

Predicting ancestral syntenies

Eric Tannier, INRIA, University of Lyon

joint work with

Cedric Chauve, Simon Fraser University, Vancouver

Frédéric Boyer, CEA Grenoble, France

Marie-France Sagot, INRIA, University of Lyon



Paleogenomics

Paleogenomics in vertebrates, or the recovery of lost genomes from the mist of times.

M. Muffato and H. R. Crollius.

BioEssays, 30:122–134, 2008.



Ancestral Animal Genome Reconstructions

V. L. Rascol, P. Pontarotti, and A. Levasseur.

Current Opinions in Immunology, 19:542–546, 2008.



Addressing chromosome evolution in the whole-genome sequence era.

T. Faraut.

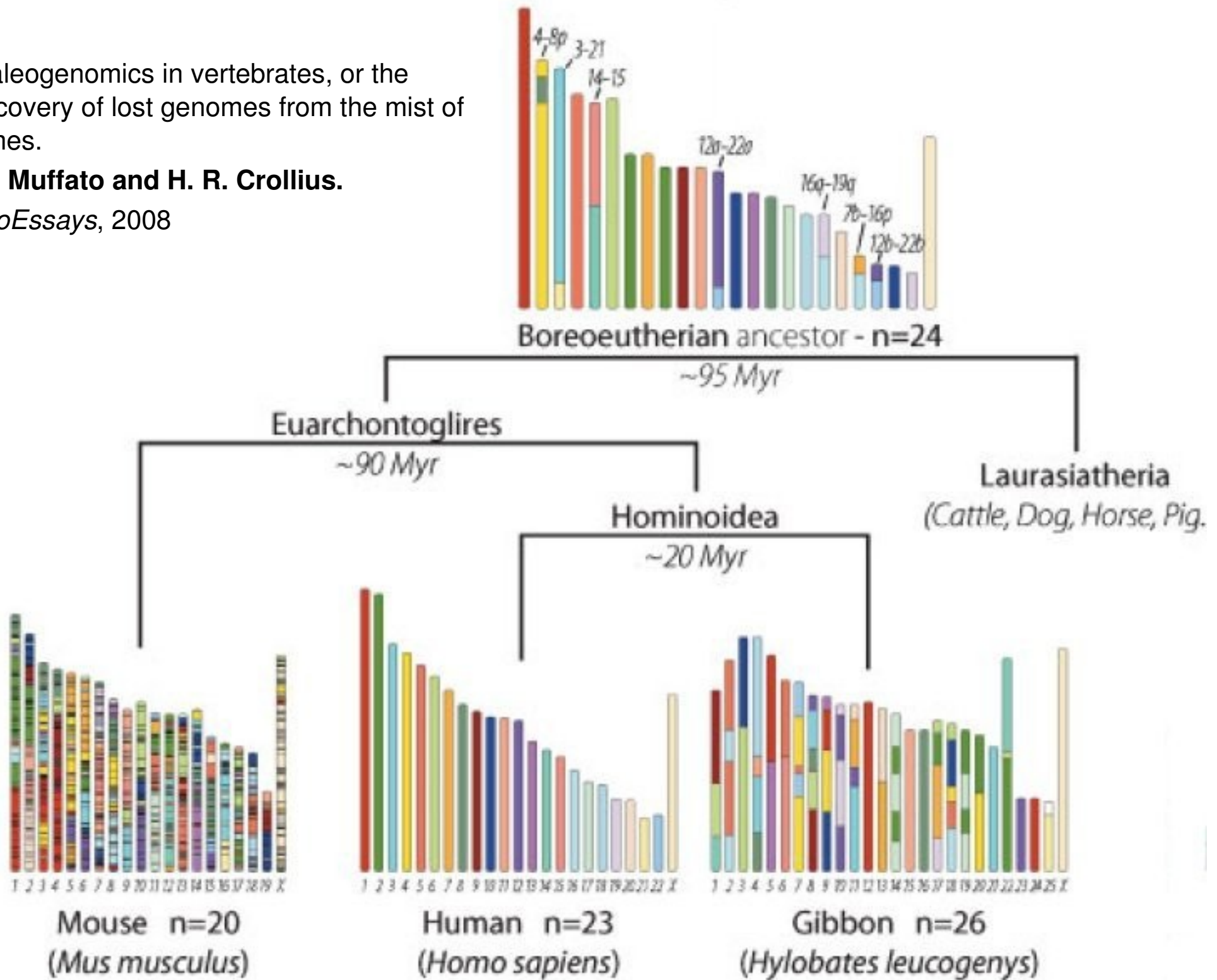
Chromosome Research, 16:5–16, 2008.



Paleogenomics in vertebrates, or the recovery of lost genomes from the mist of times.

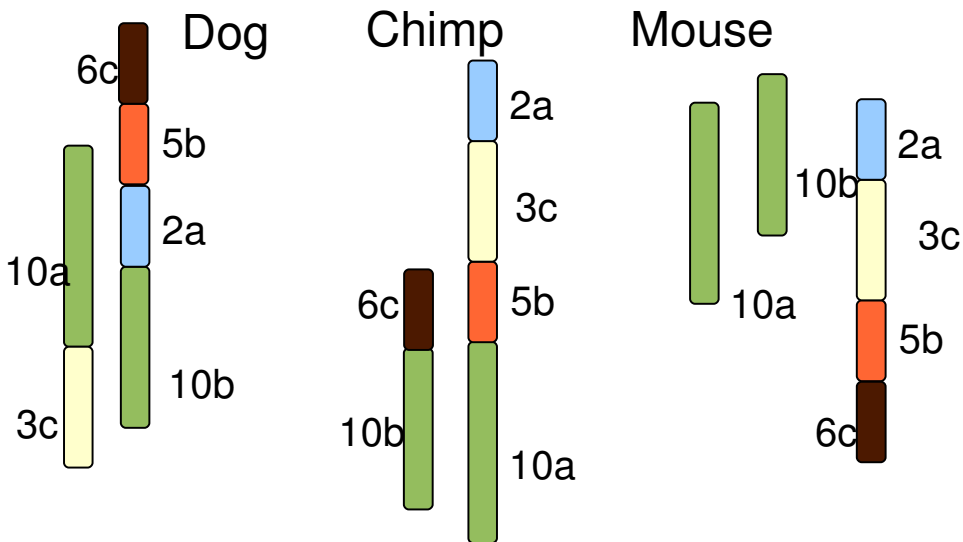
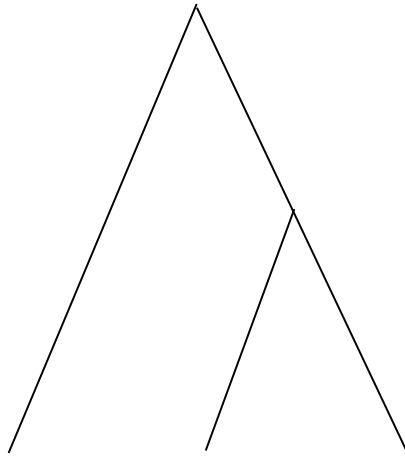
M. Muffato and H. R. Crollius.

