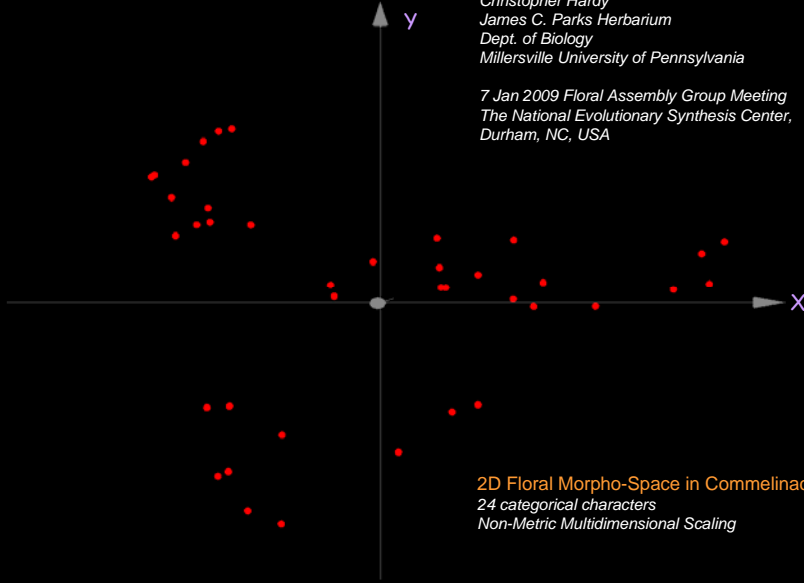


Some Aspects of Ancestor Reconstruction
in the Study of Floral Assembly

Christopher Hardy
James C. Parks Herbarium
Dept. of Biology
Millersville University of Pennsylvania

7 Jan 2009 Floral Assembly Group Meeting
The National Evolutionary Synthesis Center,
Durham, NC, USA



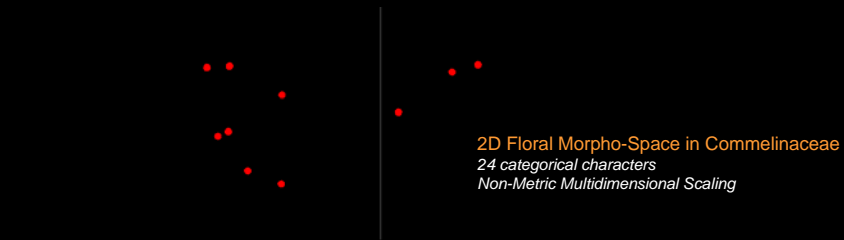
2D Floral Morpho-Space in Commelinaceae
24 categorical characters
Non-Metric Multidimensional Scaling

Some Aspects of Ancestor Reconstruction
in the Study of Floral Assembly

1. Example of Ancestor Reconstruction Methods from
Comelinaceae (spiderworts & dayflowers).

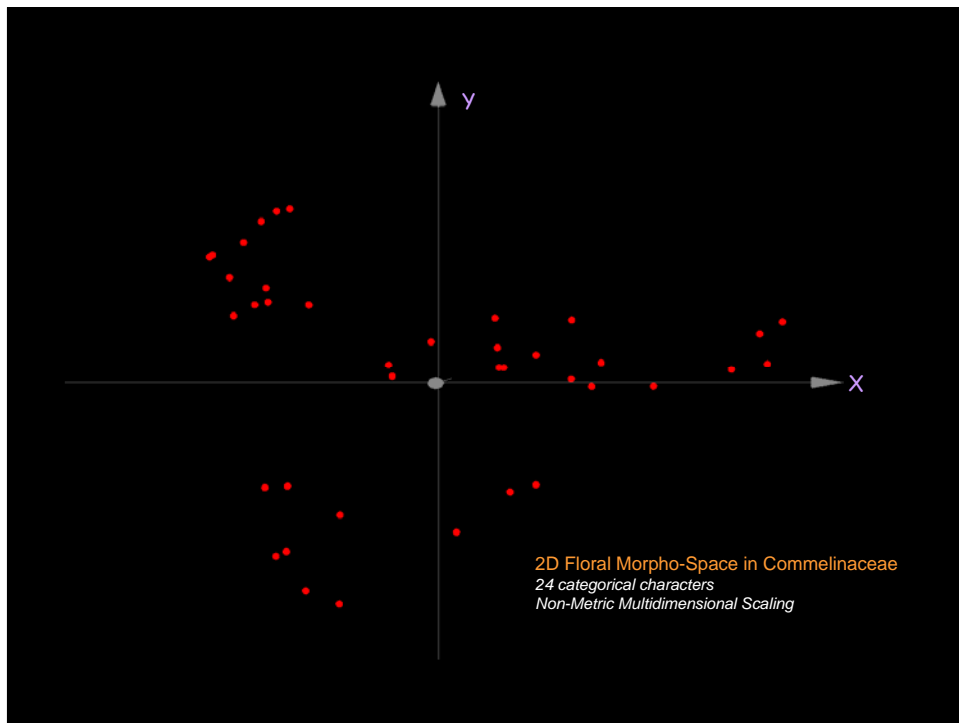
2. Assumptions of Parsimony and Model-Based Methods
-Emphasis on Branch Lengths.
-Programs for Ancestor Reconstruction.

3. Accounting for Uncertainty.

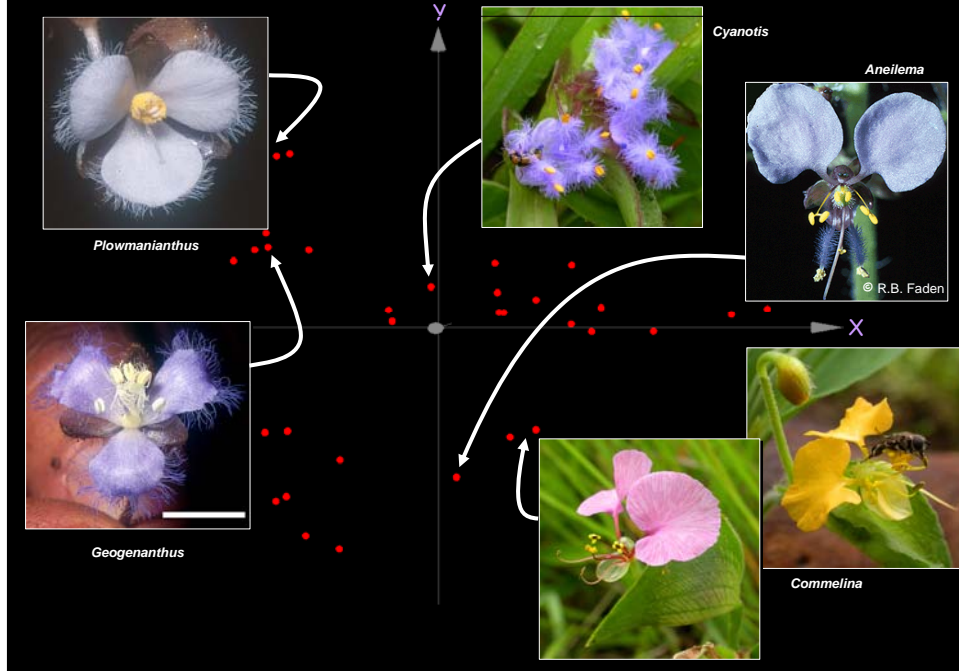


2D Floral Morpho-Space in Commelinaceae
24 categorical characters
Non-Metric Multidimensional Scaling

1. Example of Ancestor Reconstruction Methods from
Commelinaceae (spiderworts & dayflowers).



Flowers exhibit a high degree of *synorganization*.

