

# **Phylogenetic Biology**

## **Week 6**

Biology 1425

Professor: Casey Dunn, [dunnlab.org](http://dunnlab.org)

Brown University

2013

# Front matter...

All original content in this document is distributed under the following license:

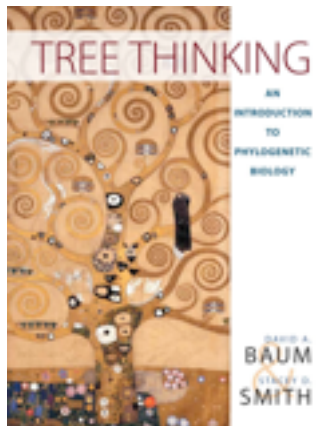


Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License  
([http://creativecommons.org/licenses/by-nc-sa/3.0/deed.en\\_US](http://creativecommons.org/licenses/by-nc-sa/3.0/deed.en_US))

See sources for copyright of non-original content

# Sources

Some non-original content is drawn from:



Baum, D and S. Smith (2012) Tree Thinking: and Introduction to Phylogenetic Biology. Roberts and Company Publishers. ISBN 9781936221165

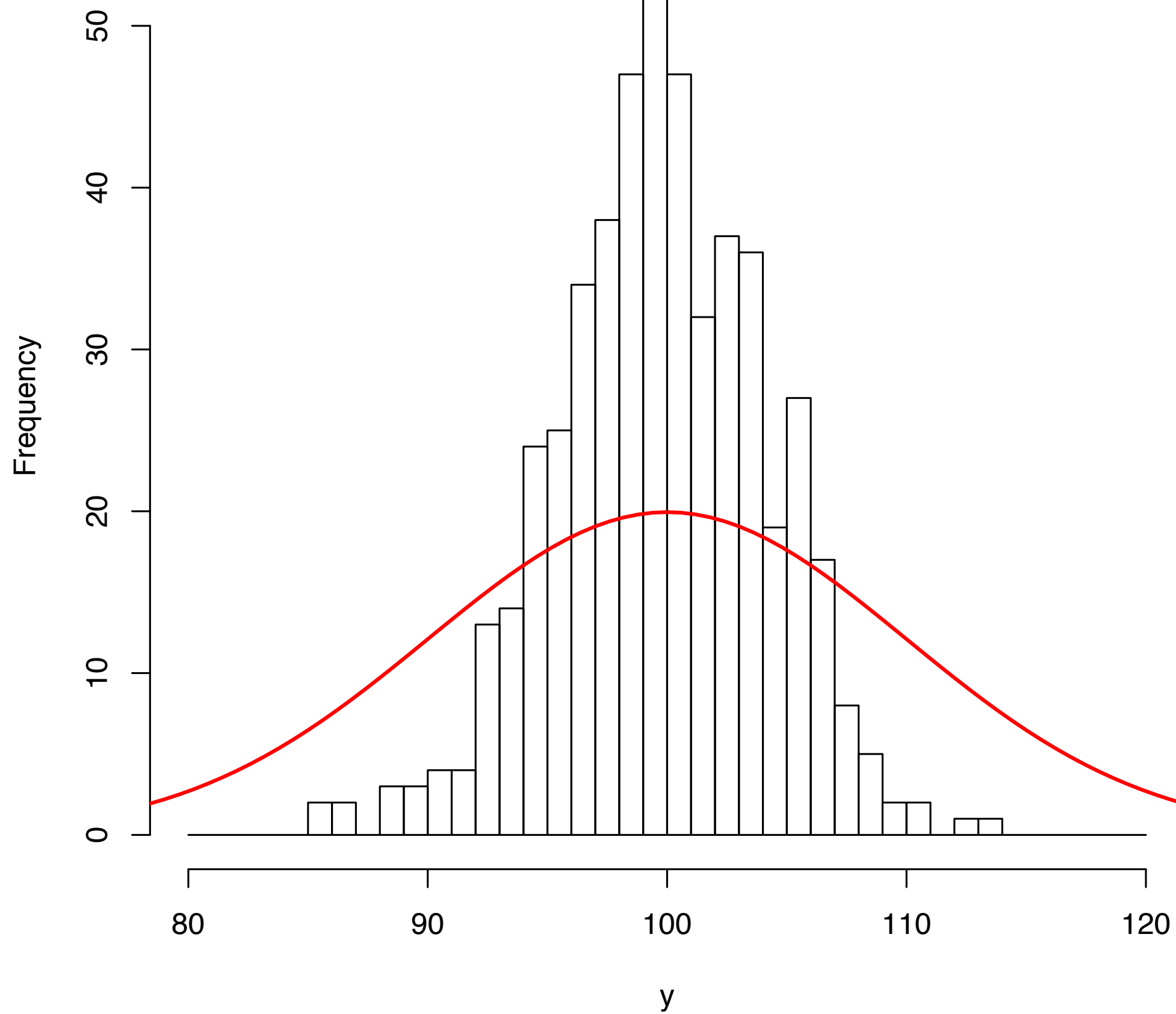
Other non-original content is referenced by url.