BioEssays, 2008

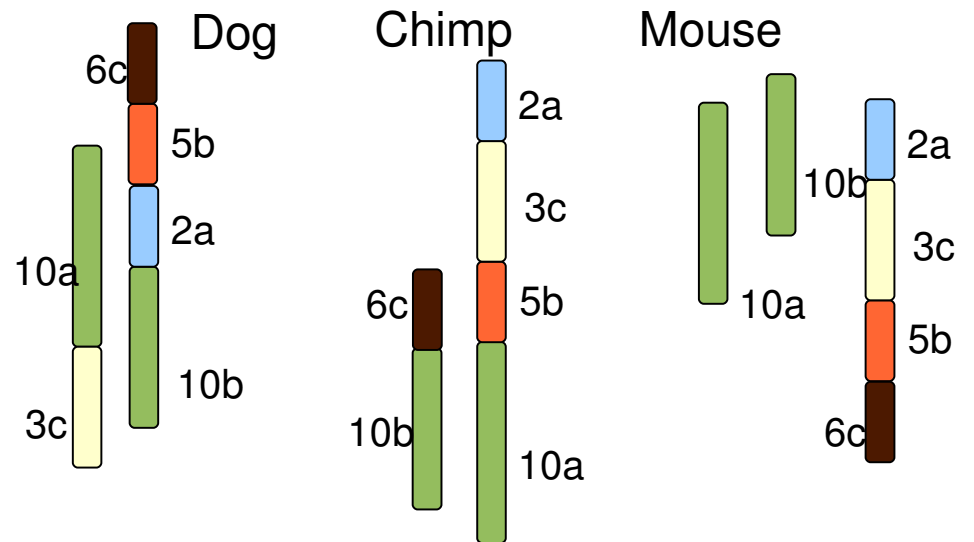
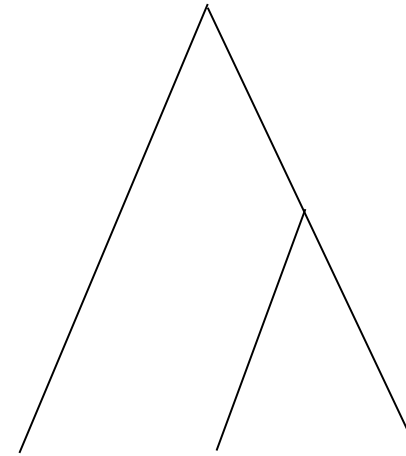


Local and global methods

If some chromosome segments are contiguous in several extant species, then they are probably contiguous in their common ancestor

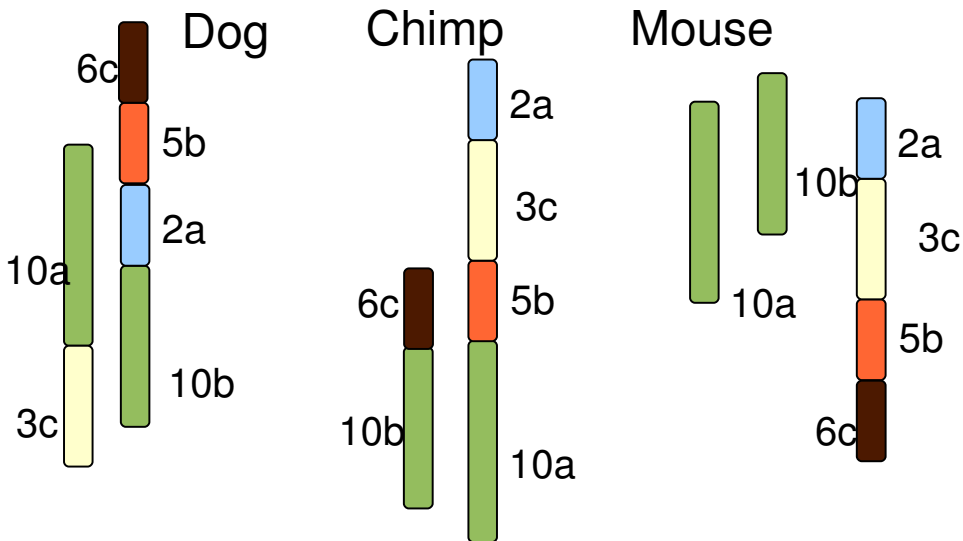
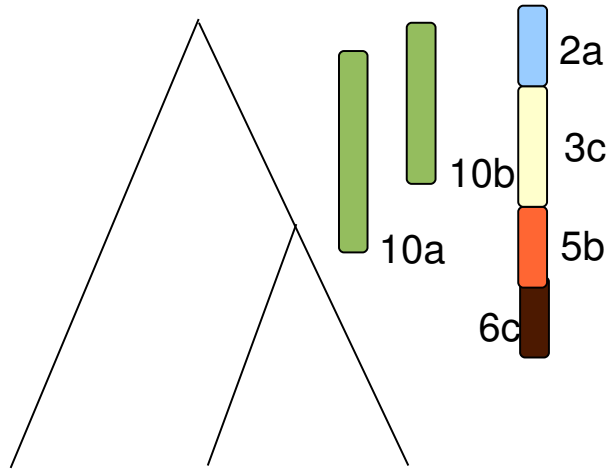


Explain the difference between extant species using a minimum number of evolutionary events