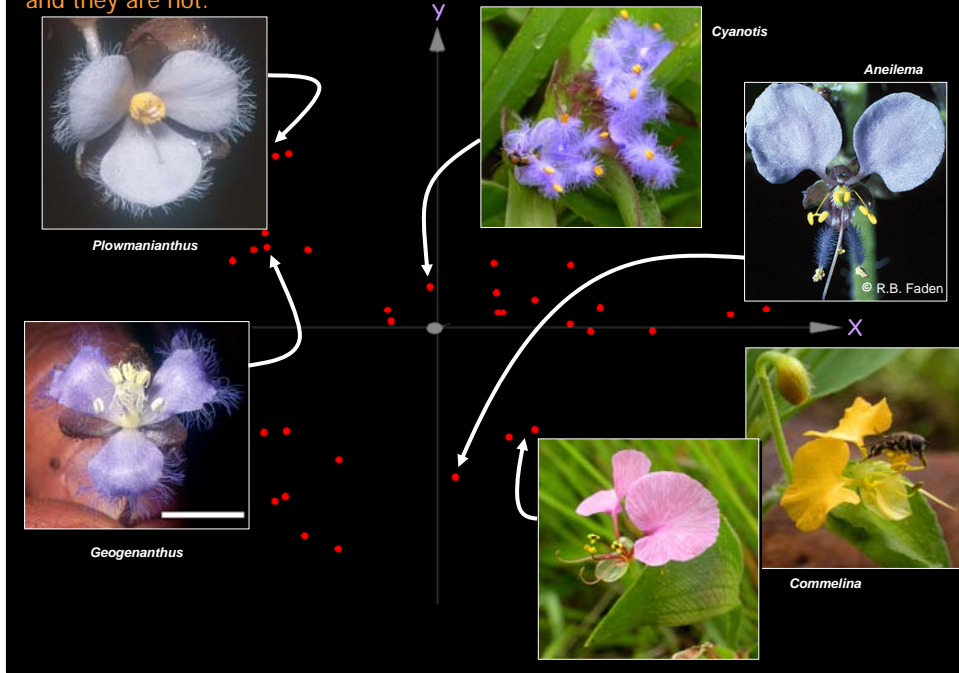


synorganization

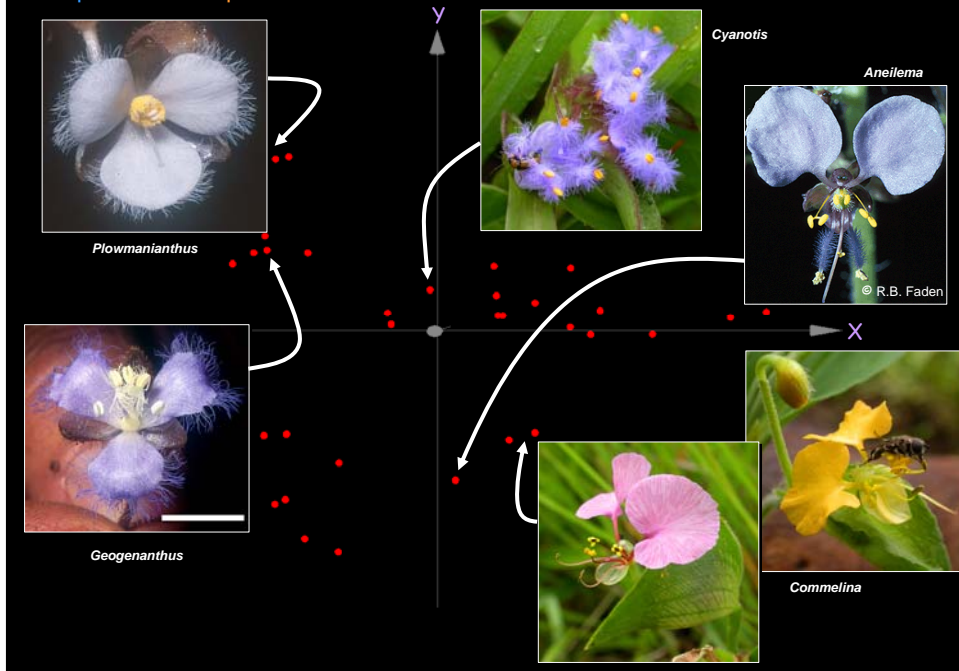
the intimate (spatial and functional) connection of organs of the same or different types to form a functional apparatus.

Endress (1994).

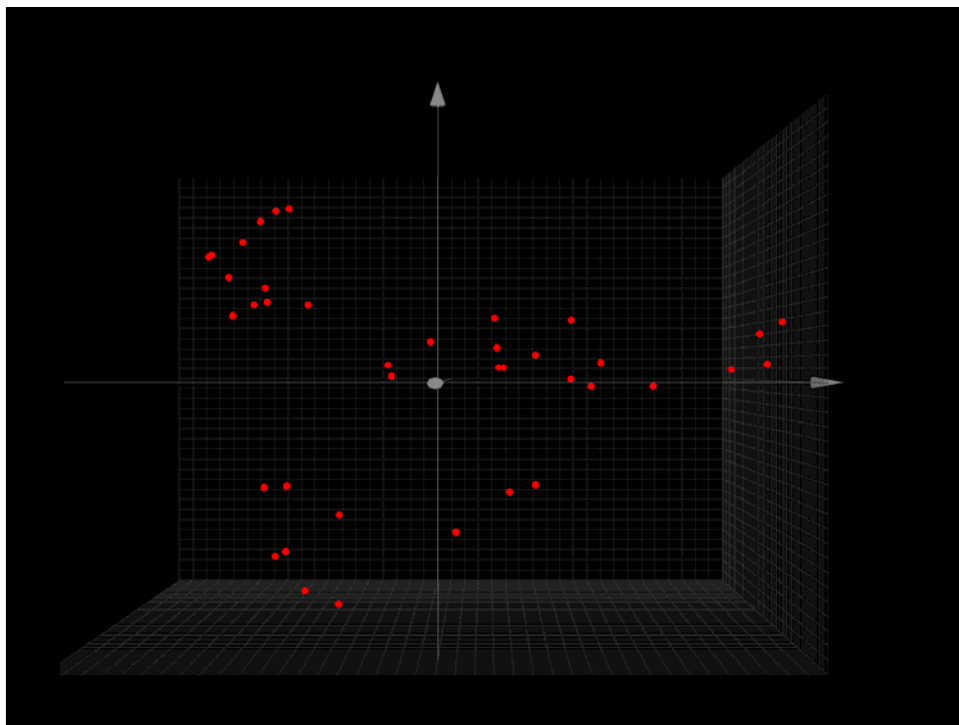
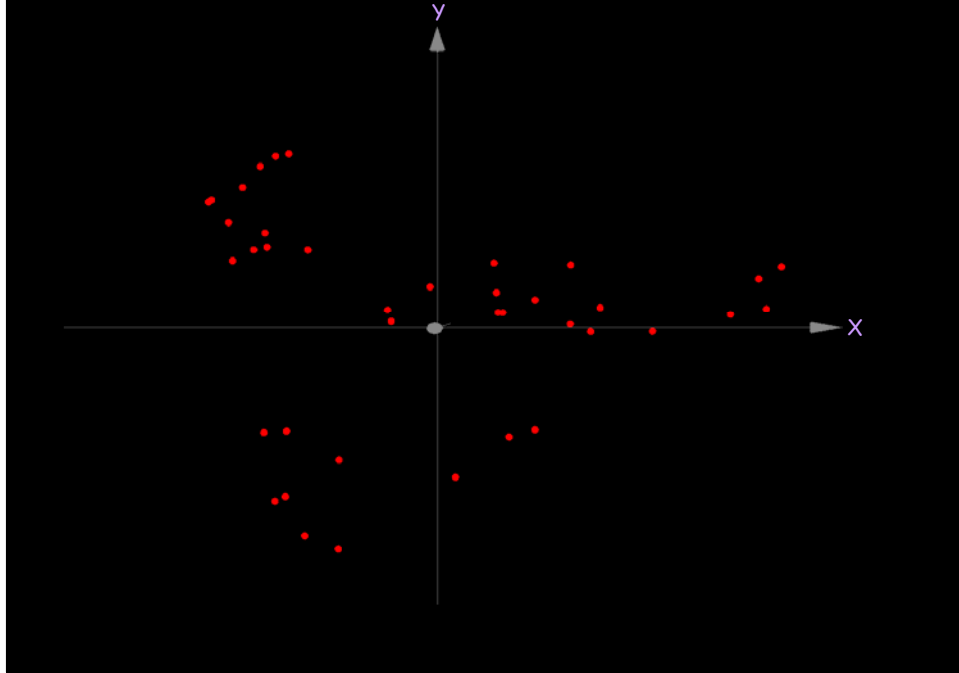
This implies that floral forms will not be uniformly distributed in morphospace: and they are not.

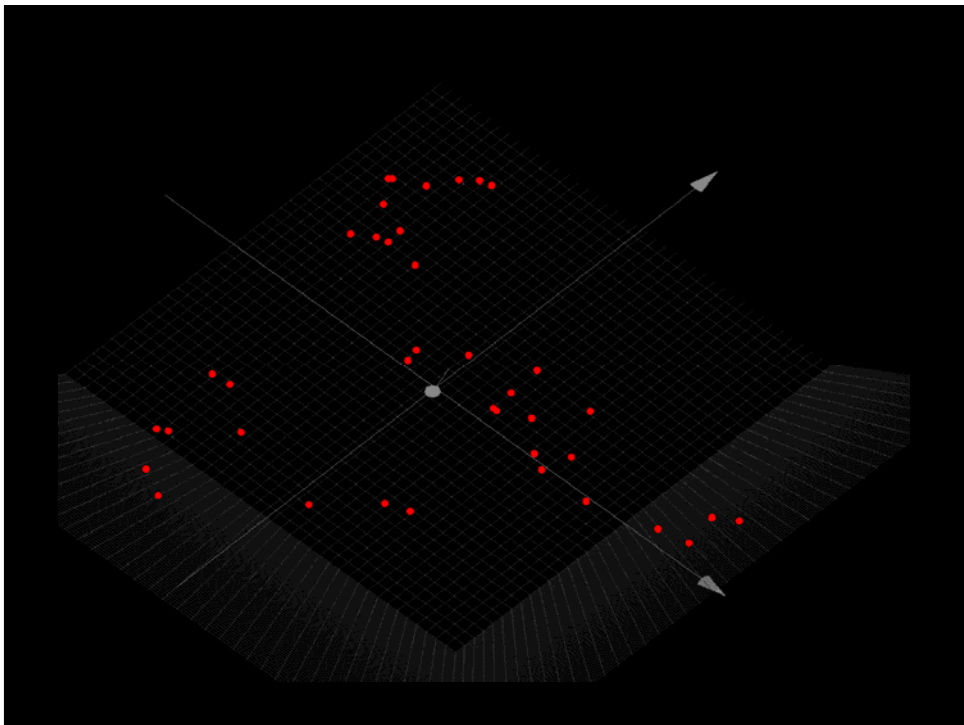
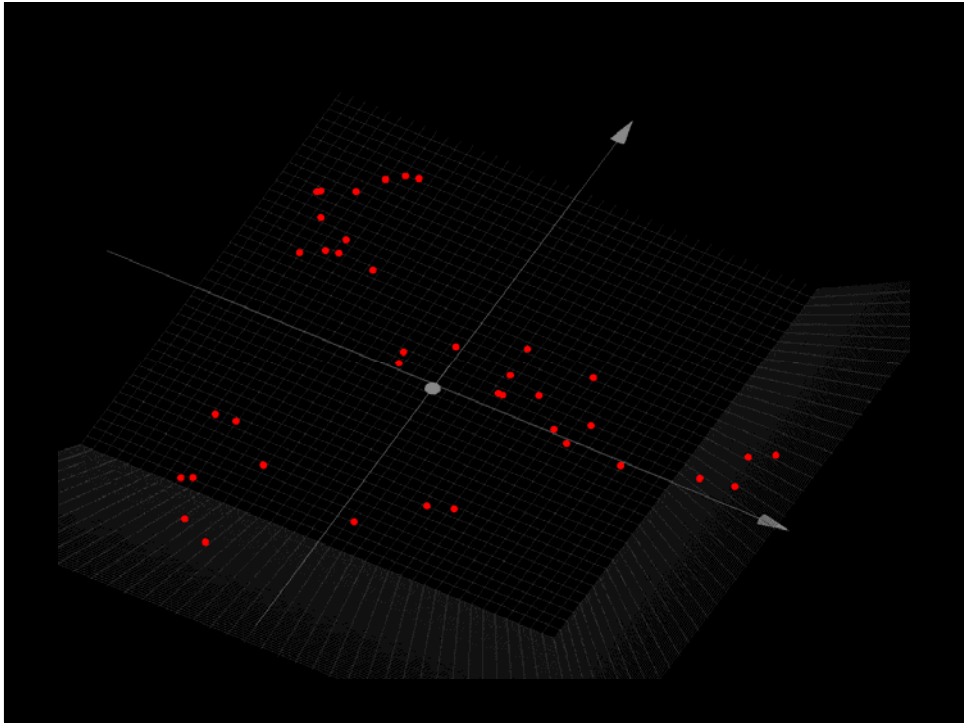


