◍
Platyhelminthes	🟢
Gastrotricha	○
Syndermata	🟢
Micrognathozoa	○
Gnathostomulida	○
Chaetognatha	○
Cycliophora	○
Entoprocta	○
Bryozoa	◍
Orthonectida	◍
Dicyemida	◍
Cnidaria	🟢
Placozoa	◍
Ctenophora	◍
Porifera	🟡

○	0
◍	1-5
🟡	5-10
🔵	10-50
🟢	50-100
🟣	100-1000
🔴	>1000

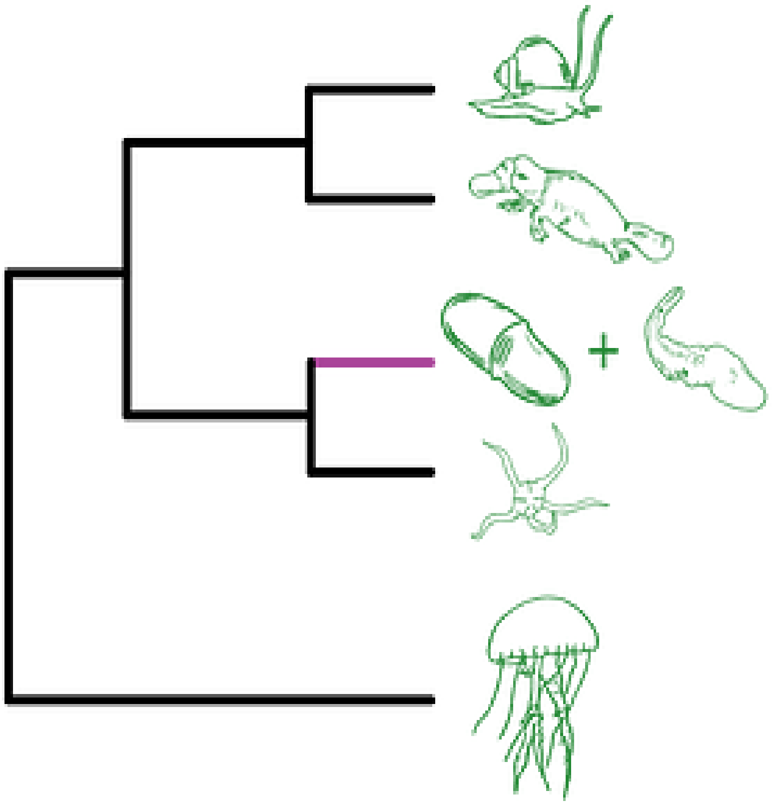
Paps et al 2023


Xenoturbella
Acoelomorpha
Chordata
Ambulacraria
Protostomia
Outgroup


Scenario 1
+
Nephrozoa
Scenario 2
+
Xenambulacraria

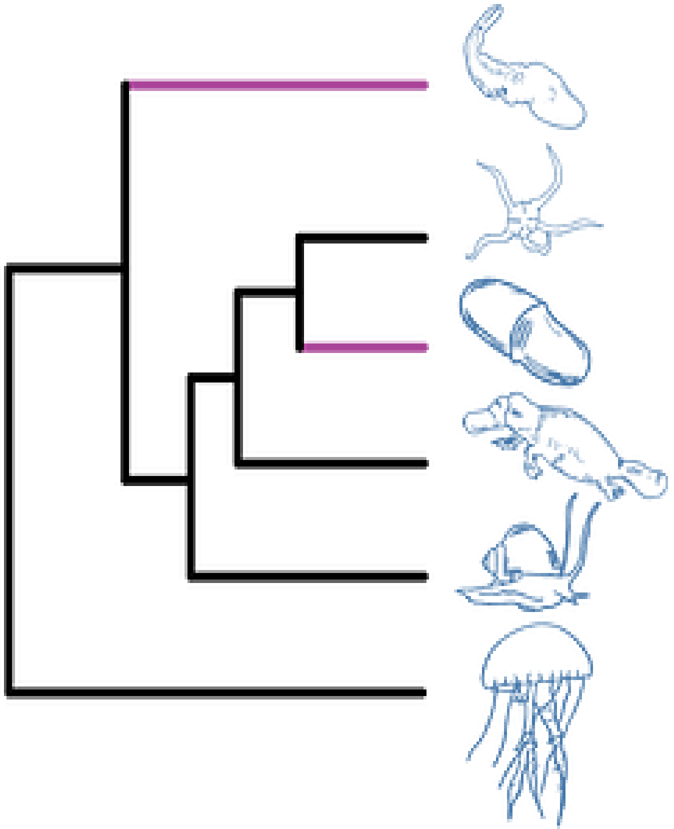


Scenario 3



Xenambulacraria & paraphyletic deuterostomes