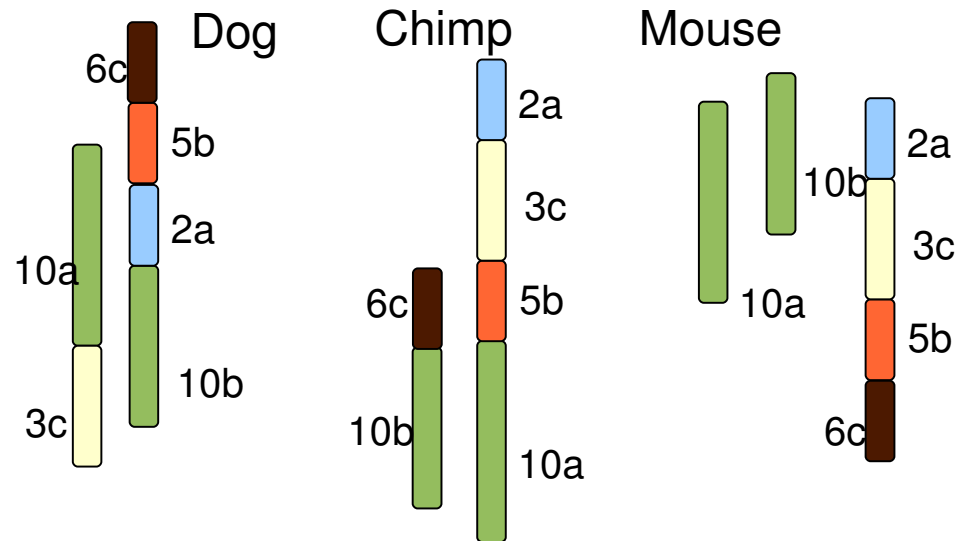
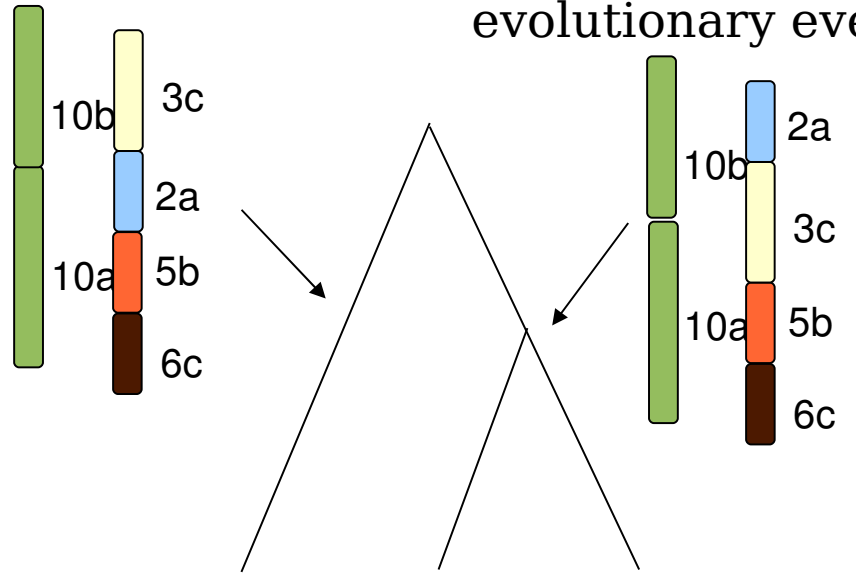


Local and global methods

If some chromosome segments are contiguous in several extant species, then they are probably contiguous in their common ancestor



Explain the difference between extant species using a minimum number of evolutionary events



Are molecular cytogenetics and bioinformatics suggesting diverging models of ancestral mammalian genomes?

Froenicke, Garcia Caldes, Graphodatsky, Muller, Lyons, Robinson, Volleth, Yang, Wienberg

Genome Research, mars 2006.

The convergence of cytogenetics and rearrangement-based models for ancestral genome reconstruction

Bourque, Pevzner, Tesler

Genome Research, mars 2006.

Ancestral genomes reconstruction: An integrated, multi-disciplinary approach is needed

Rocchi, Archidiacono, Stanyon

Genome Research, décembre 2006.



**GENOME
RESEARCH**

Two different reconstructions of the same ancestral genome (boreoeutherian)

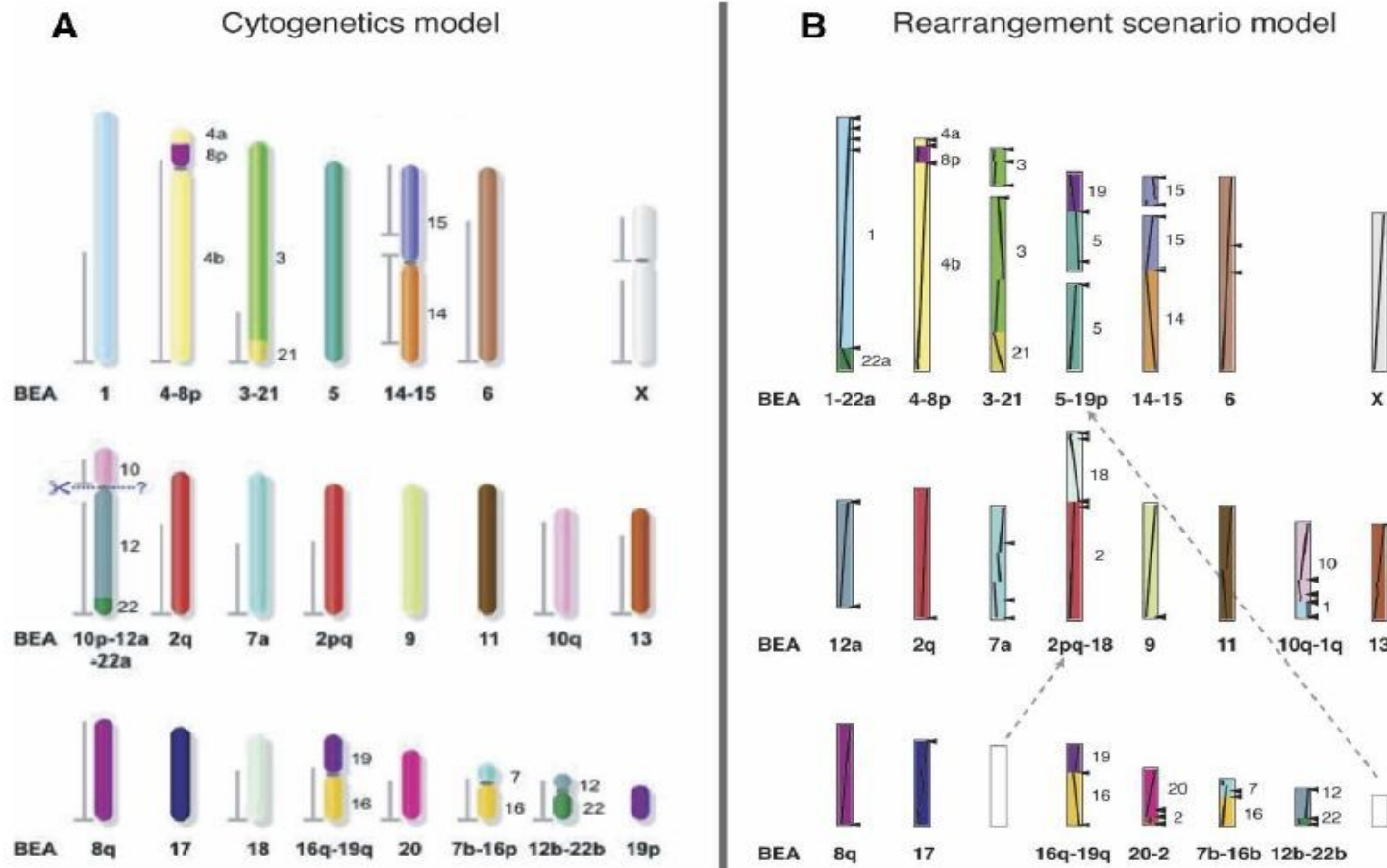
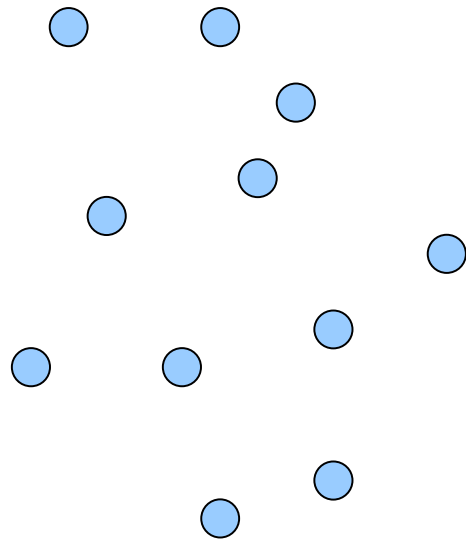