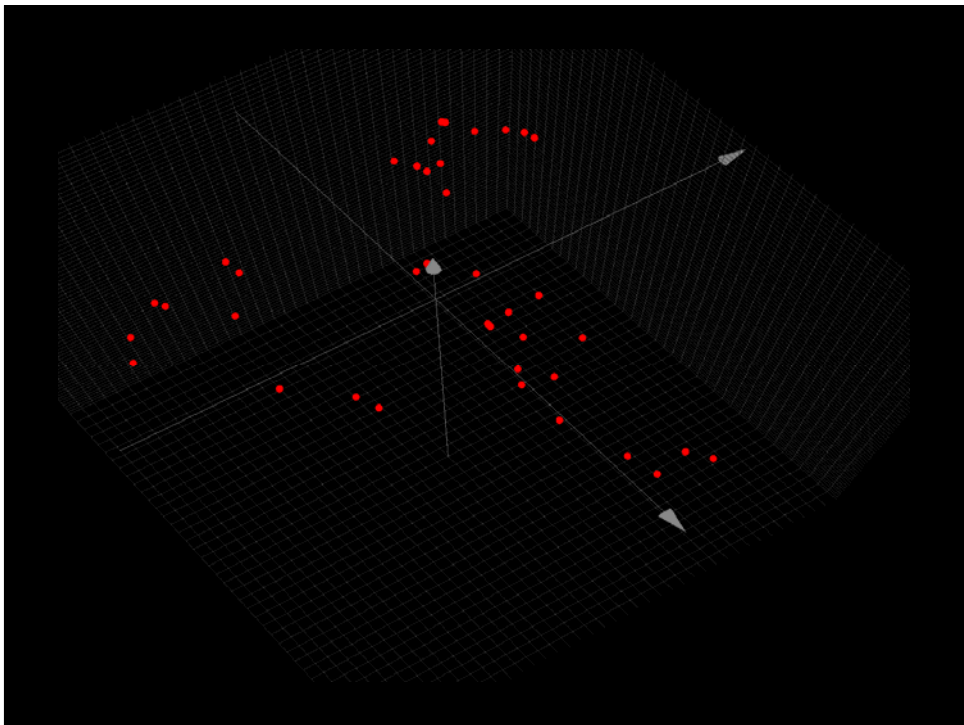
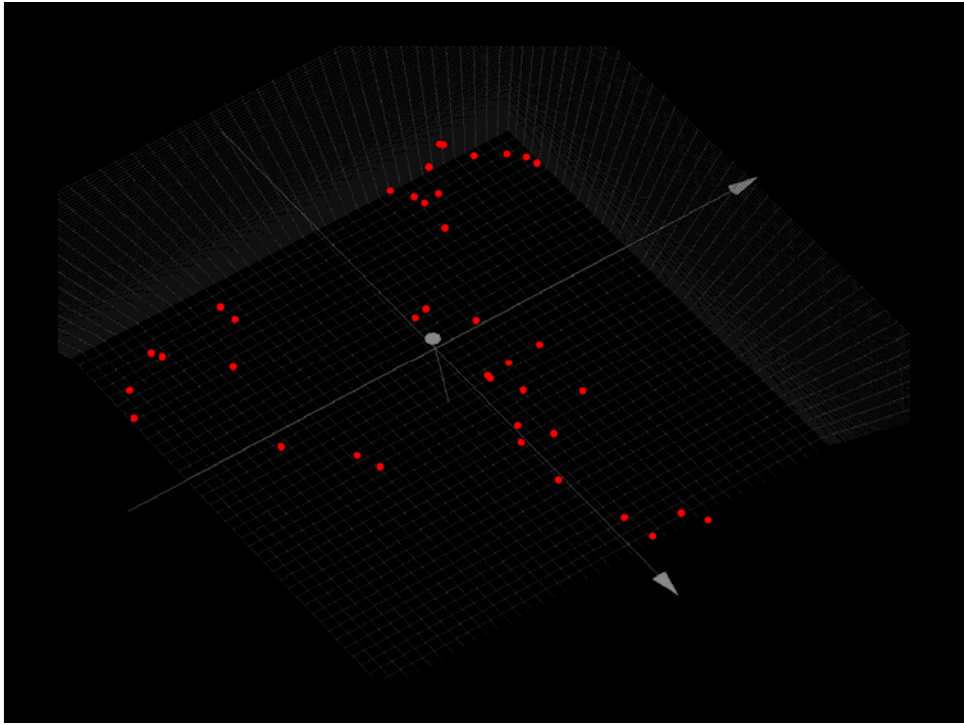
This pattern is the product of evolution via natural selection.

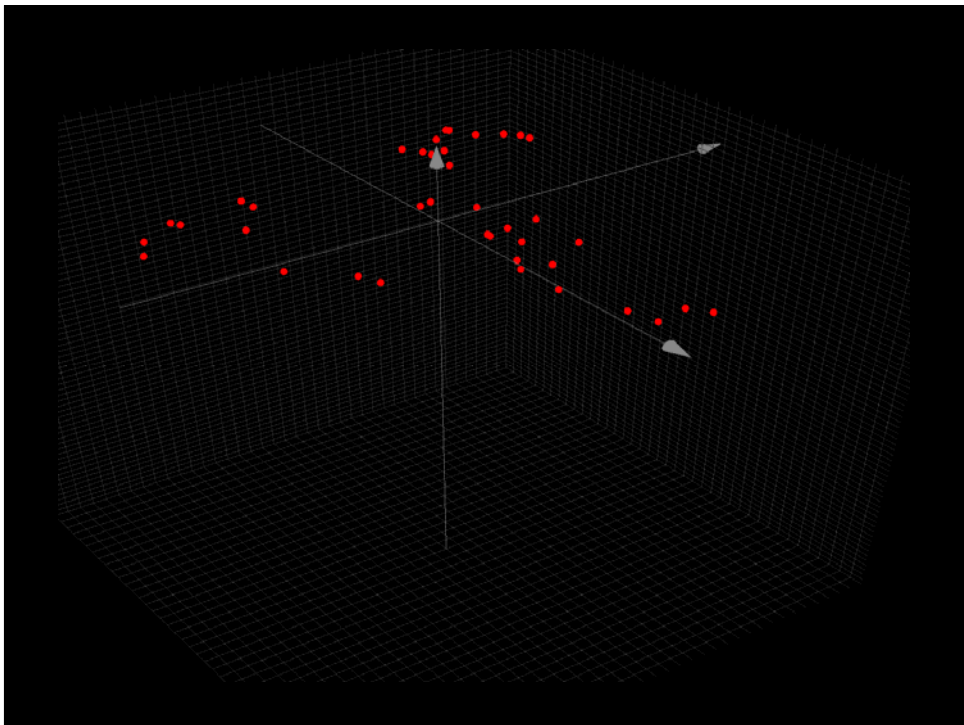
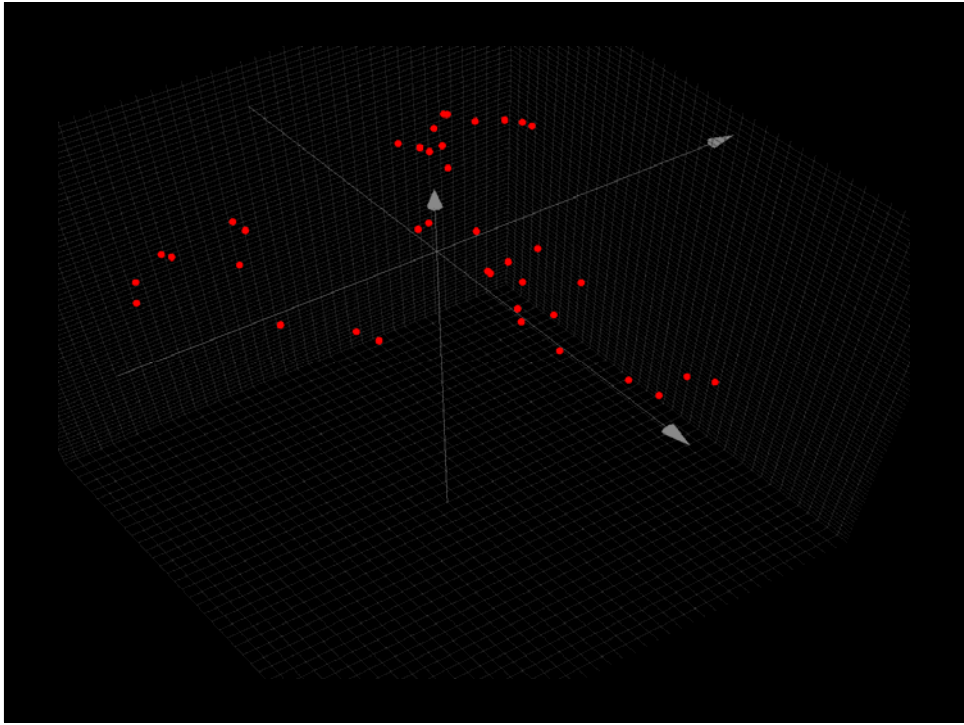