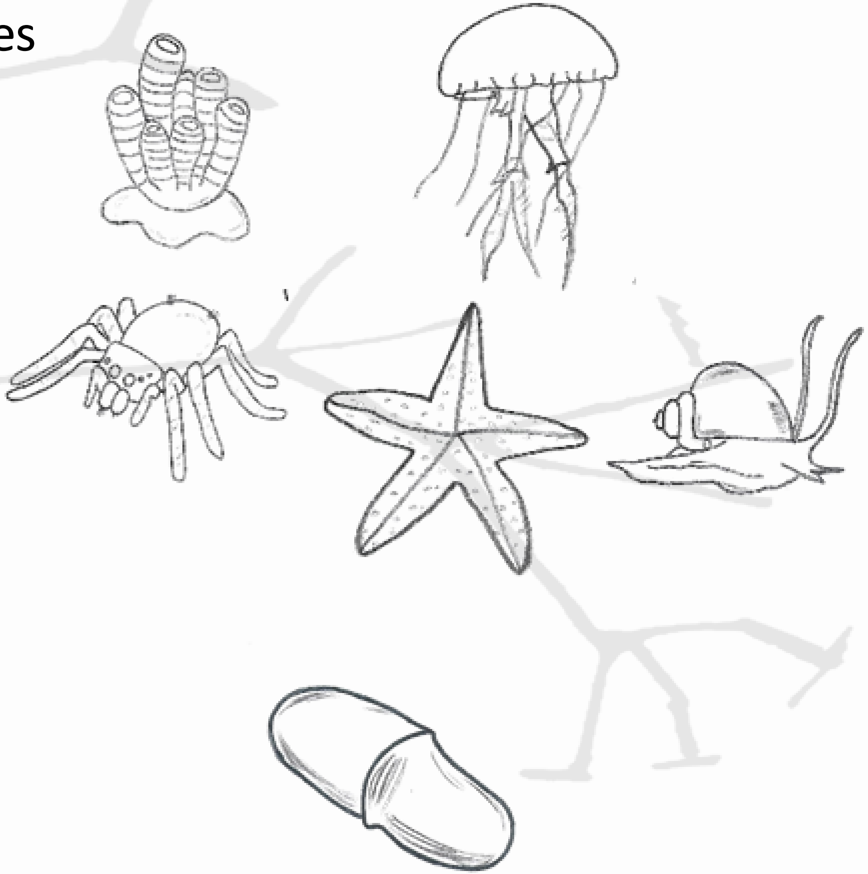
Scenario 4



Xenoturbella sister to Ambulacraria

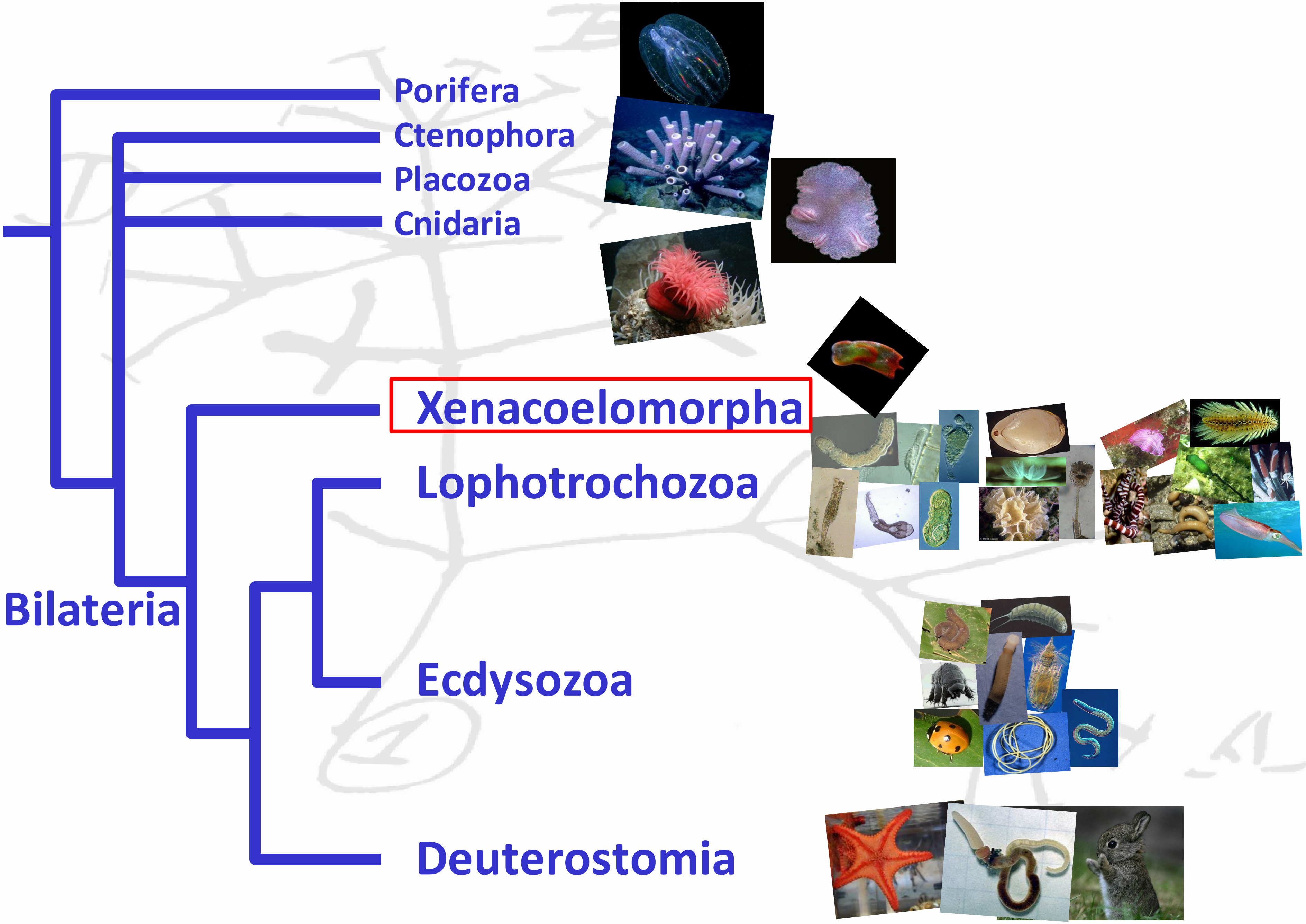
Sampling

- **100+ genomes**, no transcriptomes
- **91 BUSCO filtered -> OrthoFinder**
- **Different outgroups** (Holozoa, Metazoa, Cnidaria Bilateria)



- 16 non-animal holozoan species
- 24 non-Bilateria
  - 6 Porifera
  - 2 Ctenophora
  - 2 Placozoa
  - **14 Cnidaria**
- 12 Lophotrochozoa
- 13 Ecdysozoa
- **7 Ambulacraria**
- 13 Chordata
- **7 Xenacoelomorpha**
  - 1 Xenoturbellida
  - 4 Acoela
  - 2 Nemertodermatida

Spoilers!!!