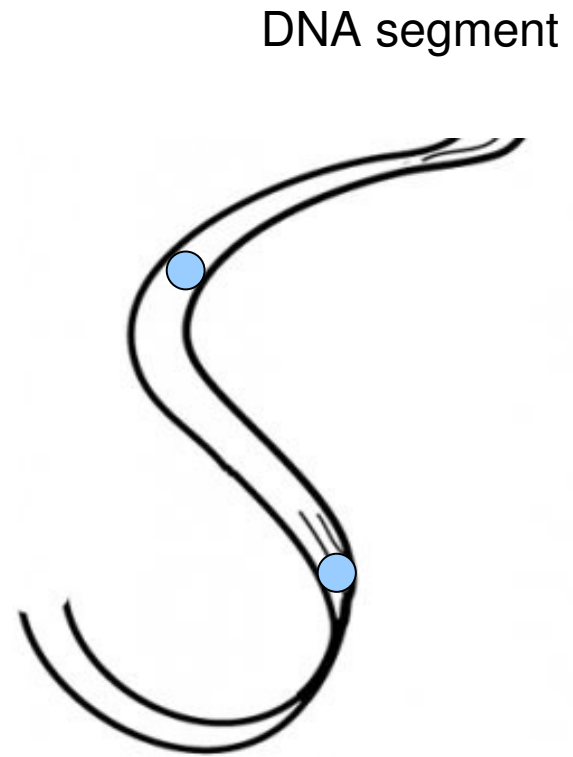
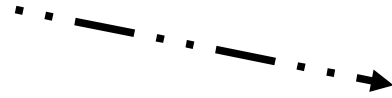
Figure 1. Putative genome architecture of the boreoeutherian ancestor from (A) cytogenetics model using over 80 eutherian genomes (Froenicke et al. 2006) and (B) a rearrangement scenario model using human–mouse–rat–cat–cattle–dog–pig (Murphy et al. 2005). In B, the black arrowheads to the right of the ancestral chromosomes indicate unresolved/weak adjacencies, and the diagonal line segments indicate original position and orientation on human. This is the same reconstruction as in Murphy et al. (2005), but reformatted, rescaled, and recolored to facilitate comparison. The two dashed arrows show the location in the rearrangement model of two of the ancestral chromosomes from the cytogenetics model. These two chromosomes are only weakly associated with two other chromosomes in the rearrangement model.