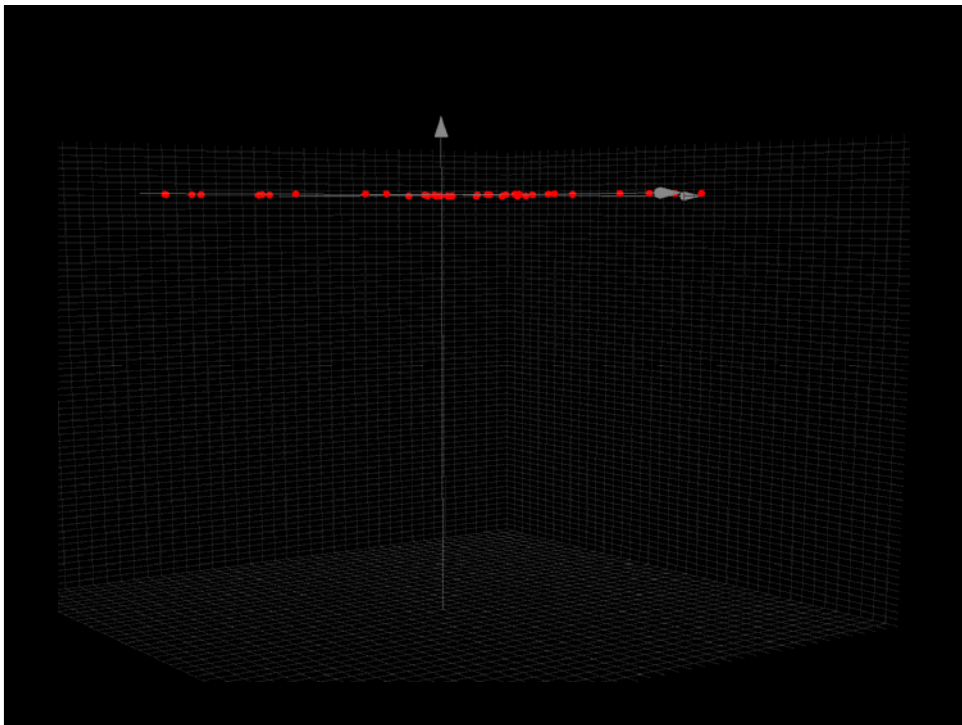
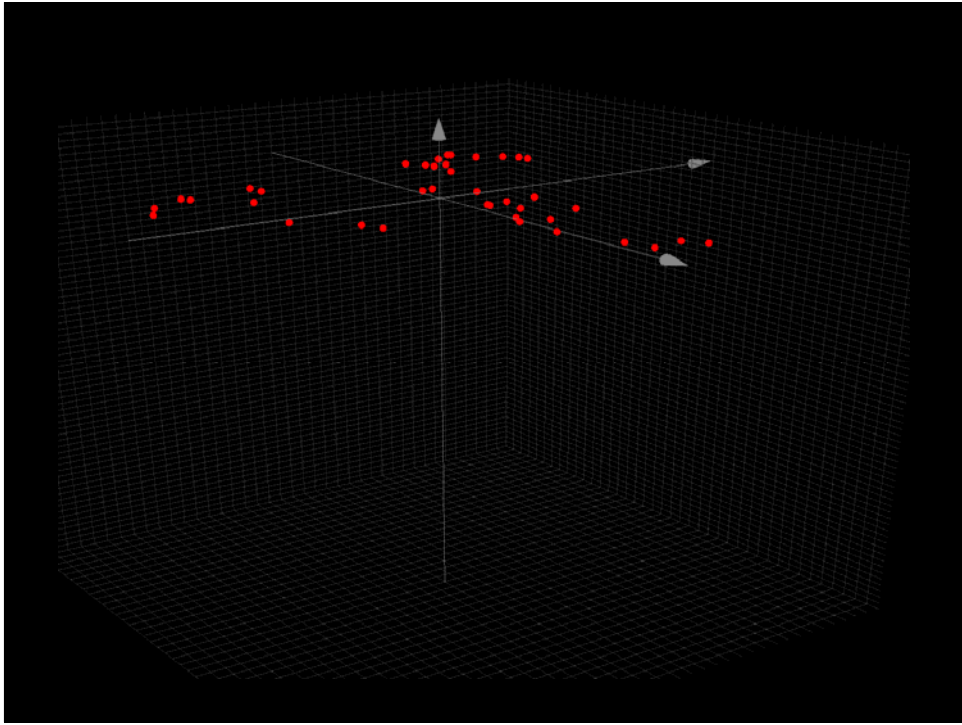
Thus, floral evolution is a [process](#) shaped by natural selection.

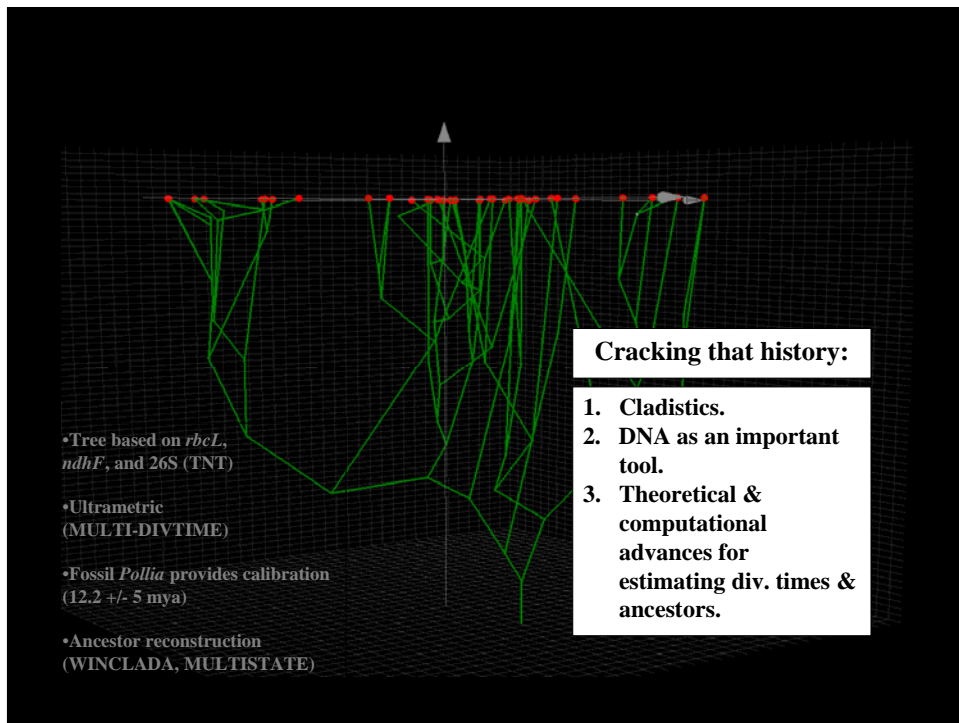
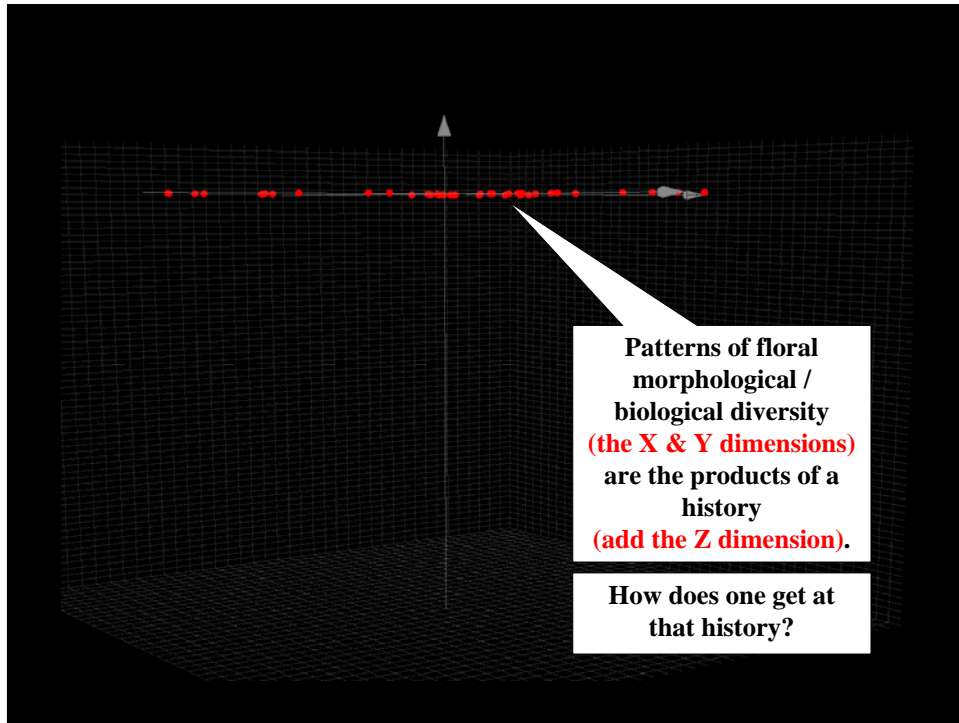