Porifera

Ctenophora

Placozoa

Cnidaria

Xenacoelomorpha

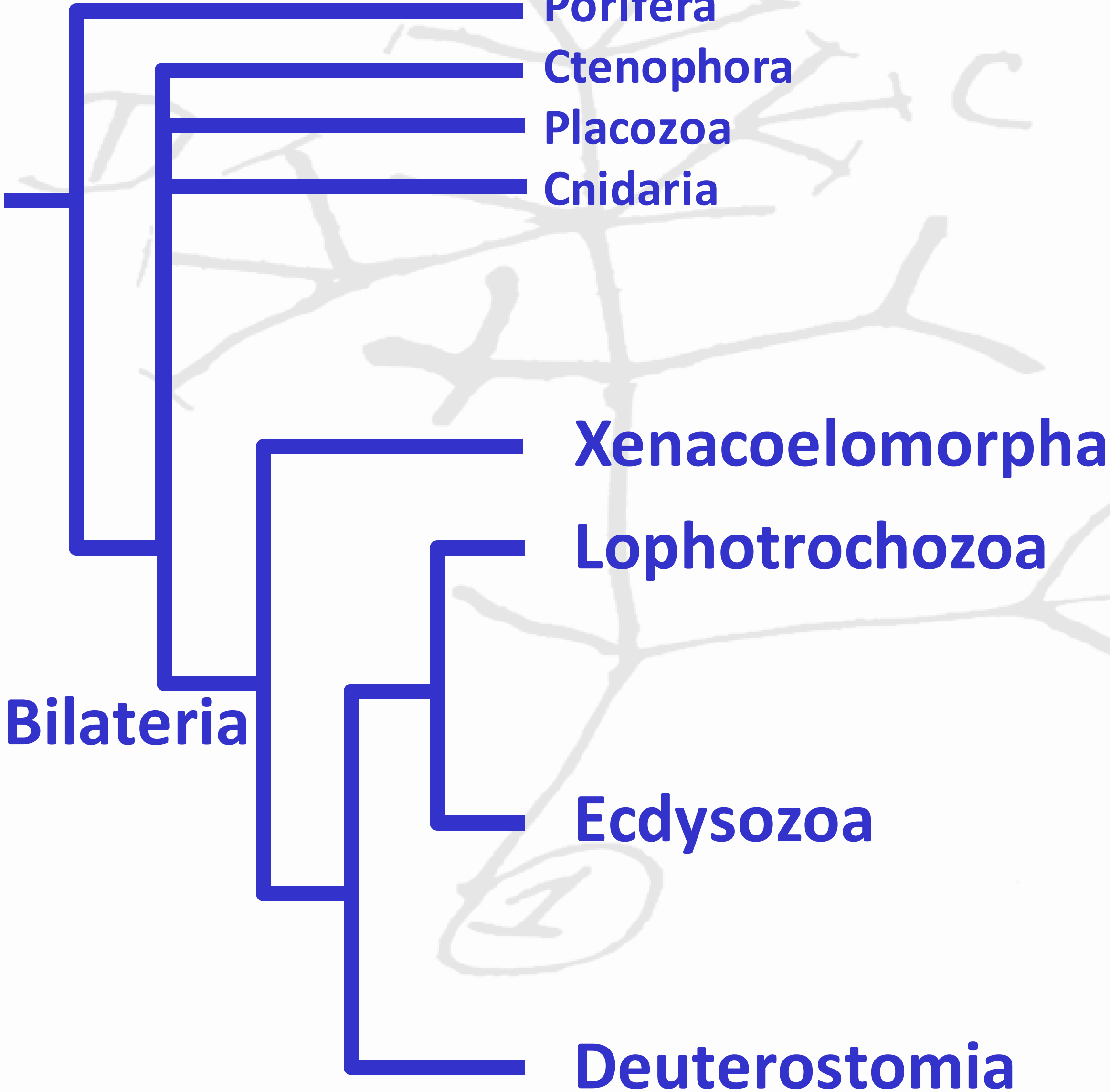
Lophotrochozoa

Ecdysozoa

Deuterostomia

Bilateria

Canary in the coal mine



Porifera

Ctenophora

Placozoa

Cnidaria

Xenacoelomorpha

Lophotrochozoa

Ecdysozoa

Deuterostomia

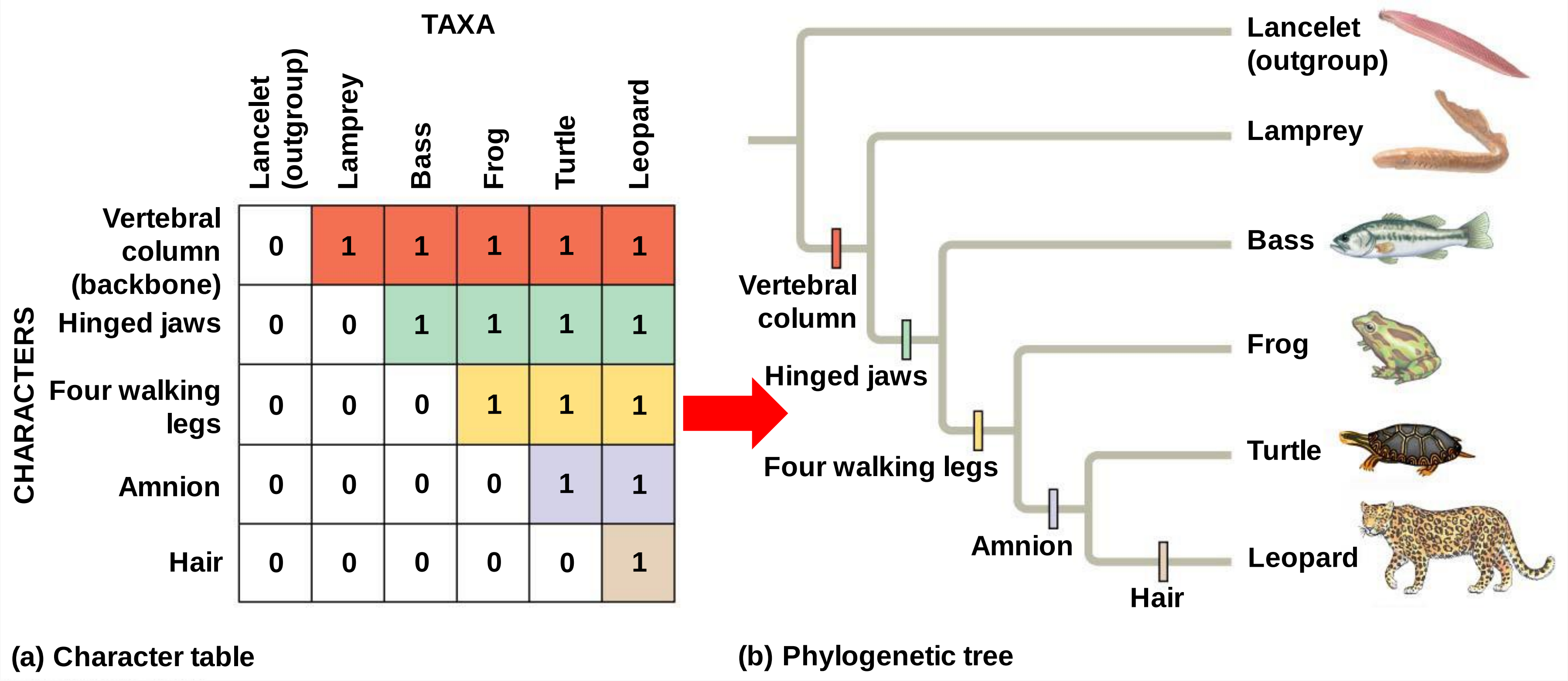
Bilateria

Genome-level analyses

1. **Gene content** for 549,544 gene families
2. **Modes of gene family evolution** (duplications, transfers, and losses; DTL)
3. **Multispecies coalescence model** (MSC)

Gene content

- **Character matrix** records presence/absence of traits