A general framework for local methods

Physical mapping technique

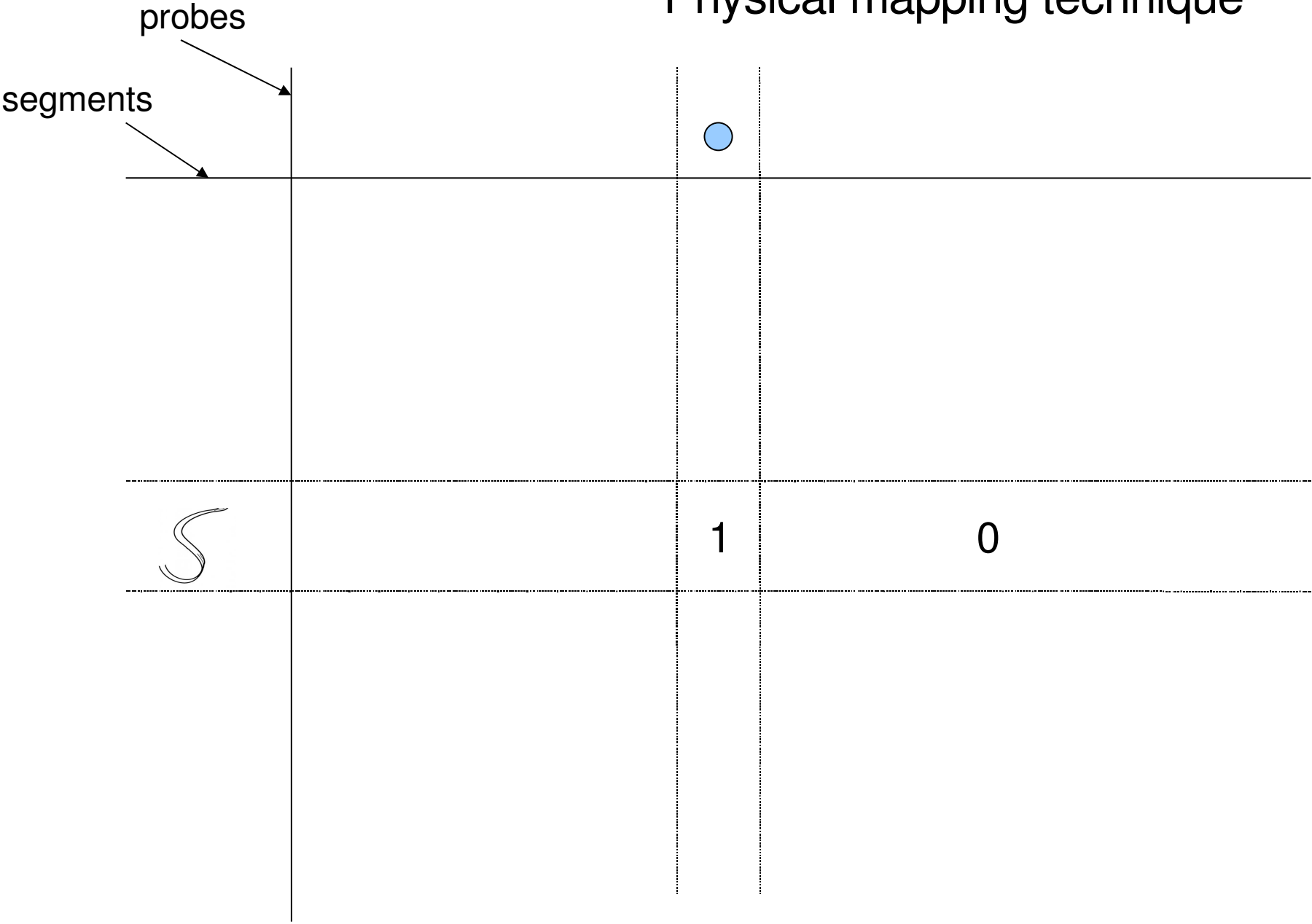


Probes



DNA segment

Physical mapping technique

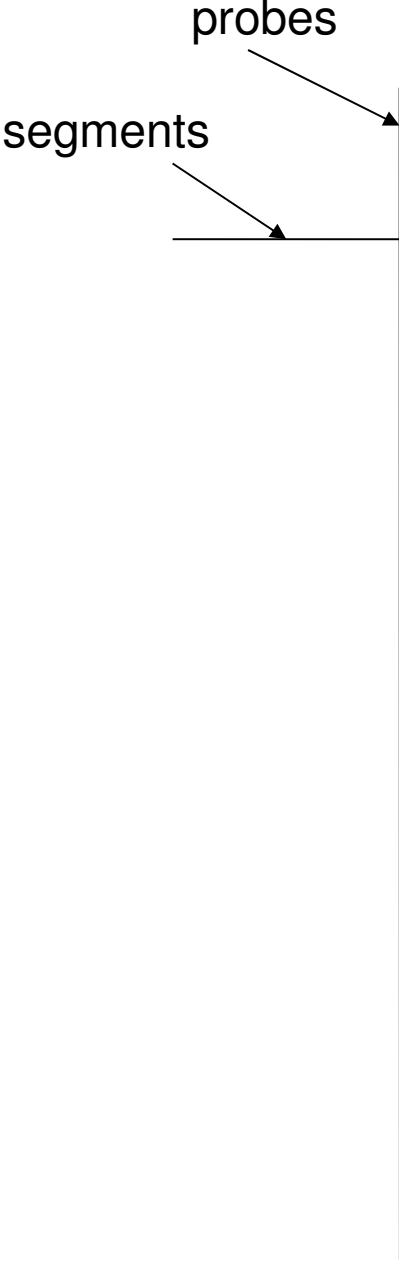


Physical mapping technique



| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 |
| 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| 4 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |

Physical mapping technique



| | G | H | A | B | C | D | E | F | I |
|--|---|---|---|---|---|---|---|---|---|
| | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |

Physical mapping technique

The diagram shows a coordinate system where the vertical axis is labeled 'probes' and the horizontal axis is labeled 'segments'. The matrix below is a 5x10 grid of binary values (0s and 1s) representing the mapping between segments and probes.

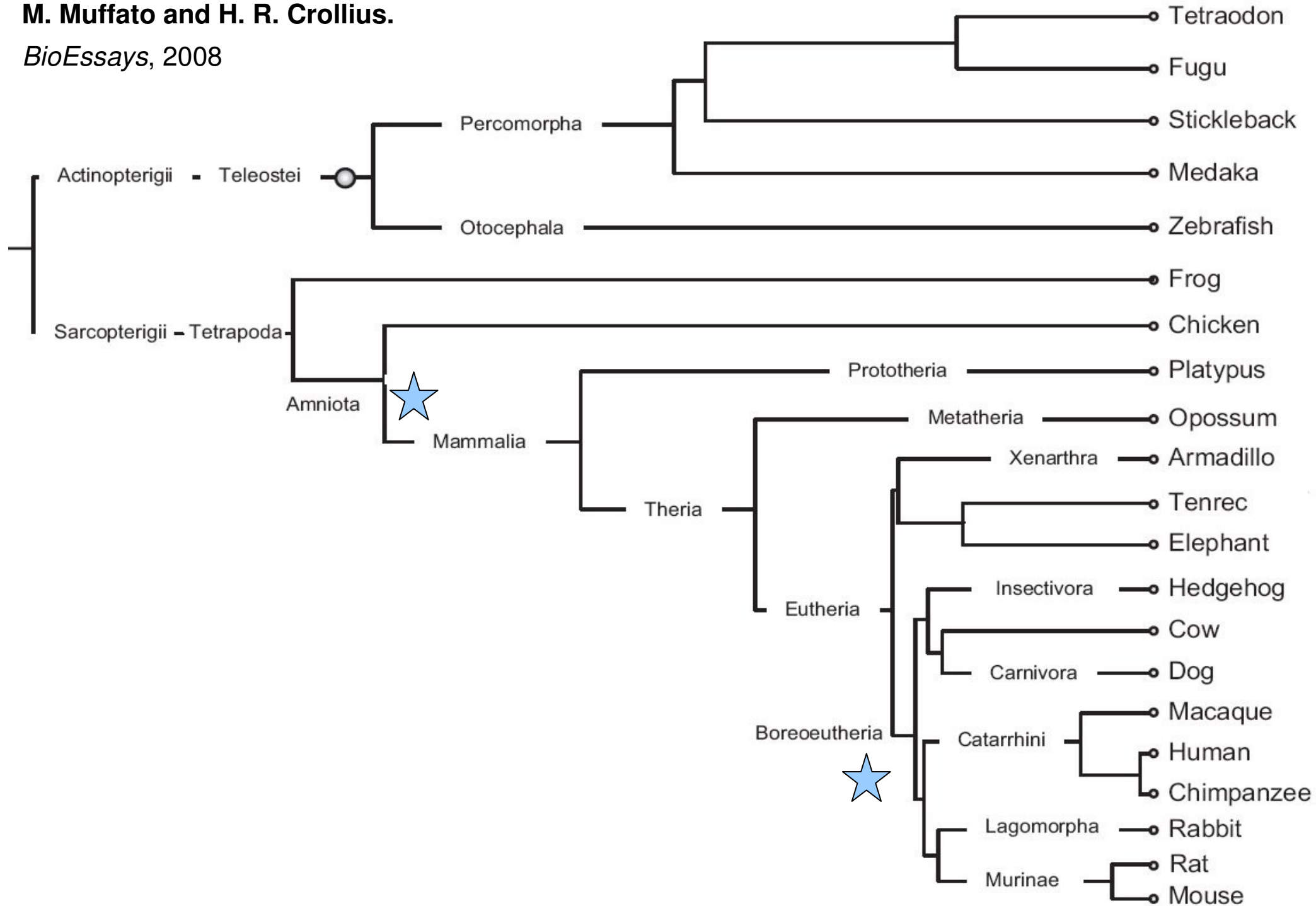
| | G | H | A | B | C | D | F | E | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |

Linear time Method: McConnell, *SODA*, 2005

How to obtain ancestral segments

M. Muffato and H. R. Crollius.

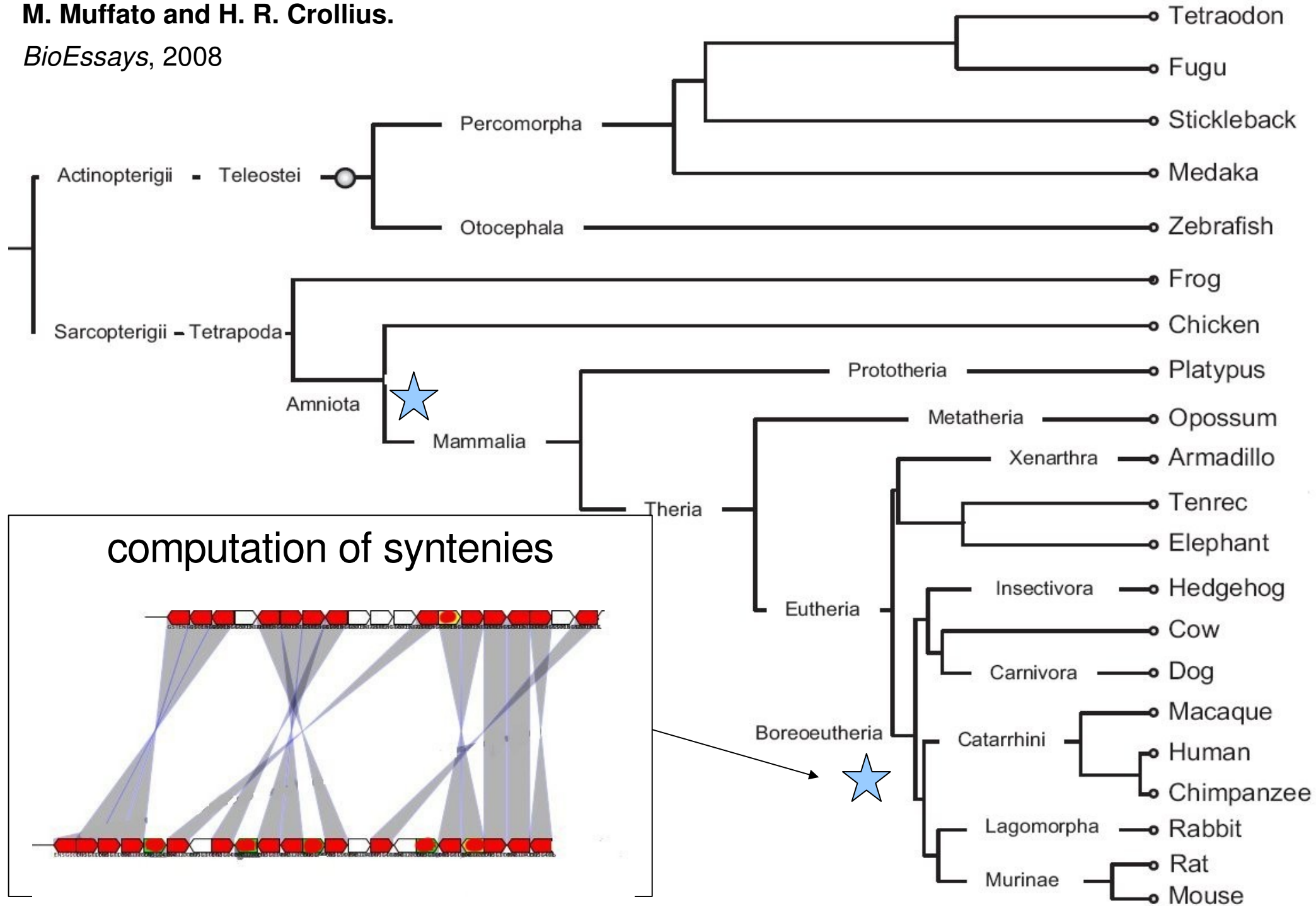
BioEssays, 2008



How to obtain ancestral segments