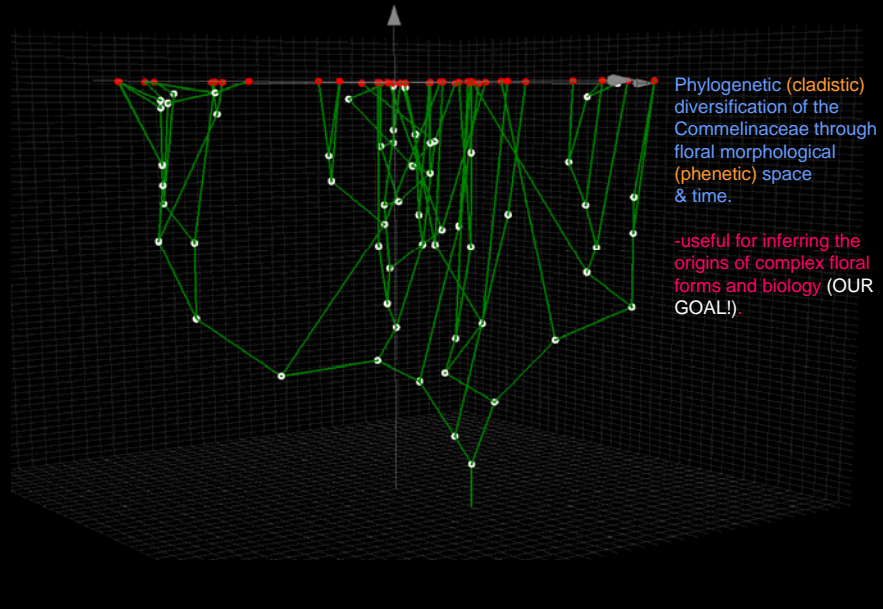






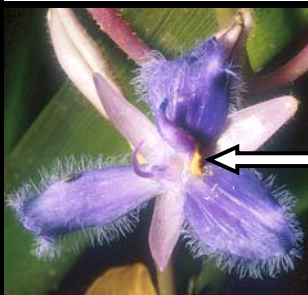
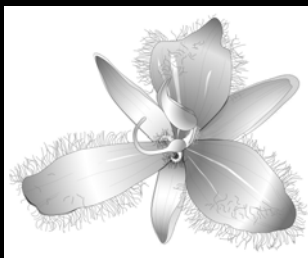


Reuniting the cladist with its not-so-distant cousin the pheneticist!

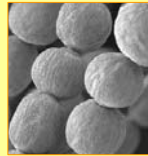


e.g., floral complexity and buzz pollination in *Cochlostema* (Commelinaceae)

- Epiphyte from NW South America lowland rainforests.
- Large, nectarless, fragrant, asymmetric flowers attract Xylocopine and Euglossine bees.
- 3 connate fertile stamens.
- 3 concealed, spirally coiled anthers.



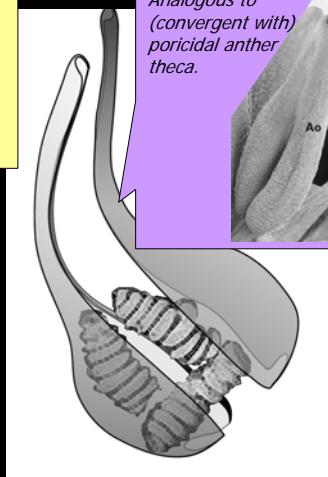
Stamen hairs mimic pollen mass.



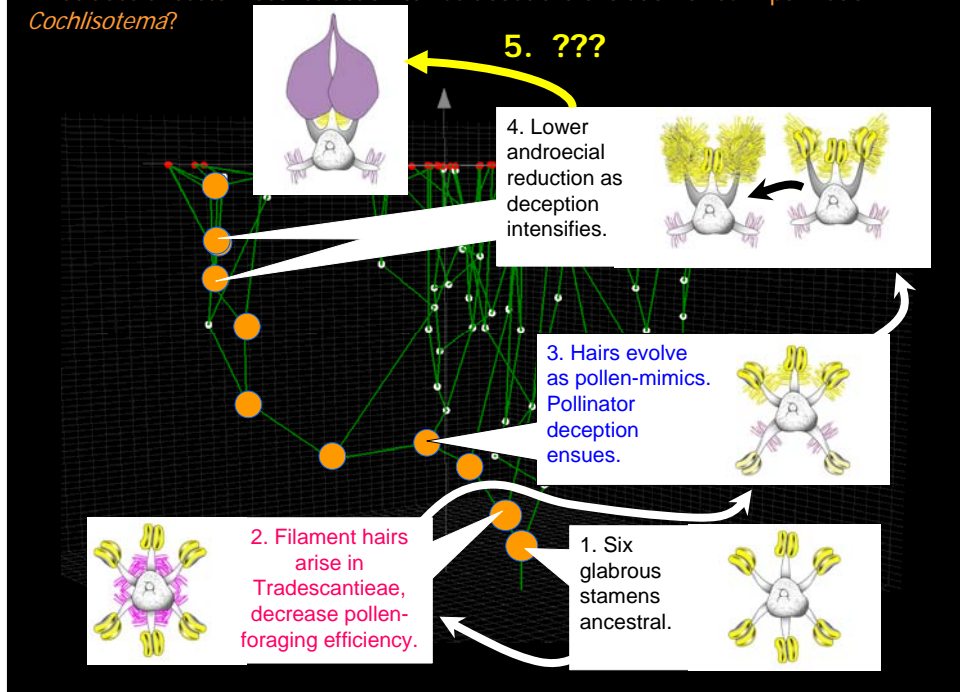
1 cm

Filament extensions conceal & conserve pollen.

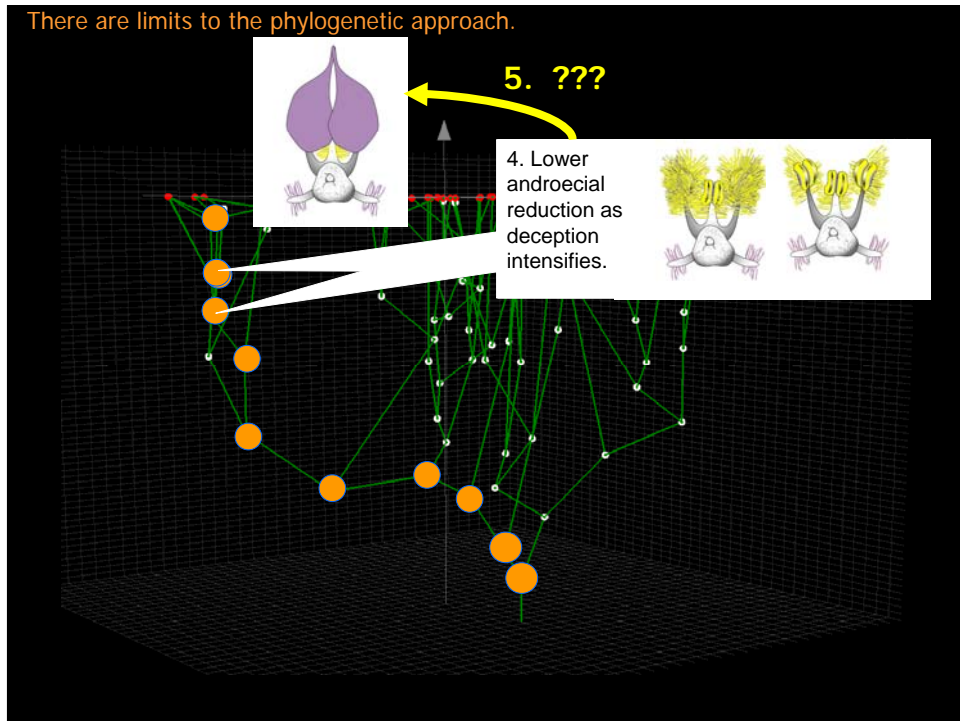
Analogous to (convergent with) poricidal anther theca.