- Then this data is used to infer the tree



TAXA

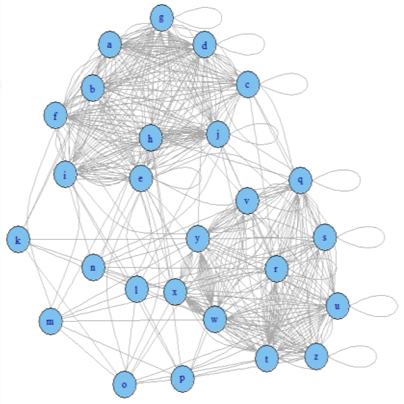
CHARACTERS	Lancelet (outgroup)	Lamprey	Bass	Frog	Turtle	Leopard
Vertebral column (backbone)	0	1	1	1	1	1
Hinged jaws	0	0	1	1	1	1
Four walking legs	0	0	0	1	1	1
Amnion	0	0	0	0	1	1
Hair	0	0	0	0	0	1

(a) Character table

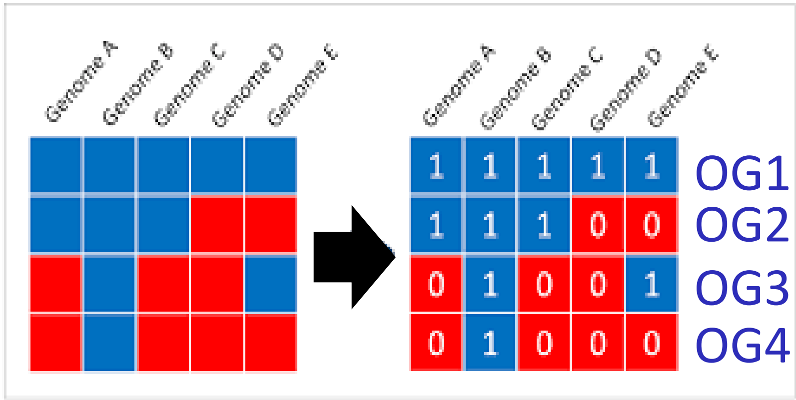
(b) Phylogenetic tree

Gene content: 549,544 gene families

Orthogroups



Presence / Absence Matrix

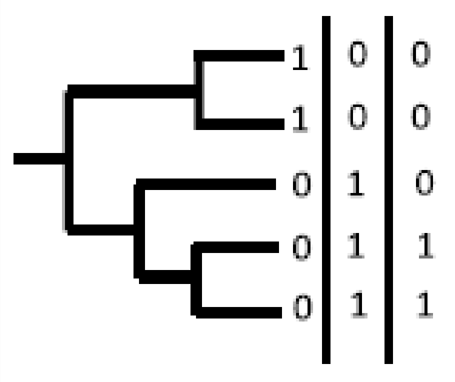


	Genome A	Genome B	Genome C	Genome D	Genome E
OG1	1	1	1	1	1
OG2	1	1	1	0	0
OG3	0	1	0	0	1
OG4	0	1	0	0	0