# Consistency

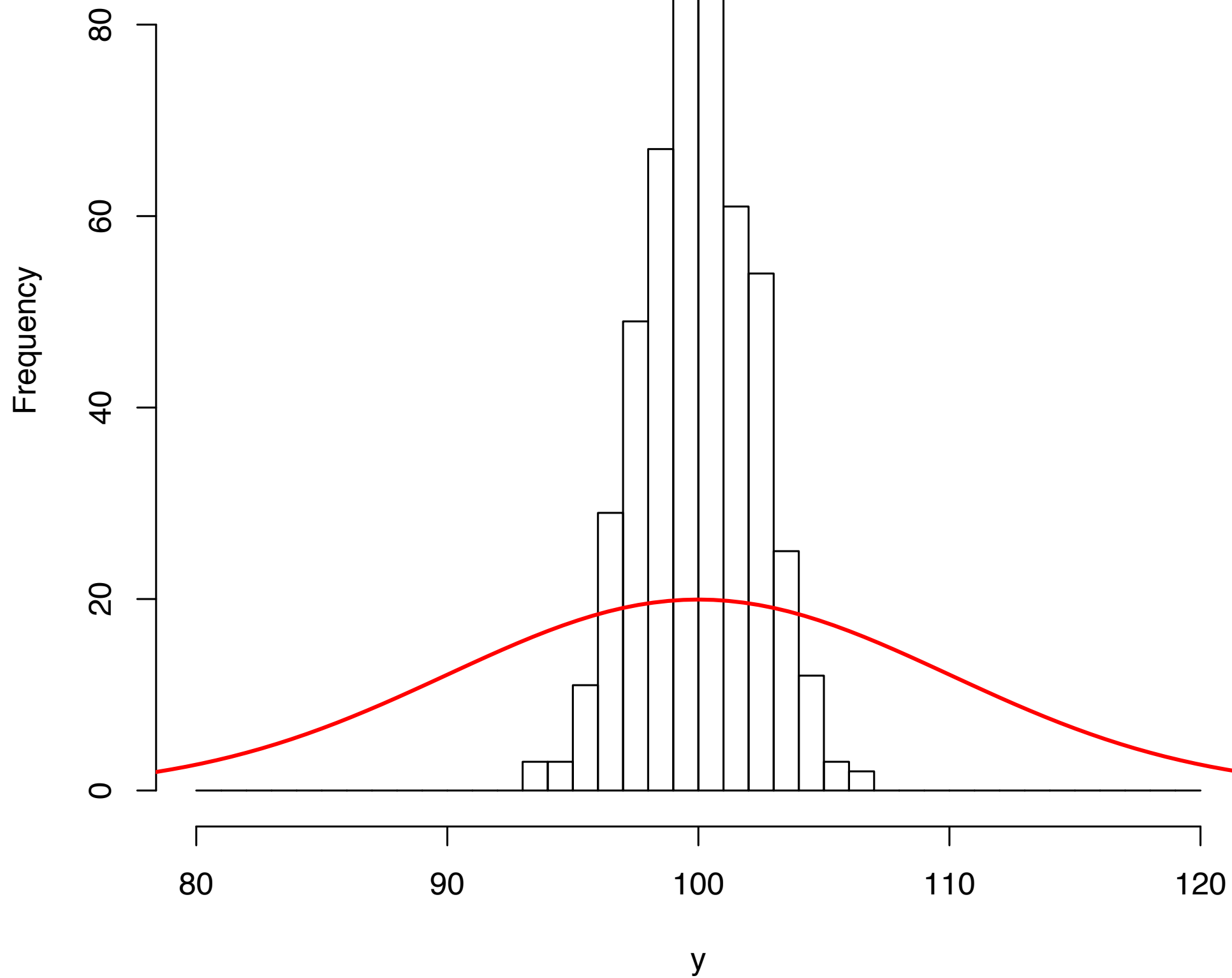
The ability of an estimation method to converge to a true value as more data are added

A trivial example: Does increasing sample size improve our ability to estimate the mean of a normally distributed population?