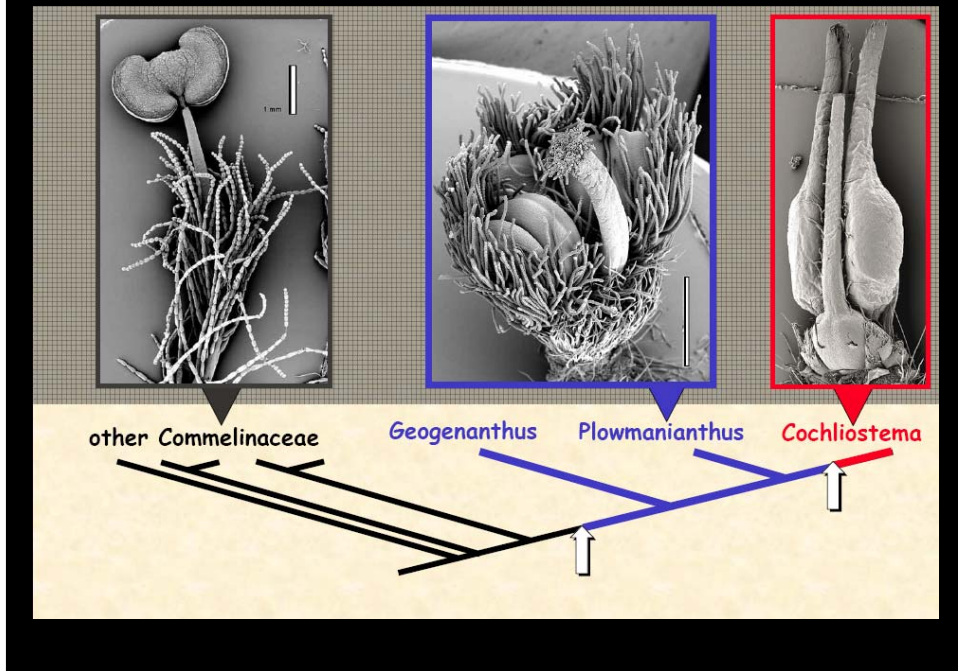
What does ancestor reconstruction tell us about the evolution of buzz pollination in *Cochlisotema*?



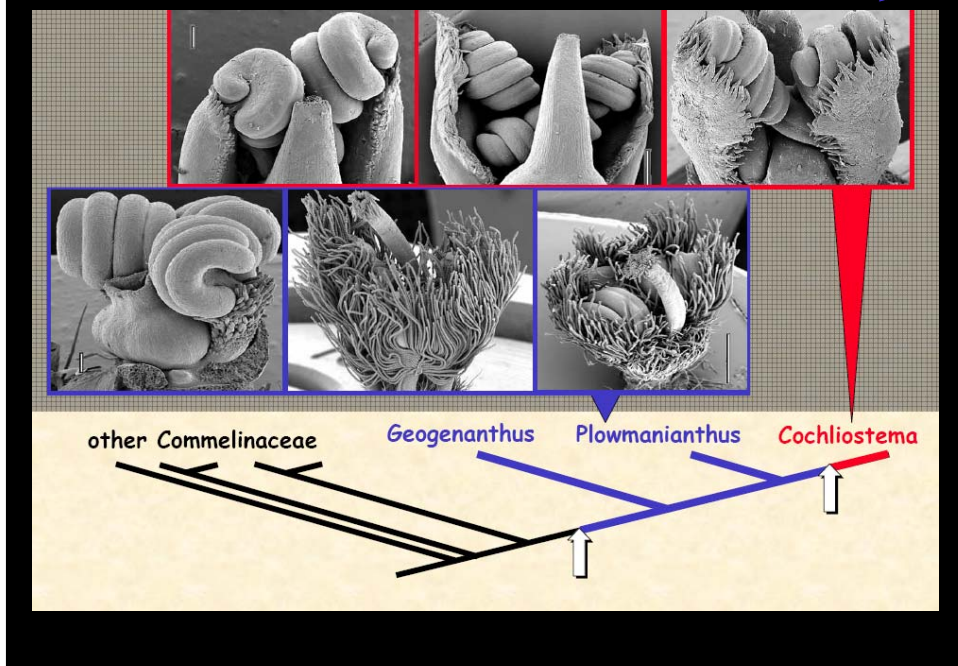
There are limits to the phylogenetic approach.

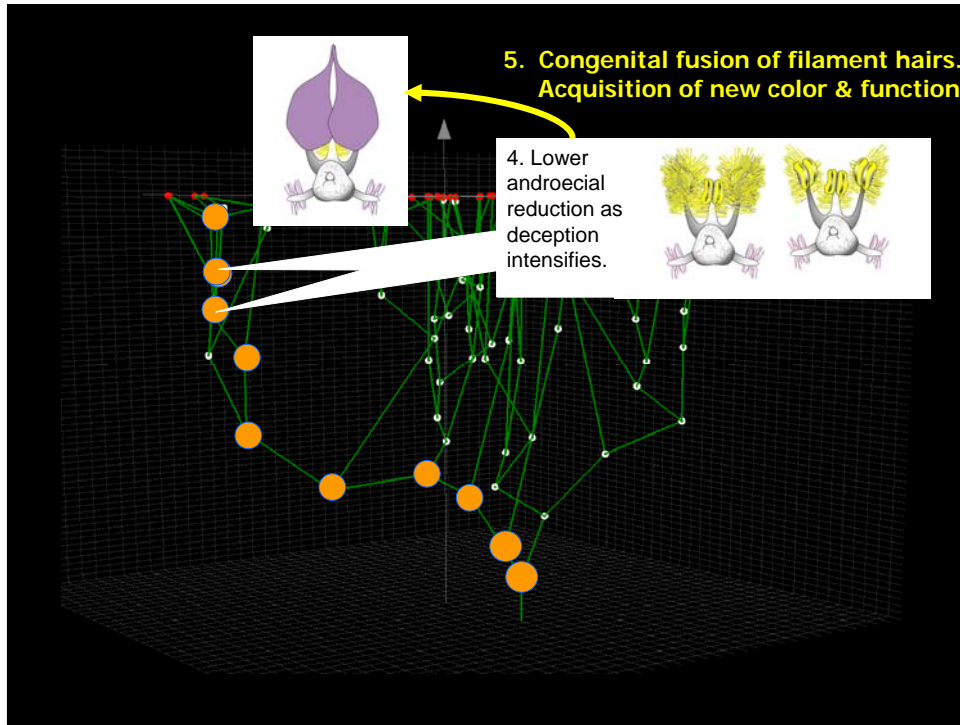


Stamen morphology in the phylogenetic vicinity of *Cochliostema*.



Stamen developmental morphology in the phylogenetic vicinity of *Cochliostema*.

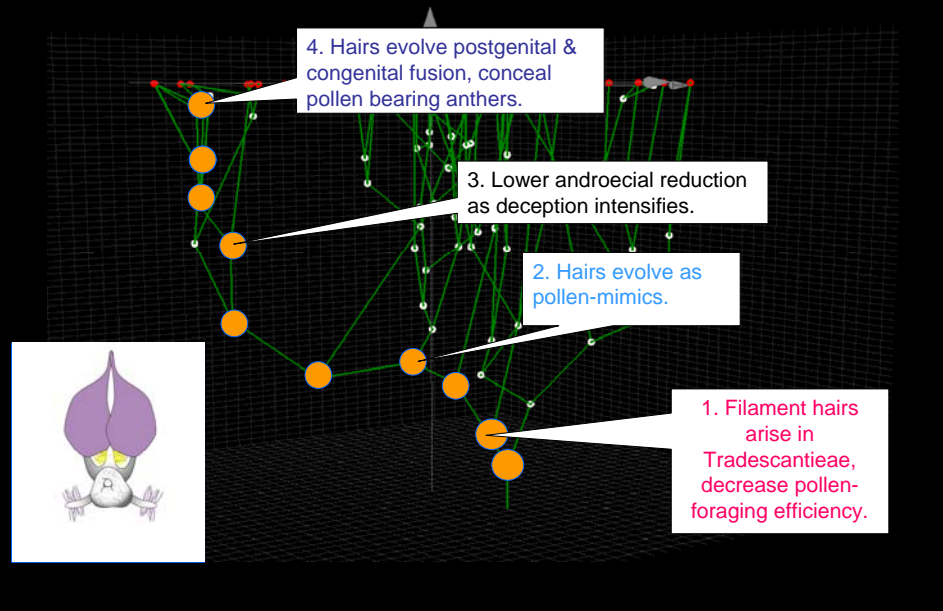




Important side note:

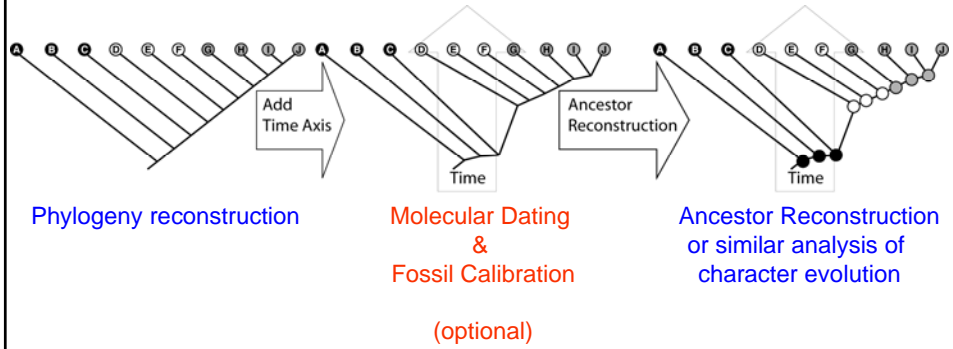
There are limits to the phylogenetic approach & ancestor reconstruction.
Phylogenetics and developmental studies can be reciprocally illuminating.

A fascinating story of adaptation & exaptation in *Cochlostema*.