M. Muffato and H. R. Crollius.

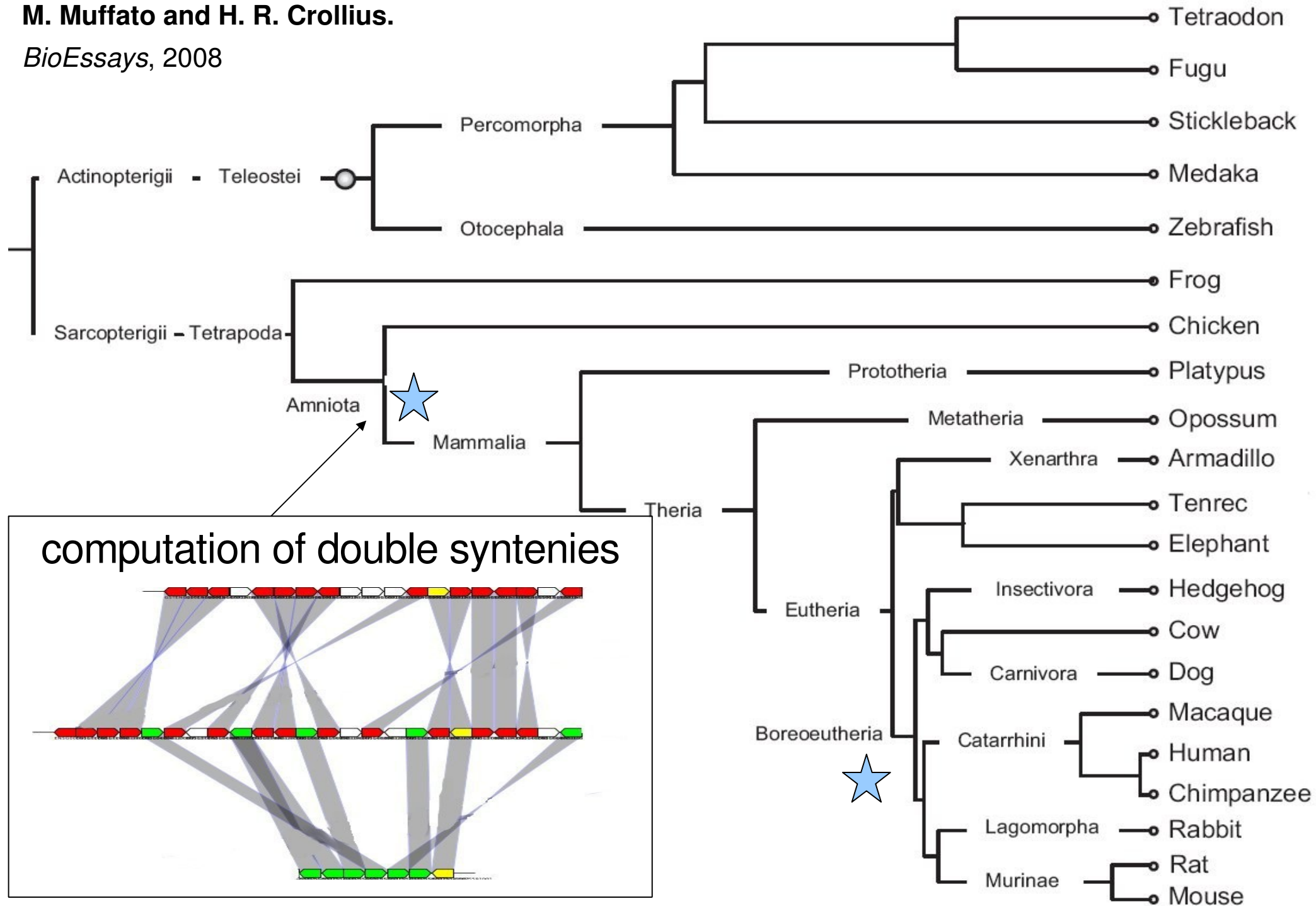
BioEssays, 2008



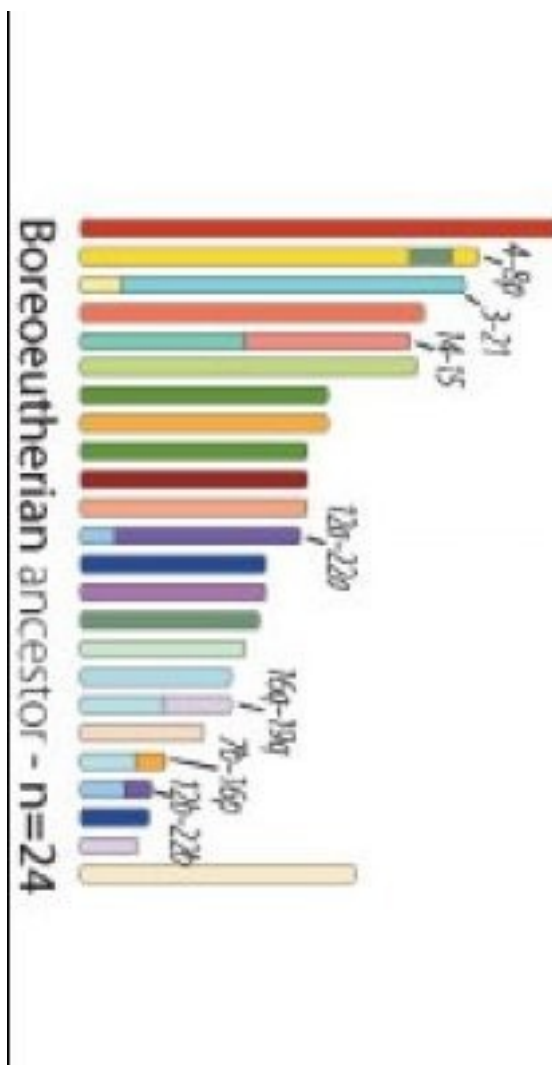
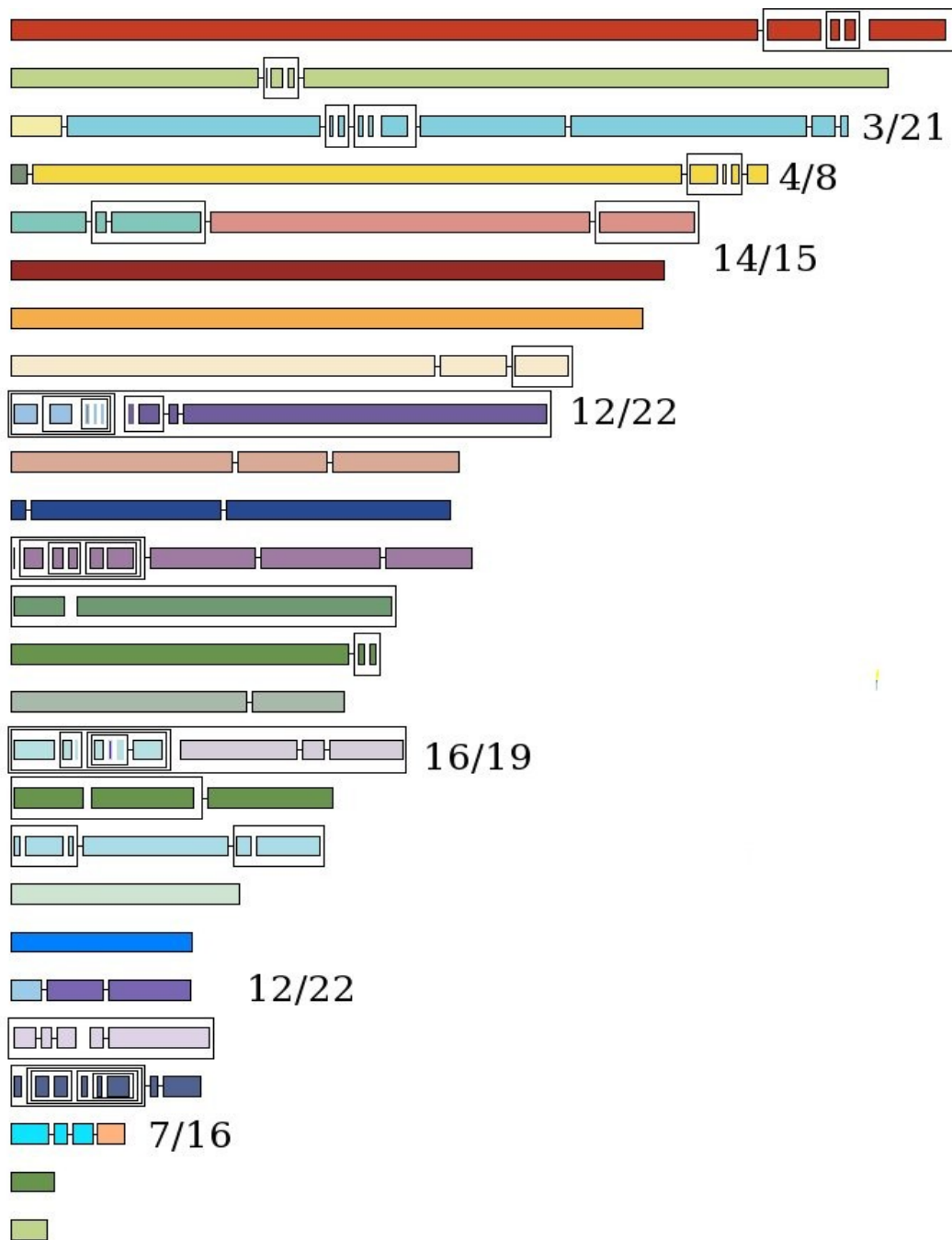
How to obtain ancestral segments