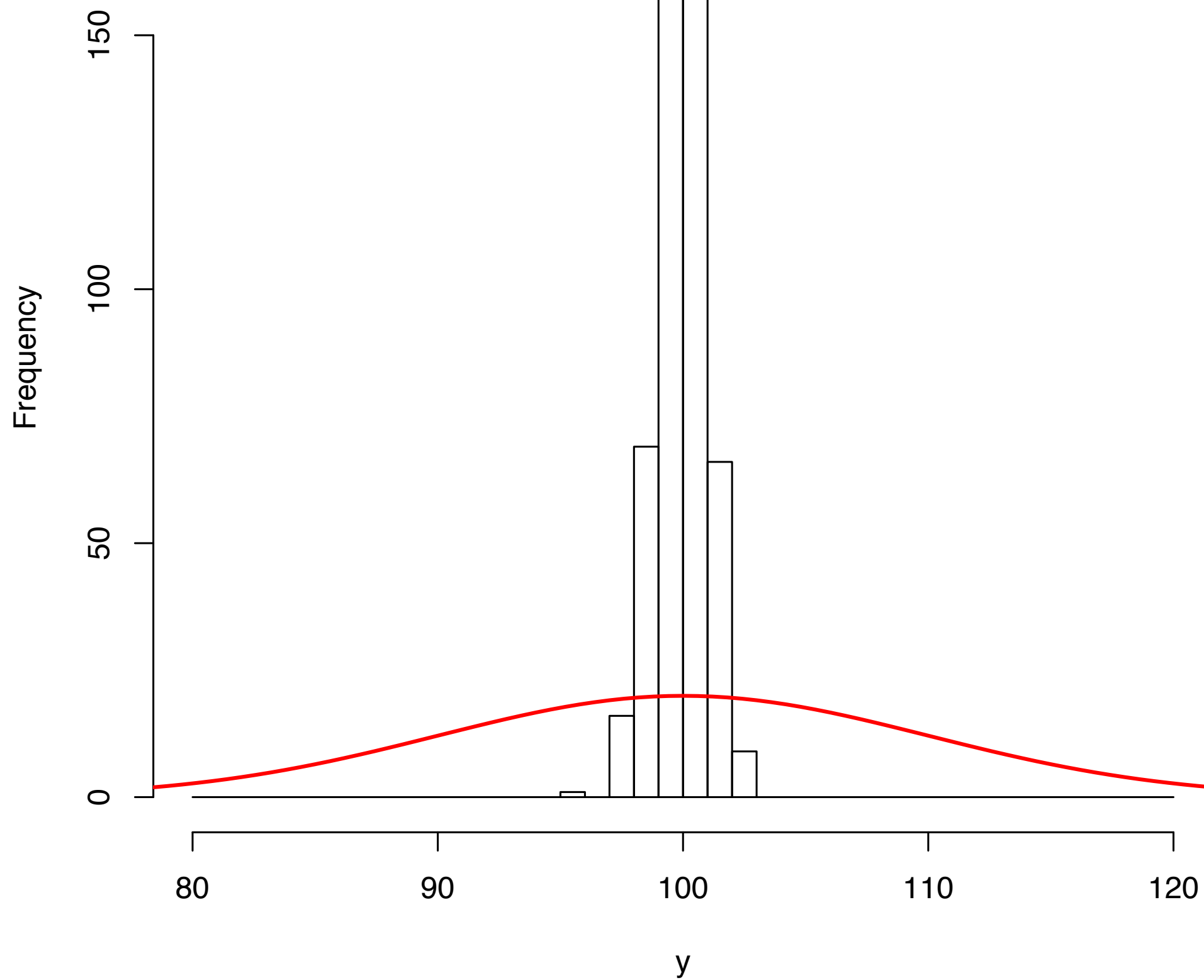
**Sample size: 5**



**Sample size: 20**

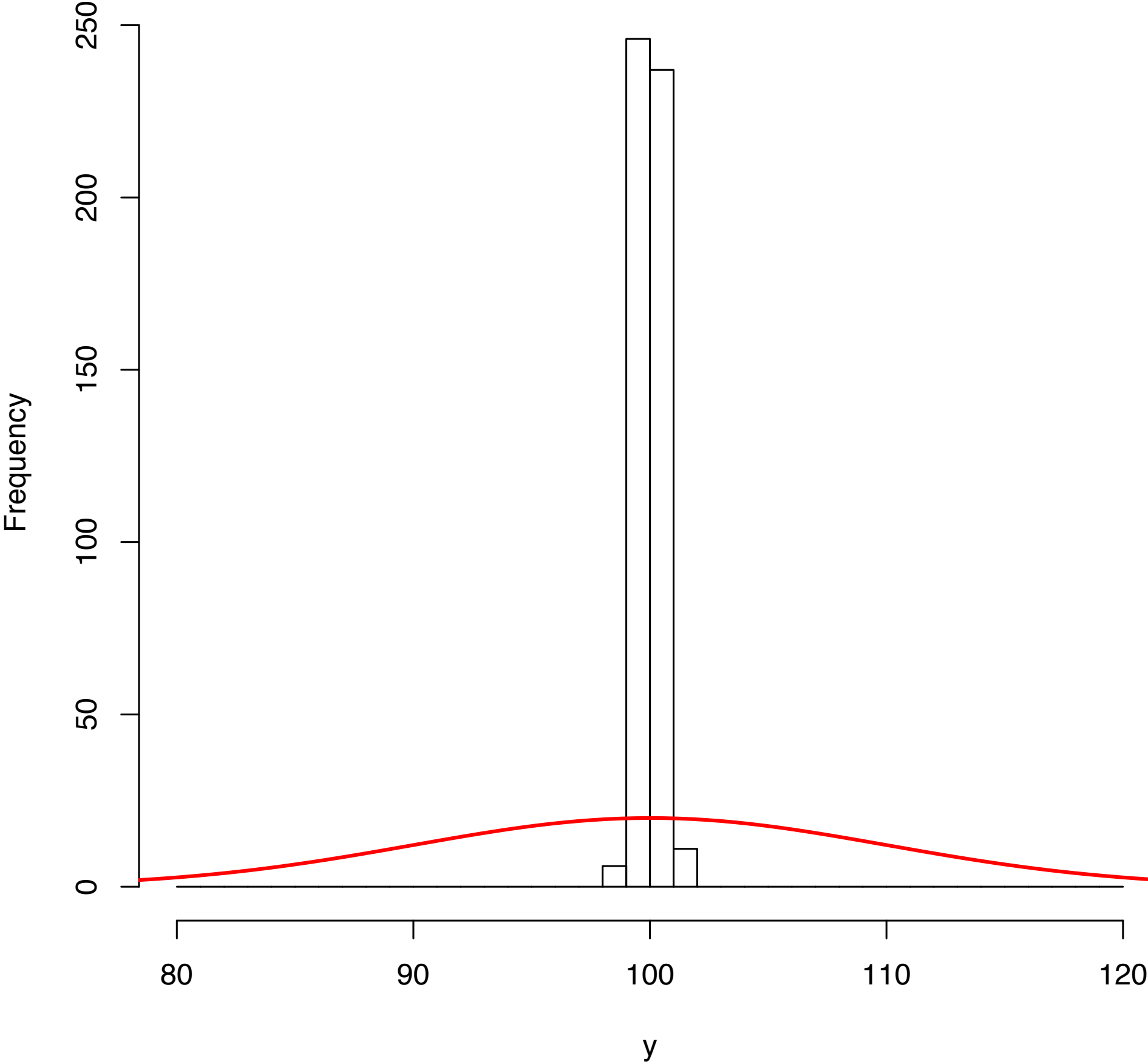


**Sample size: 100**





**Sample size: 500**



**Is phylogenetic inference  
consistent?**

# Distance methods

The distance between two taxa is the sum of the branch lengths that separate them.

# Distance methods

The distance between two taxa is the sum of the branch lengths that separate them.