2. Assumptions of Parsimony and Model-Based Methods
-Emphasis on Branch Lengths.
-Programs for Ancestor Reconstruction.

Ancestor Reconstruction: The basic process.



31

The first question: Should we choose Parsimony OR Model-based Methods?

32

How about: Parsimony AND Model-based Methods

33

How about: Parsimony AND Model-based Methods

-**Politically speaking:** A pluralistic approach gives you more journal options
(e.g., Syst Biol vs. Cladistics).

34

How about: Parsimony AND Model-based Methods

- Politically speaking: A pluralistic approach gives you more journal options (e.g., Syst Biol vs. Cladistics).
- Cost-benefit analysis:** Parsimony is quick, easy, costs nothing to do. Widely available in user-friendly packages.

35

How about: Parsimony AND Model-based Methods

- Politically speaking: A pluralistic approach gives you more journal options (e.g., Syst Biol vs. Cladistics).
- Cost-benefit analysis: Parsimony is quick, easy, costs nothing to do. Widely available in user-friendly packages.
- Common familiarity:** Most evolutionary biologists have and still do use parsimony.

36

How about: Parsimony AND Model-based Methods

- Politically speaking: A pluralistic approach gives you more journal options (e.g., Syst Biol vs. Cladistics).
- Cost-benefit analysis: Parsimony is quick, easy, costs nothing to do. Widely available in user-friendly packages.
- Common familiarity: Most evolutionary biologists have and still do use parsimony.
- Expanded Toolkit:** Model-based methods are powerful & flexible tools to explore data and to estimate various parameters relating to the evolution of flowers.

37

How about: Parsimony AND Model-based Methods

- Politically speaking: A pluralistic approach gives you more journal options (e.g., Syst Biol vs. Cladistics).
- Cost-benefit analysis: Parsimony is quick, easy, costs nothing to do. Widely available in user-friendly packages.
- Common familiarity: Most evolutionary biologists have and still do use parsimony.
- Expanded Toolkit: Model-based methods are powerful & flexible tools to explore data and to estimate various parameters relating to the evolution of flowers.
- Paul O. Lewis (UConn, 2007):** In reality, there are too many unknowns to be able to argue persuasively either for or against allowing branch length heterogeneity.

This much, at least, I think we can safely conclude:

- likelihood models offer a **second opinion** that can be compared to parsimony
- development of likelihood models for morphology stimulates valuable discussion about character evolution, which will ultimately provide better models for morphology

38

FYI: A list of widely available programs.

Table 1 Some programs available for ancestor reconstruction and the basic methods they implement. Abbreviations are as follows: ML, maximum likelihood; GLS, generalized least squares; binary, binary characters; multistate, multistate characters

	Discrete characters		Continuous characters	
	Parsimony	ML	Parsimony	ML/GLS
PAUP* 4.0 (Swofford, 2002)	Fitch & Wagner	Yes (binary and multistate)	Linear (Wagner) & squared-change	Yes
MACCLADE 4.0 (Maddison & Maddison, 1992)	Fitch & Wagner	—	Linear (Wagner) & squared-change	—
WINCLADA 1.0 (Nixon, 2002)	Fitch & Wagner	—	—	—
MESQUITE 1.05 (Maddison & Maddison, 2003)	Fitch & Wagner	Yes (binary and multistate)	Linear (Wagner) & squared-change	—
TNT 1.0 (Goloboff <i>et al.</i> , 2004)	Fitch & Wagner	—	Linear (Wagner)	—
BAYESMULTISTATE and DISCRETE (Pagel <i>et al.</i> , 2004)	—	Yes (binary and multistate)	—	—
DISCRETE (Pagel, 1997)	—	Yes (binary)	—	—
MULTISTATE (Pagel, 1997)	—	Yes (binary and multistate)	—	—
ANCML (Schluter, 1997)	—	—	—	Yes
CONTINUOUS* (Pagel, 1997)	—	Yes (binary)	—	Yes
COMPARE 4.6 (Martins, 2004)	—	—	—	Yes
SIMMAP 1.0 (Bollback, 2006)	—	Yes (binary and multistate)	—	Yes
BAYESTRAITS 1.0 (Pagel & Meade, 2004 onwards)	—	Yes (binary and multistate)	—	Yes

*CONTINUOUS actually only provides reconstructions for the root node: the program is really only designed for testing for correlated evolution between traits, adaptive evolution and differential rates or trends in character evolution.

Diversity and Distributions, (Diversity Distrib.) (2006) 12, 7–19

FYI: A list of widely available programs.

User-friendly GUIs

Table 1 Some programs available for ancestor reconstruction and the basic methods they implement. Abbreviations are as follows: ML, maximum likelihood; GLS, generalized least squares; binary, binary characters; multistate, multistate characters

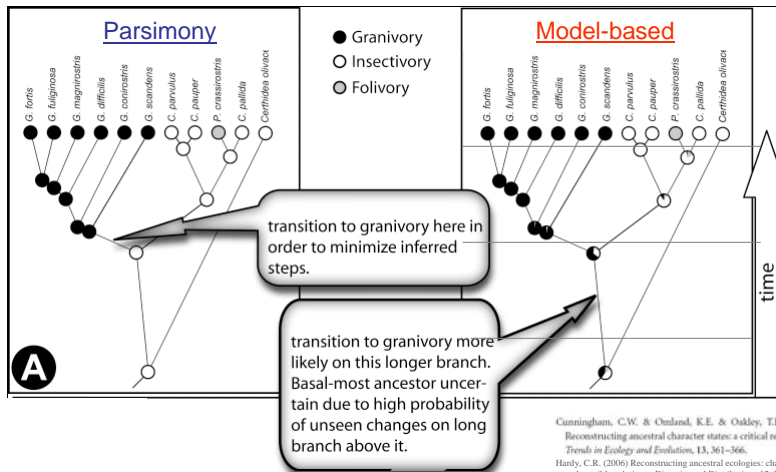
	Discrete characters		Continuous characters	
	Parsimony	ML	Parsimony	ML/GLS
PAUP* 4.0 (Swofford, 2002)	Fitch & Wagner	Yes (binary and multistate)	Linear (Wagner) & squared-change	Yes
MACCLADE 4.0 (Maddison & Maddison, 1992)	Fitch & Wagner	—	Linear (Wagner) & squared-change	—
WINCLADA 1.0 (Nixon, 2002)	Fitch & Wagner	—	—	—
MESQUITE 1.05 (Maddison & Maddison, 2003)	Fitch & Wagner	Yes (binary and multistate)	Linear (Wagner) & squared-change	—
TNT 1.0 (Goloboff <i>et al.</i> , 2004)	Fitch & Wagner	—	Linear (Wagner)	—
BAYESMULTISTATE and DISCRETE (Pagel <i>et al.</i> , 2004)	—	Yes (binary and multistate)	—	—
DISCRETE (Pagel, 1997)	—	Yes (binary)	—	—
MULTISTATE (Pagel, 1997)	—	Yes (binary and multistate)	—	—
ANCML (Schluter, 1997)	—	—	—	Yes
CONTINUOUS* (Pagel, 1997)	—	Yes (binary)	—	Yes
COMPARE 4.6 (Martins, 2004)	—	—	—	Yes
SIMMAP 1.0 (Bollback, 2006)	—	Yes (binary and multistate)	—	Yes
BAYESTRAITS 1.0 (Pagel & Meade, 2004 onwards)	—	Yes (binary and multistate)	—	Yes

*CONTINUOUS actually only provides reconstructions for the root node: the program is really only designed for testing for correlated evolution between traits, adaptive evolution and differential rates or trends in character evolution.

Diversity and Distributions, (Diversity Distrib.) (2006) 12, 7–19

What are your assumptions when you choose one method over the other?

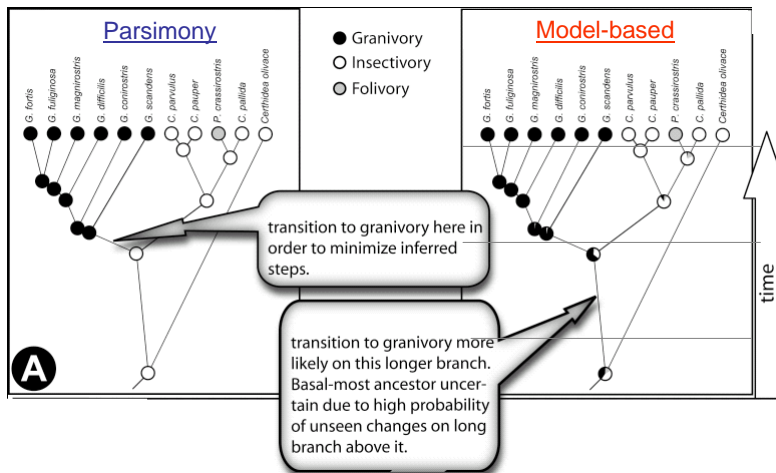
Where different reconstructions exist, these are due to branch lengths.



What are your assumptions when you choose one method over the other?

Implicit Assumption
 P(change) equal on all branches (i.e., branch lengths, if had, are not considered).