RevBayes Phylogeny

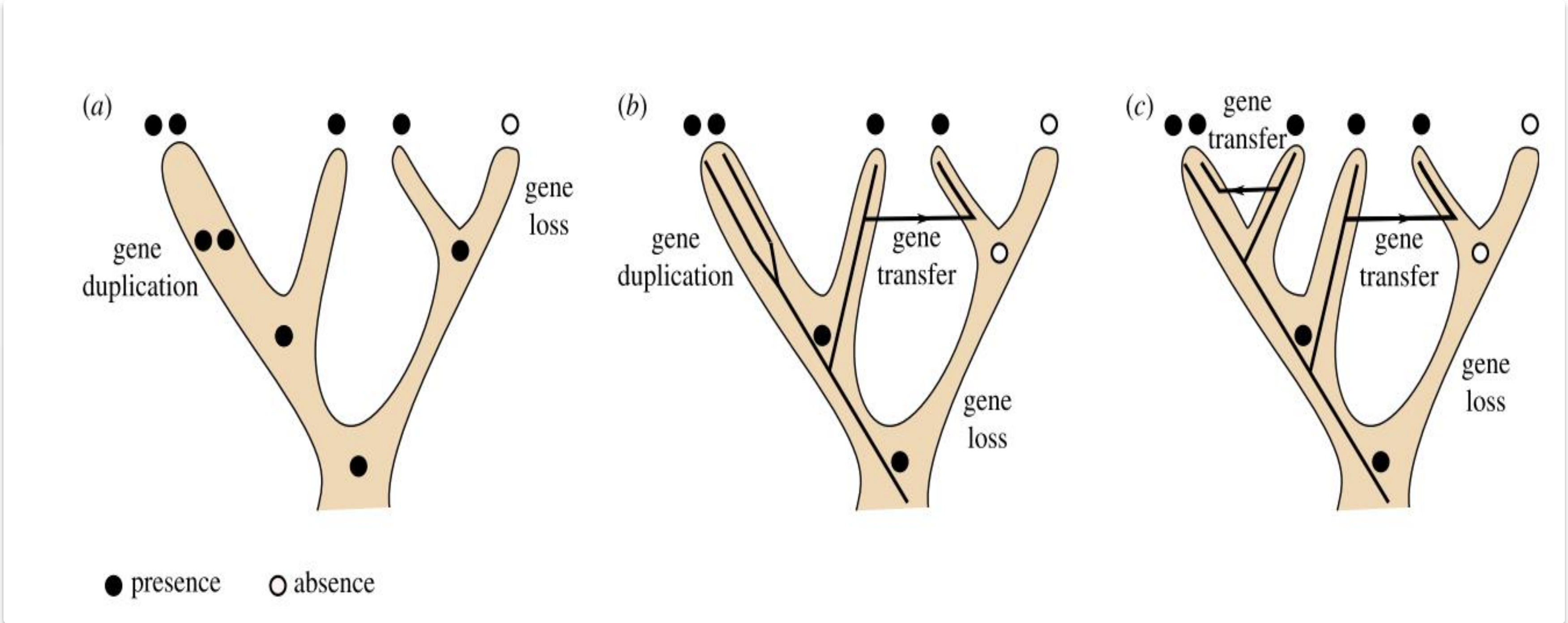
Markov Chain (CTCM) & Dollo models

Correction for unobserved gene losses



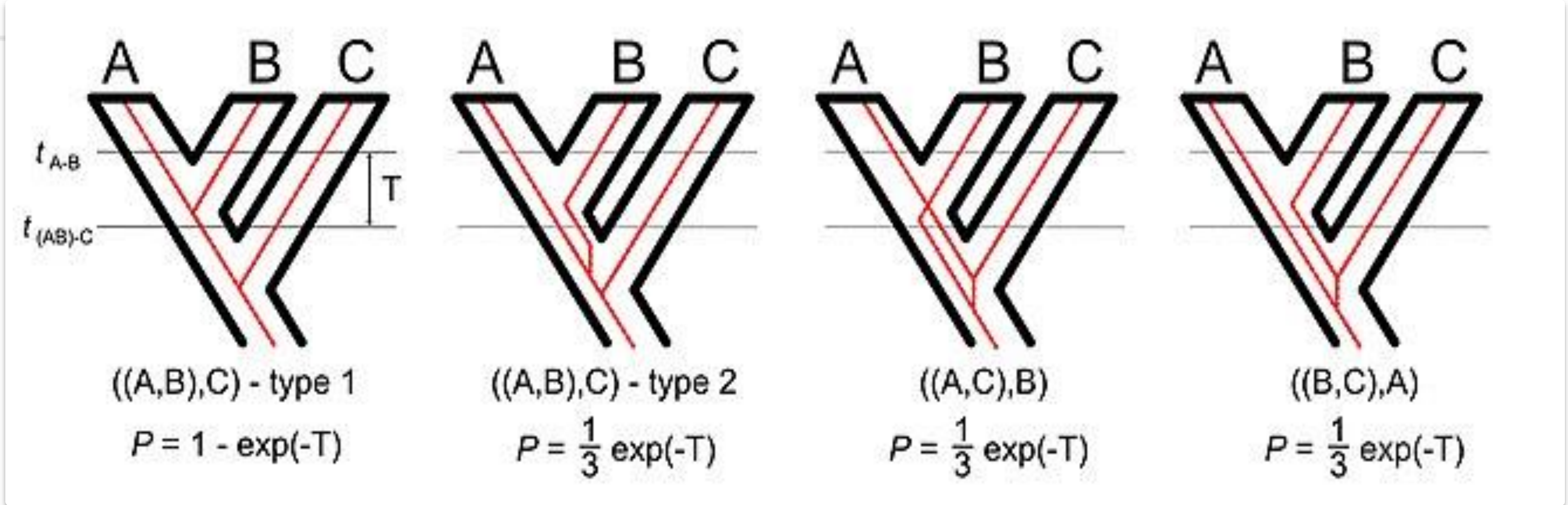
1	0	0
1	0	0
0	1	0
0	1	1
0	1	1

Duplications, Transfers, Losses (DTLs)