If you know all the distances, you can reconstruct the tree.

# Distance methods

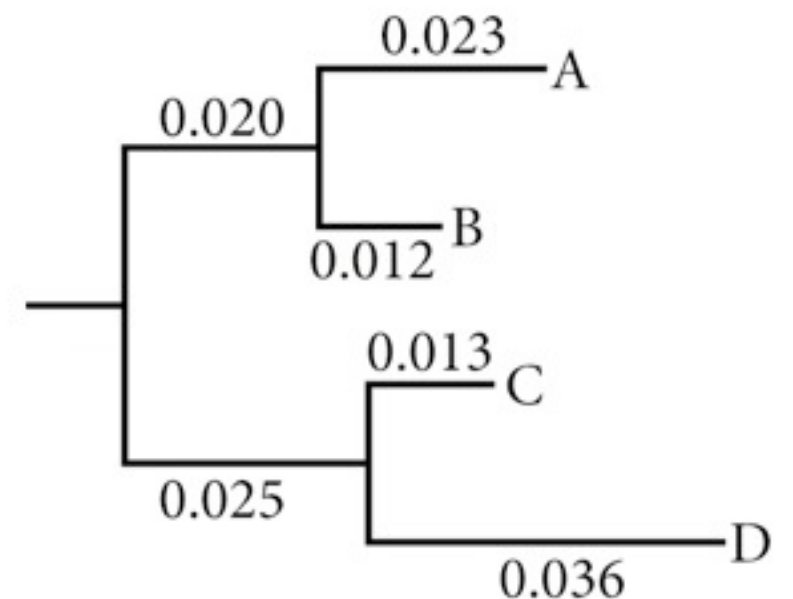
The fraction of characters that differ between taxa A, B, C, D:

	A	B	C	D
A	0			
B	0.035	0		
C	0.081	0.070	0	
D	0.104	0.093	0.049	0

# Distance methods

The fraction of characters that differ between taxa A, B, C, D:

	A	B	C	D
A	0			
B	0.035	0		
C	0.081	0.070	0	
D	0.104	0.093	0.049	0



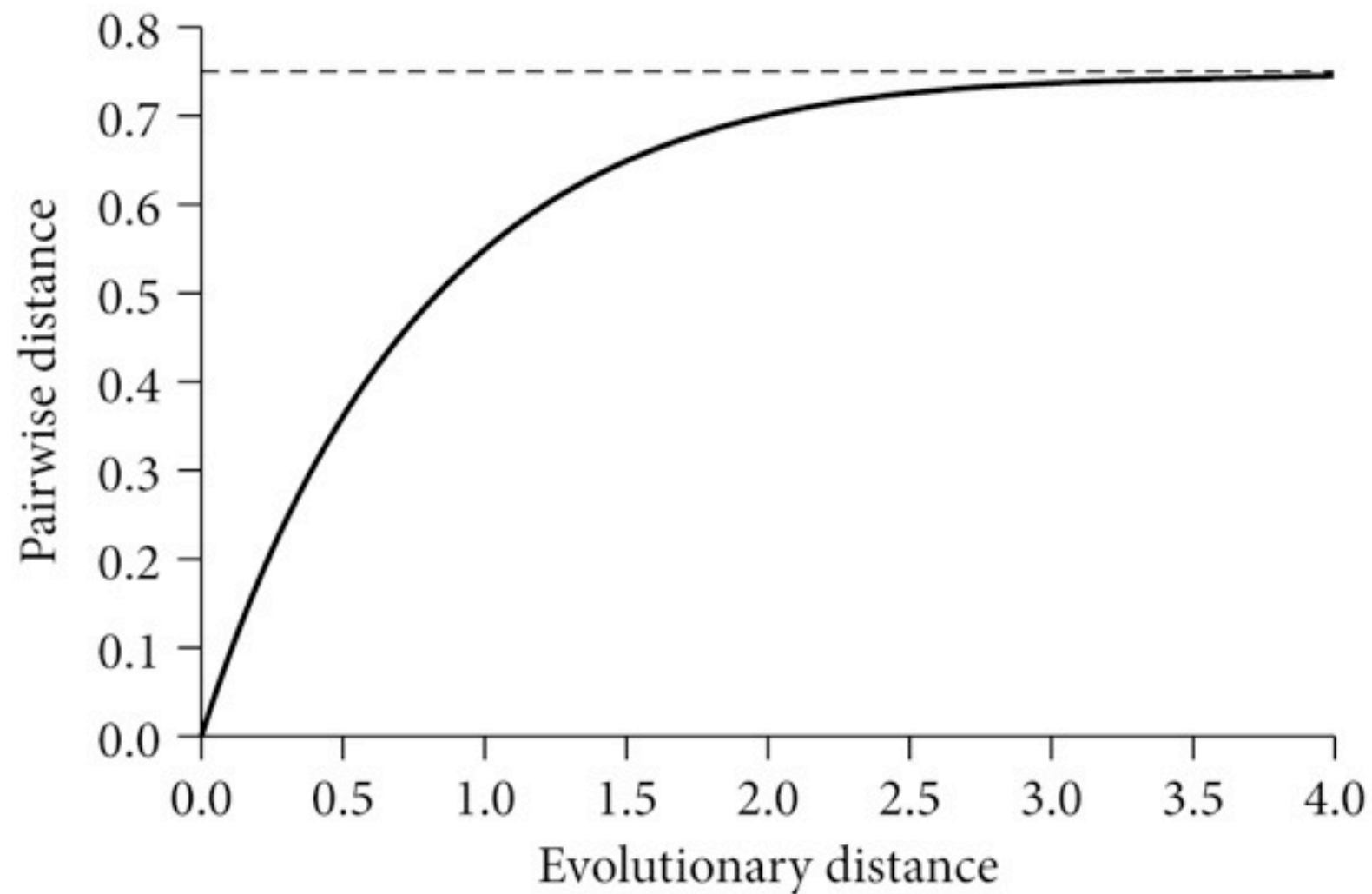
Baum and Smith 2012, Figure 8.11

# Distance methods

That was easy. Why did we bother going through parsimony, maximum likelihood, and Bayesian methods?

# Distance methods

Because observed distances aren't evolutionary distances...