M. Muffato and H. R. Crollius.

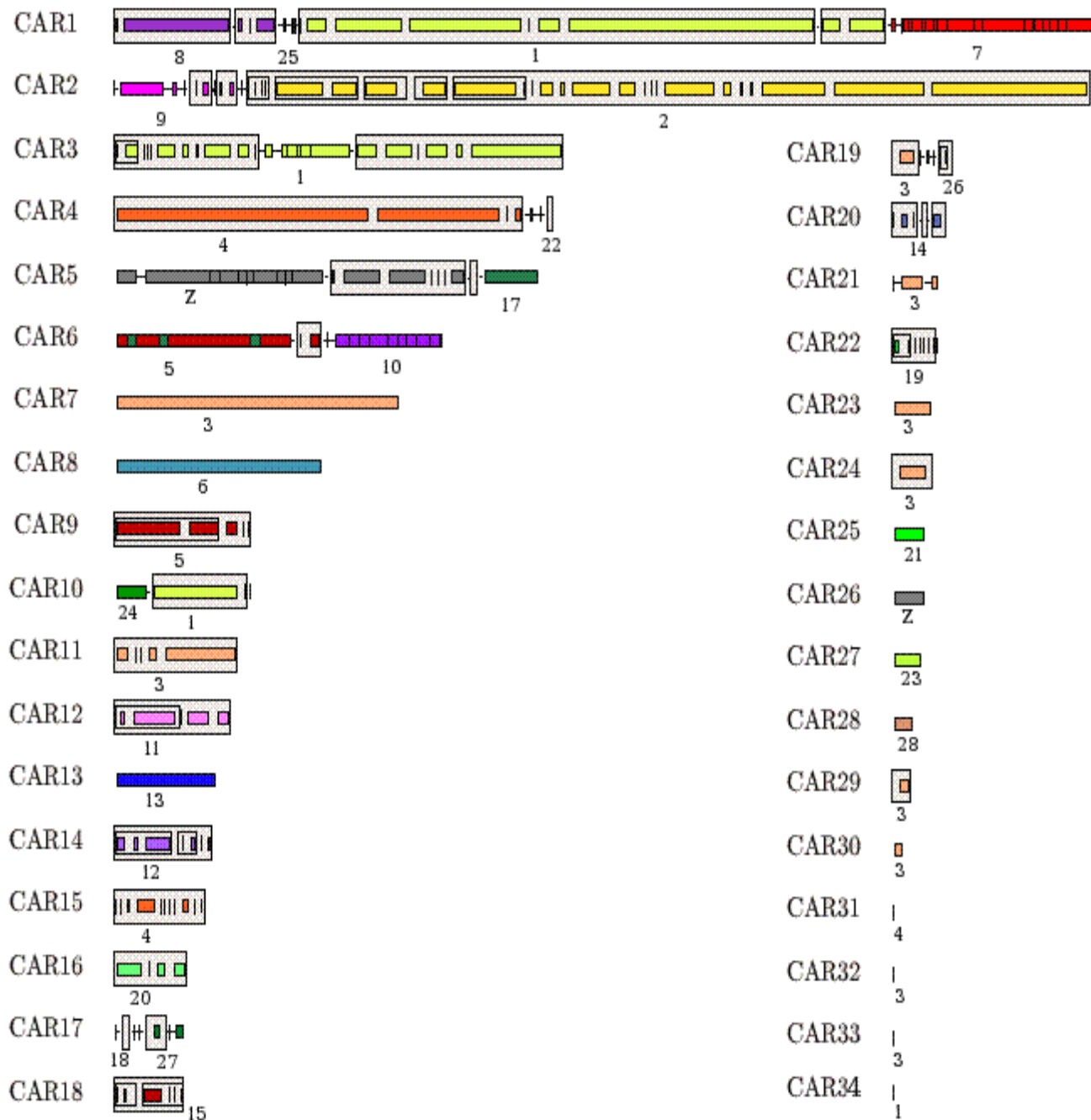
BioEssays, 2008



26 segments of the boreoeutherian ancestor



Segments of the amniote ancestor



Segments of the amniote ancestor

Nakatani et al, *Genome Research*, 2007

Kohn et al, *Trends in Genetics*, 2006

present method

| | | |
|---|---------------------------------|--|
| 2-9-16, 1-24, 3-14, 5-10, 17-Z, 4-22, 8-18, 18-27-19, 21-23-26-32 | 2-9, 1-14-18 13-17-Z, 1-7 | 2-9, 1-24, 5-10 17-Z, 4-22, 18-27 8-25-1-7, 3-26 |
|---|---------------------------------|--|

syntenic associations between chicken chromosome segments

Future works

- ✗ In case of absence of solution to the consecutive one's (C1) problem, an optimisation phase is needed (we discard a minimum number of putative ancestral segments).
- ✗ A better understanding of the combinatorics of non C1 matrices would improve this framework.