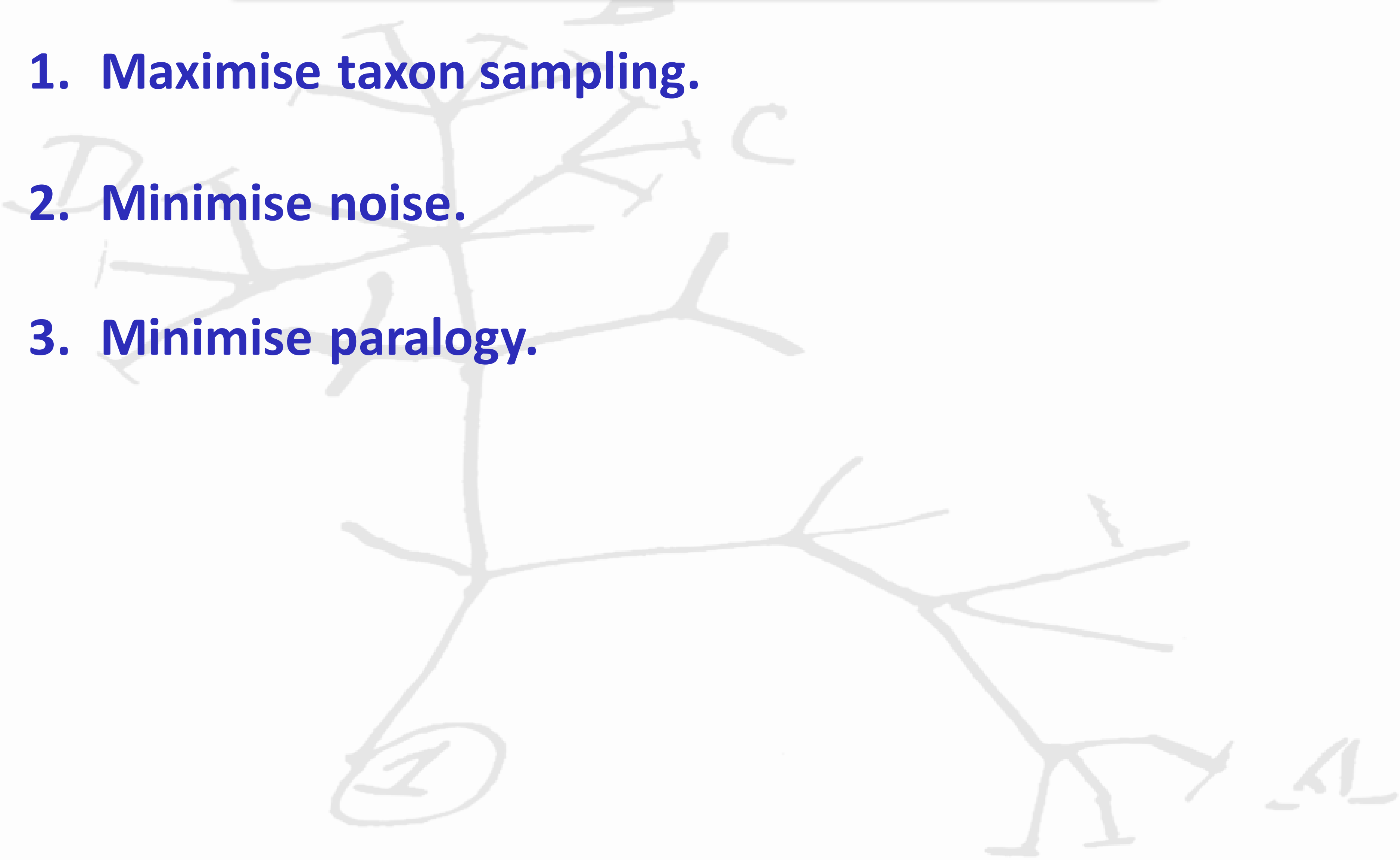
- **ALE**: cannot infer tree, but can compare topologies statically
- **SpeciesRax:** infers tree from DTL in all the gene trees of all the genes in all genomes

Multispecies Coalescent (MSC)