Baum and Smith 2012, Figure 8.12

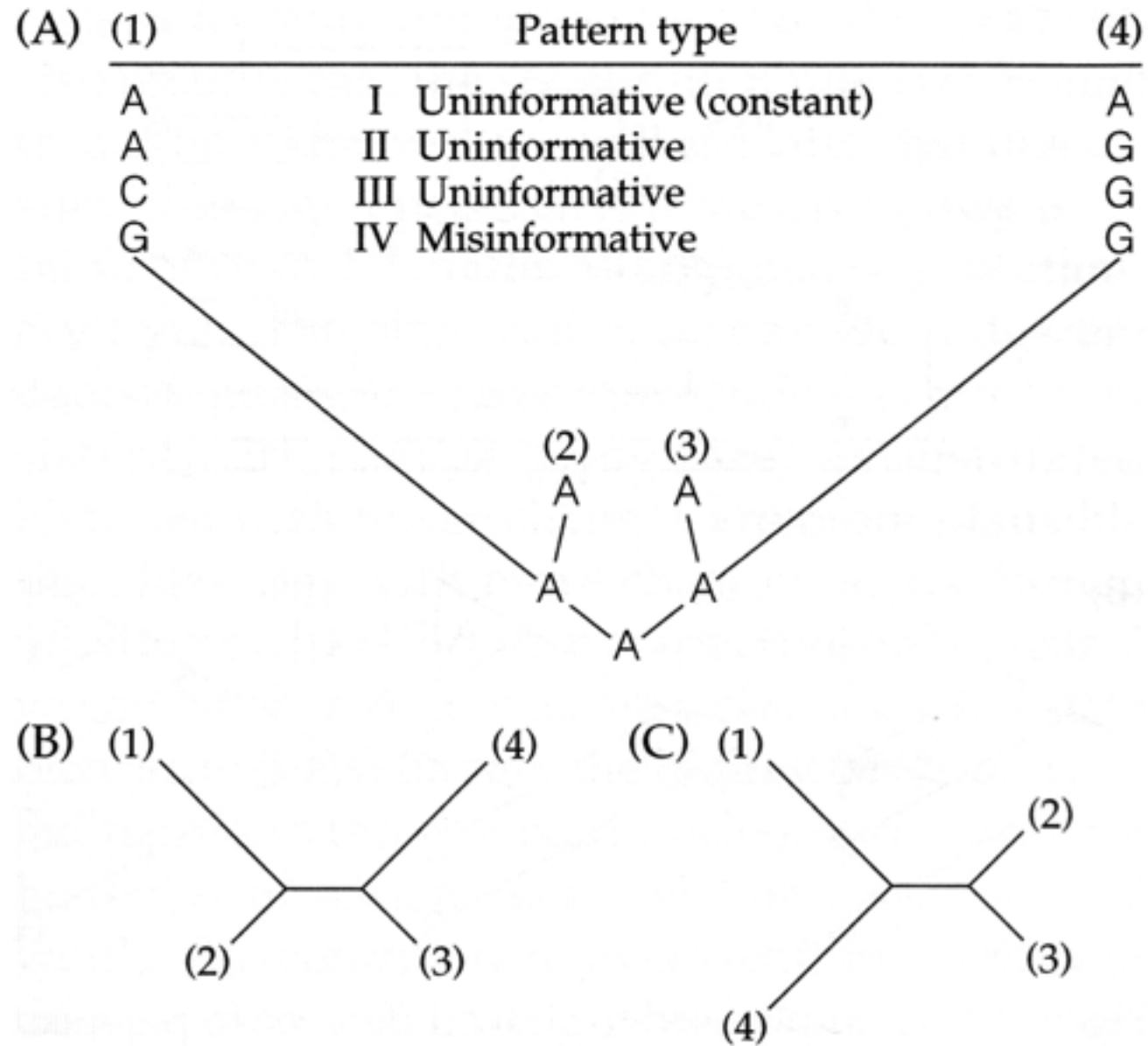


# Is phylogenetic inference consistent?

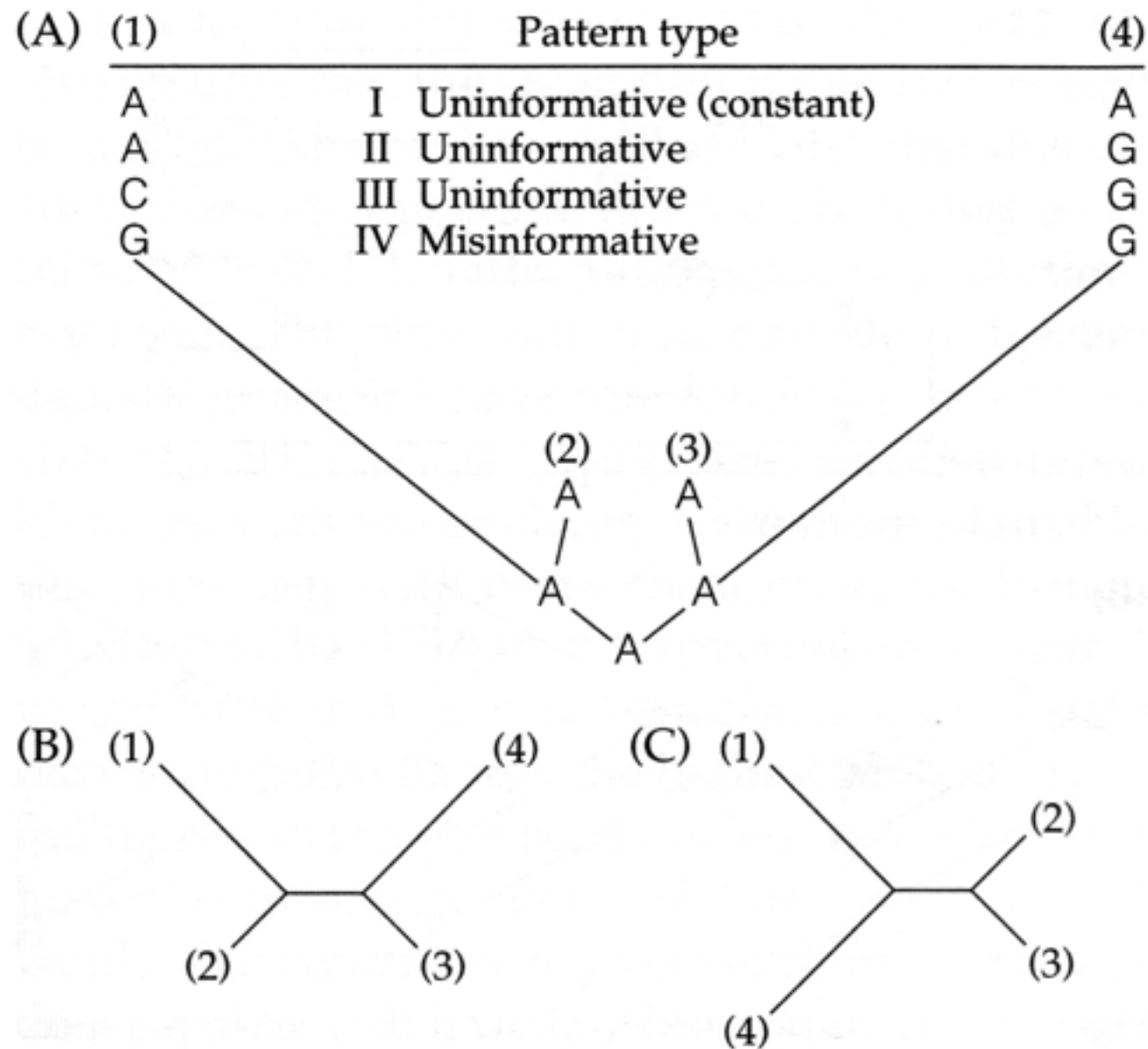
Not necessarily. This depends on a variety of factors, including:

- Inference method
- Model adequacy
- The structure of the true tree

# The Felsenstein Zone

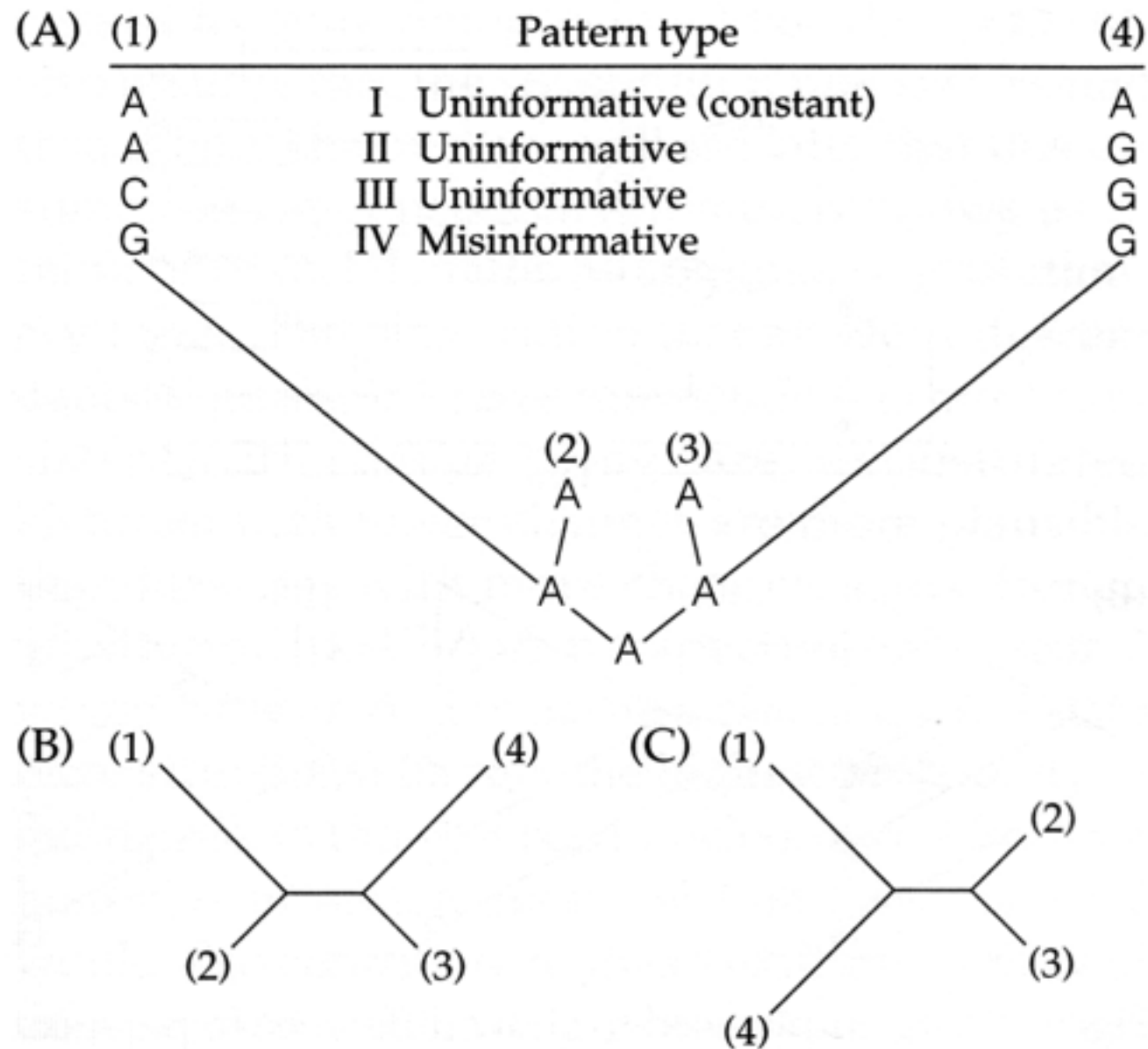


# The Felsenstein Zone



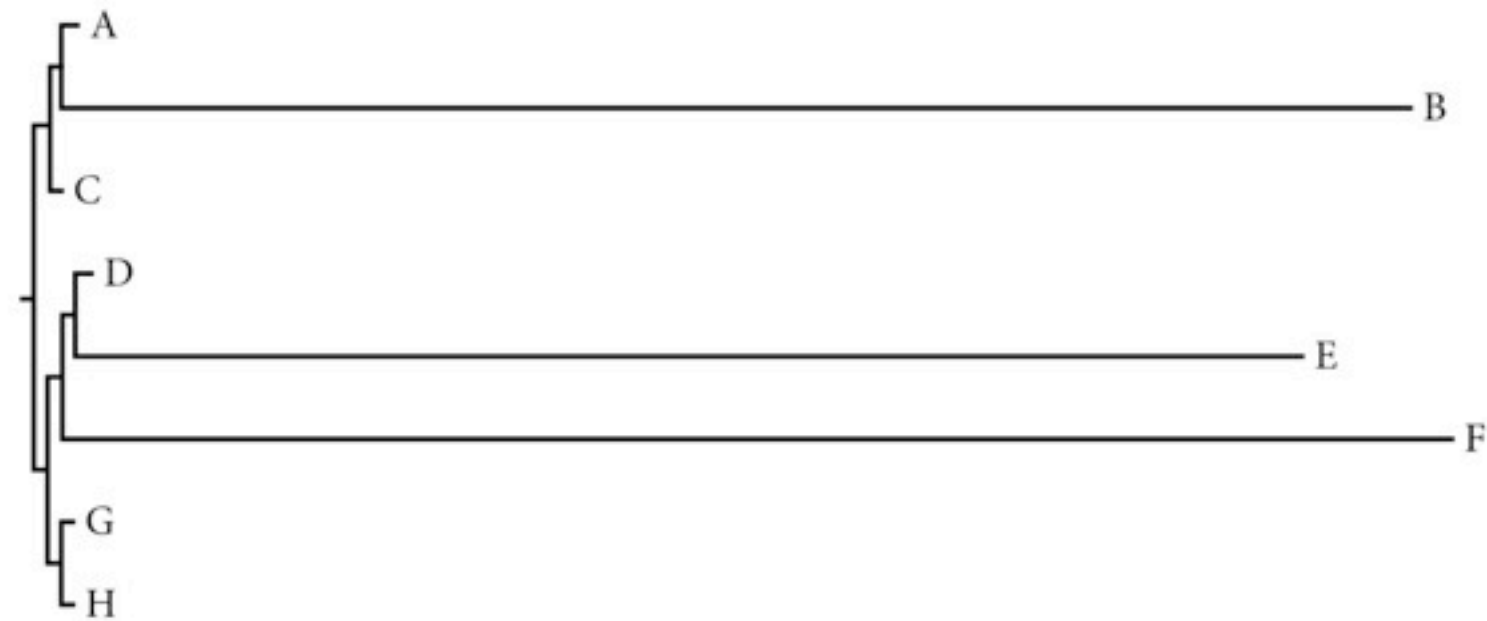
Methods that do account for unobserved changes can be consistent.

# The Felsenstein Zone

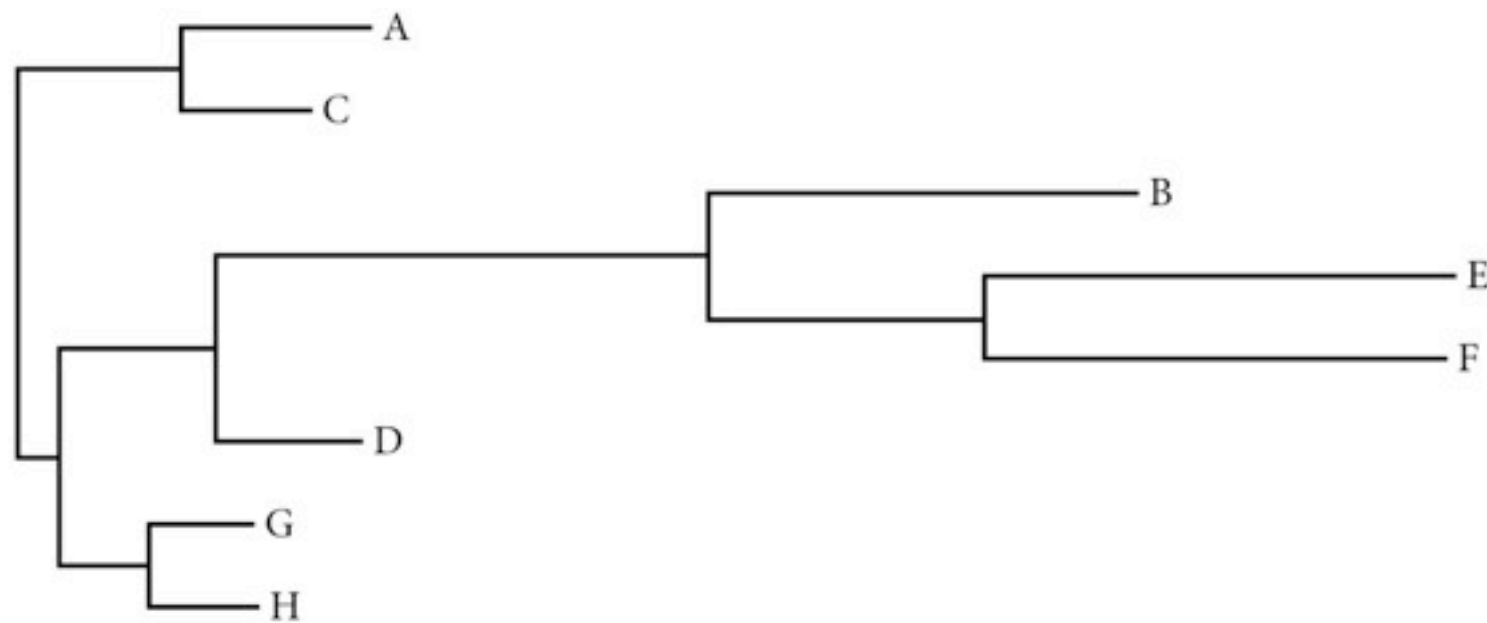


Methods that don't account for unobserved changes will be inconsistent.

# Long branch attraction



True tree



Unrooted tree estimated by parsimony

Baum and Smith 2012, Figure 7.14

Inadequate models can lead to  
inconsistent results by failing to  
accommodate unobserved changes

Does this mean that more complex models are always better?

Does this mean that more complex models are always better?

No. Just because it is more complex doesn't mean it is correctly accounting for unobserved changes.



Does this mean that more complex models are always better?

No. Just because it is more complex doesn't mean it is correctly accounting for unobserved changes.

More complex models come with their own costs, such as higher computational requirements. This means that less tree space is explored with a given amount of computational resources.

What to do?

Among other things, sensitivity tests.