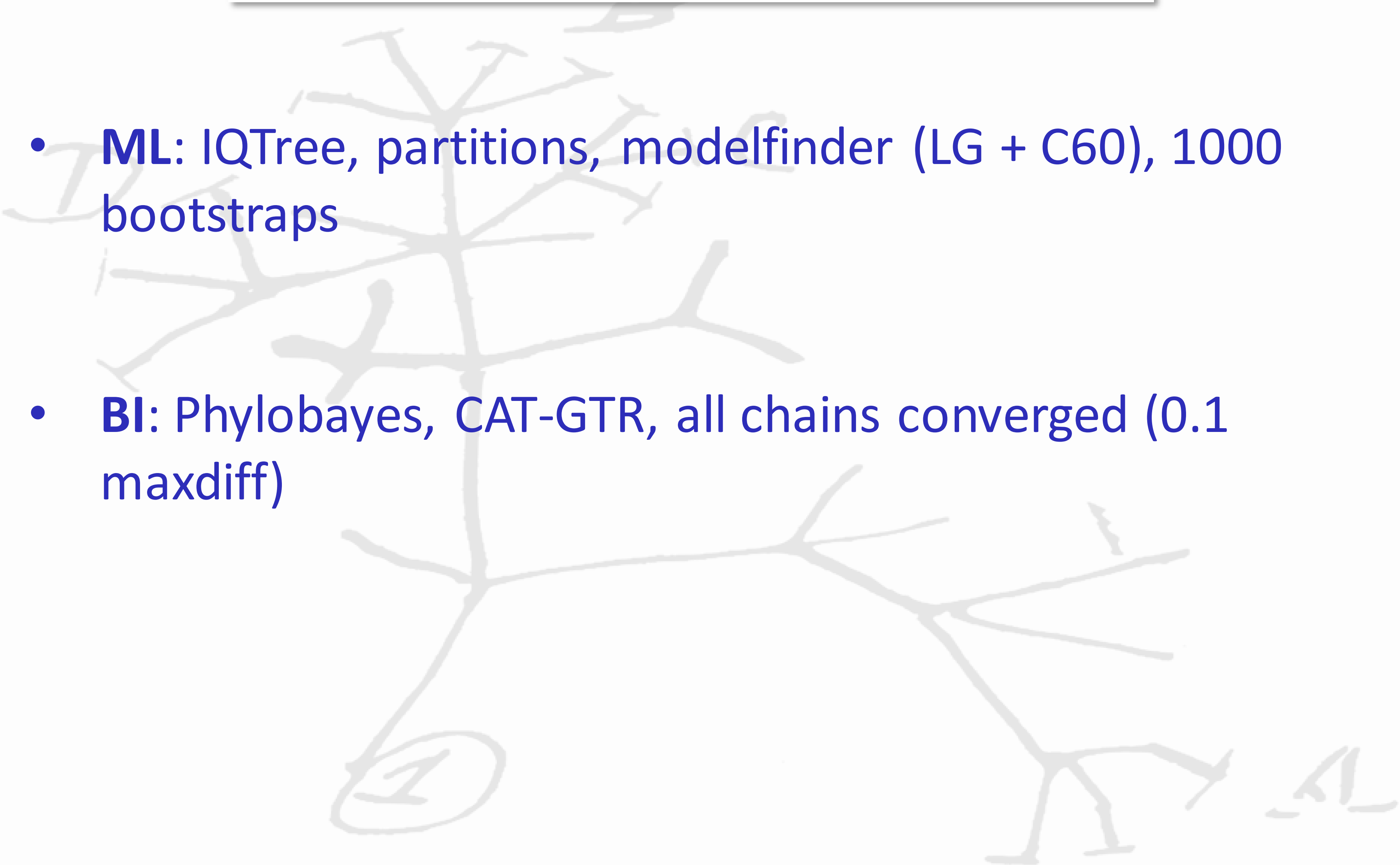
- **ASTRAL:** infers all gene trees of all the genes in all genomes using the MSC
- Then it summarises all the gene trees into one.

Phylogenomics: three approaches

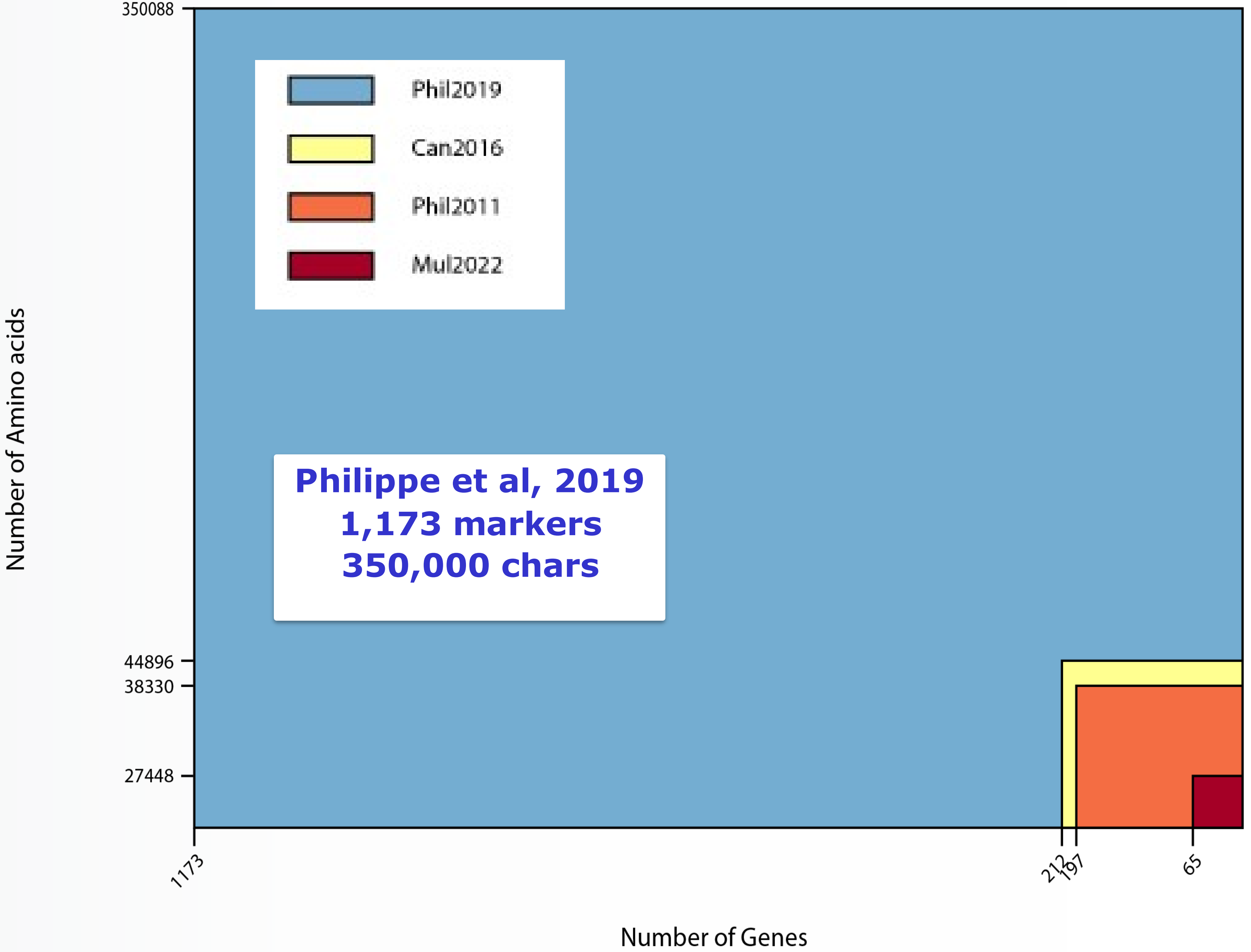


1. **Maximise taxon sampling.**
2. **Minimise noise.**
3. **Minimise paralogy.**

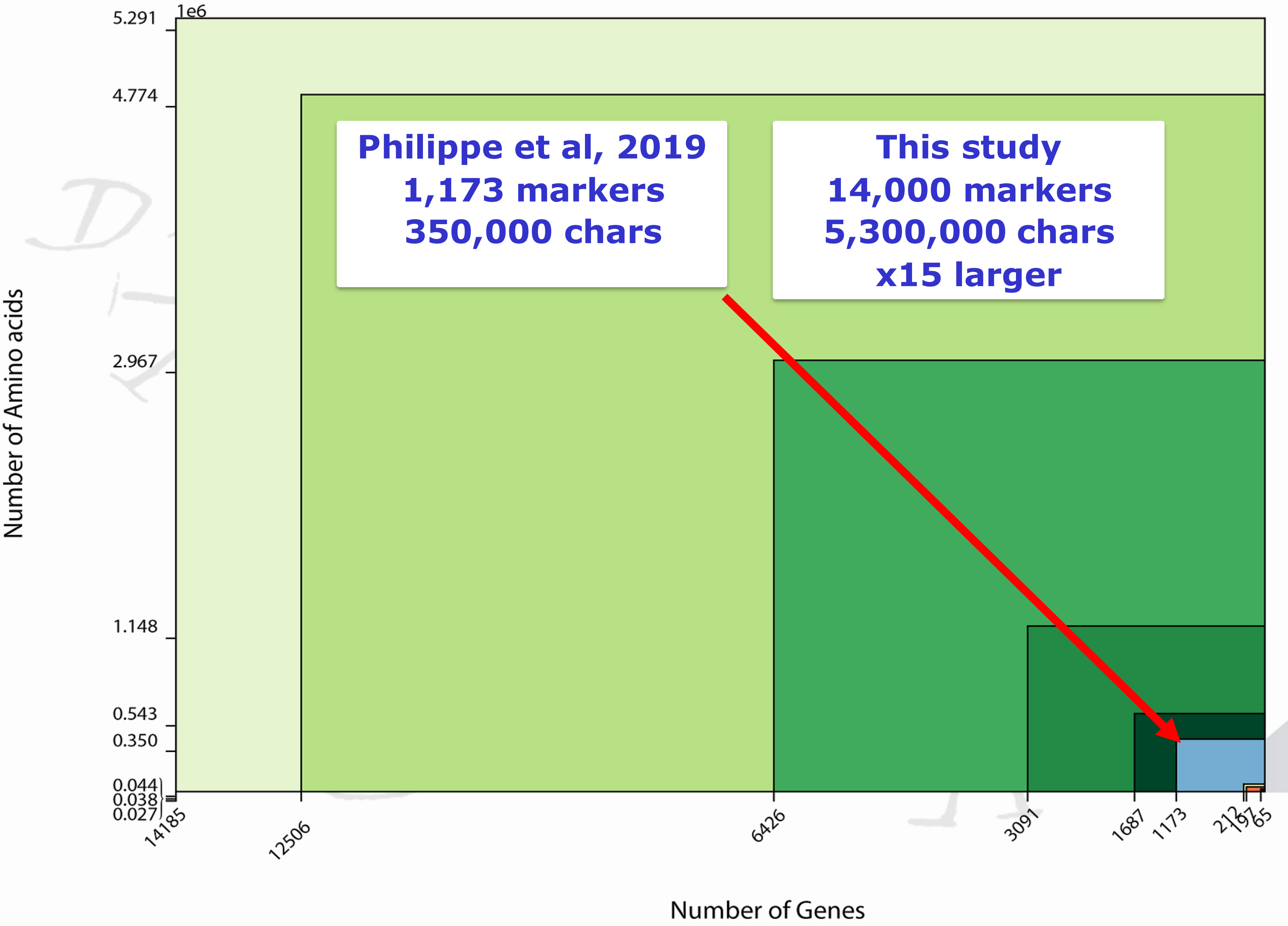
Phylogenomics: three approaches



- **ML**: IQTree, partitions, modelfinder (LG + C60), 1000 bootstraps
- **BI**: Phylobayes, CAT-GTR, all chains converged (0.1 maxdiff)



Philippe et al, 2019
1,173 markers
350,000 chars



Philippe et al, 2019
1,173 markers
350,000 chars

This study
14,000 markers
5,300,000 chars
x15 larger

Phylogenomics: three approaches

1. **Maximise taxon sampling.**
    - 14k genes and 5,3 M positions
2. **Minimise noise:** ALE25T
    - 6,4k genes and 3M positions
3. **Minimise paralogy**
    - Single-copy filtered (3k genes and 1M, and smaller)