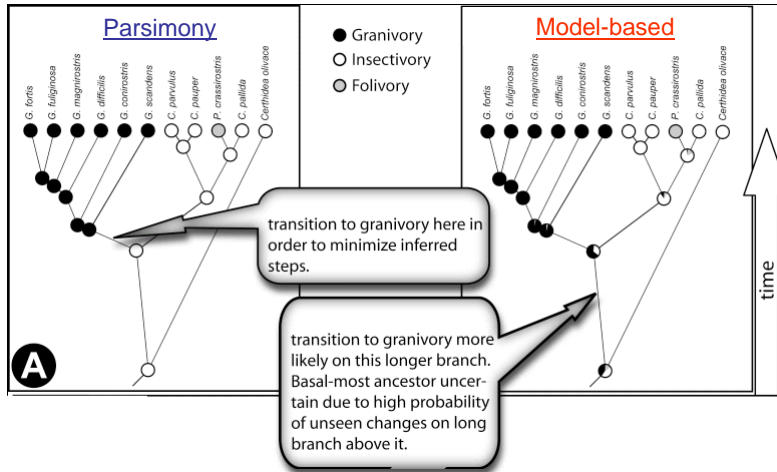
Explicit Assumption
 P(change) ~ branch length



What are your assumptions when you choose one method over the other?

Implicit Assumption
Punctuational Equilibrium?

Explicit Assumption
Gradual Evolution?



43

What are your assumptions when you choose one method over the other?

Implicit Assumption
Punctuational Equilibrium?

Explicit Assumption
Gradual Evolution?

Does this mean that models are inappropriate for ancestor reconstruction in adaptive radiations?
Does that mean that such methods should not be used in cases where adaptation is thought to have occurred (i.e., in flower evolution)?

Adaptive radiation: family Proteaceae from the Cape region in South Africa



<http://hydrodictyon.eeb.uconn.edu/people/silander/>

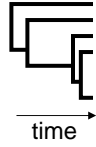


Proteaceae example borrowed from Paul Lewis.

44

Types of branch lengths.

Ultrametric branch lengths (e.g., following a molecular clock, Penalized Likelihood or MultiDivtime analysis):

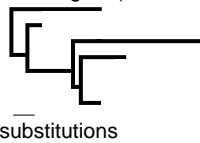


Default assumption in most programs: probability of change is proportional to time (equilibria and punctuated change more difficult to detect).

45

Types of branch lengths.

Raw genetic branch lengths (as seen on a "phylogram"):

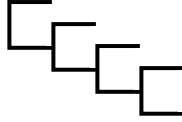


Default assumption in most programs: degree of character evolution is proportional to degree of molecular evolution in the gene(s) analyzed. (punctuated morphological change during adaptive radiations difficult to detect).

46

Types of branch lengths.

Equal branch lengths (e.g., all branches are assigned the same length):



Default assumption in most programs:

Change equiprobable on all branches. Parsimony-like.

(Punctuated morphological change during adaptive radiations potentially easier to detect).

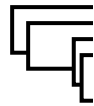
47

One of BayesTraits' best kept secret: the kappa (κ) parameter.



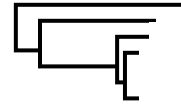
$\kappa < 1.0$
(branch lengths
shortened, degree in proportion
to their original length)

Starting tree



time →

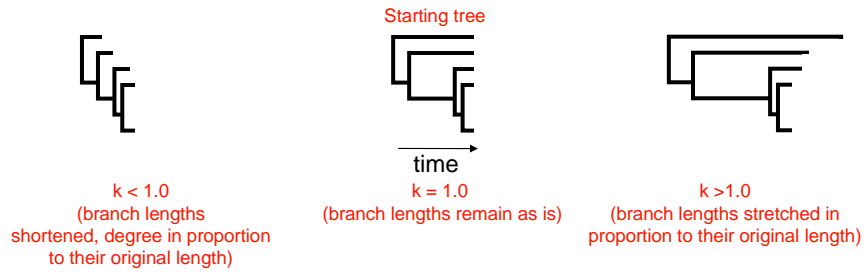
$\kappa = 1.0$
(branch lengths remain as is)



$\kappa > 1.0$
(branch lengths stretched in
proportion to their original length)

48

One of BayesTraits' best kept secret: the kappa (k) parameter.

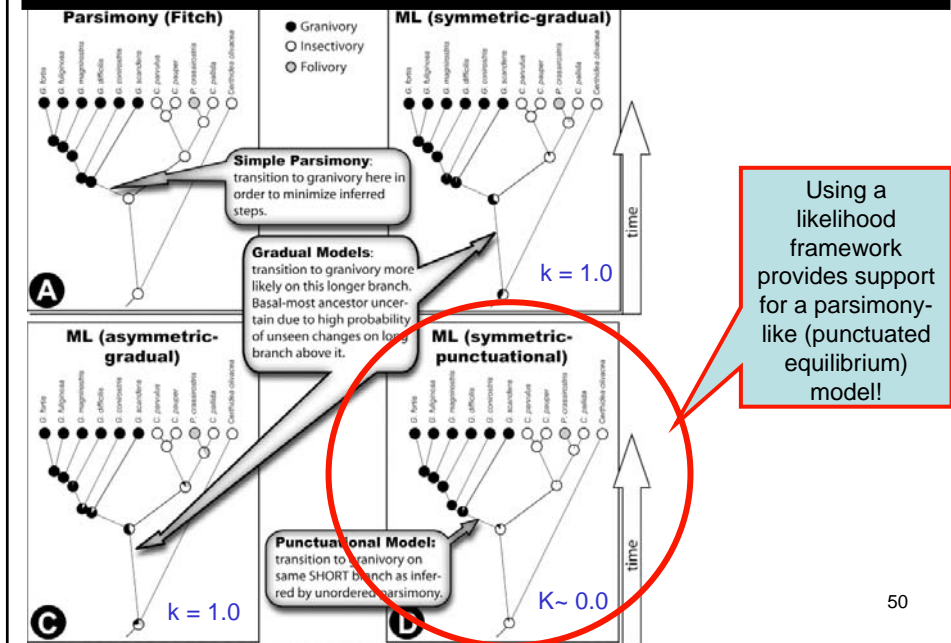


How to use it (my suggestions):

1. LR Test: Does including k increase likelihood relative to not including it. No k is better than a poorly estimated k.
2. If it's appropriate, is the ML (or near ML) estimate of its value closer to 0 than to 1? This gives an indication of rates of character evolution.
3. If not appropriate to include k, check for whether punctuated or gradual model is better: Compare likelihoods of tree and reconstruction with k fixed at 1.0 vs. fixed at 0.0.

49

One of BayesTraits' best kept secret: the kappa (k) parameter.



50

4. Accounting for Uncertainty.

Accounting for Uncertainty.

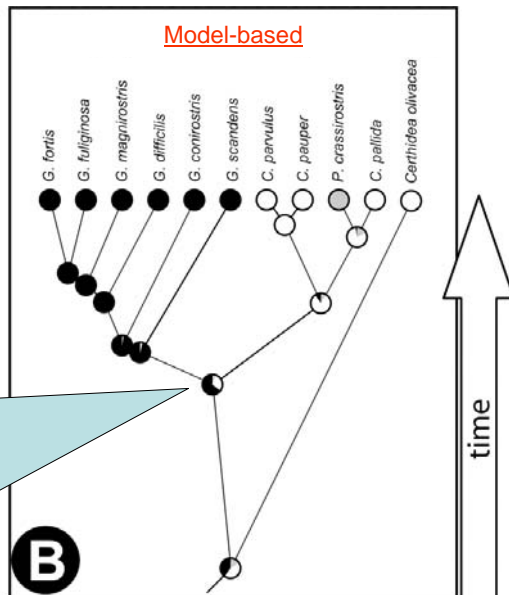
- Granivory
- Insectivory
- Folivory

Model-based methods are explicitly probabilistic:

-the probability of states for any particular node is automatically estimated.

-pie diagrams represent good graphical way of depicting proportional or relative likelihoods.

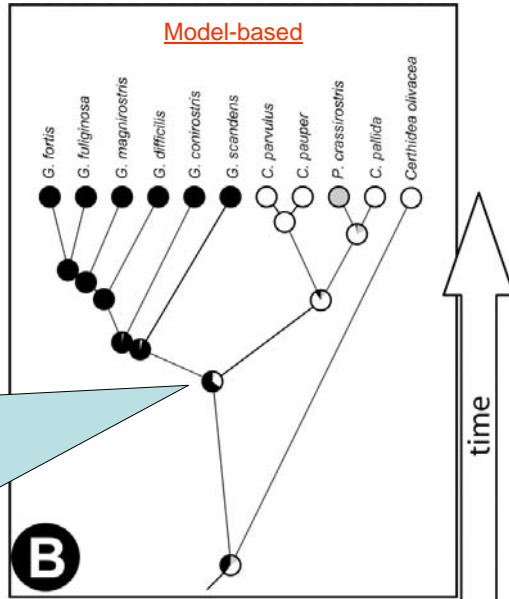
-these are valid as long as the model is valid.



Accounting for Uncertainty.

- Granivory
- Insectivory
- Folivory

-these are valid as long as the model is valid.



B

Accounting for Uncertainty: **Asymmetric model is bad here.**

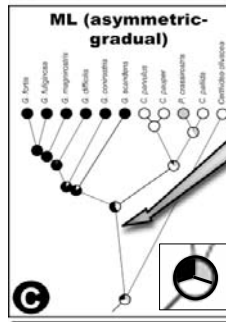
Adding separate parameters for forward and reverse rates (something that sounds logical) is actually not valid here.

- Granivory
- Insectivory
- Folivory

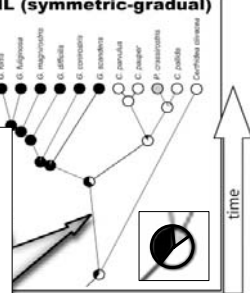
Over-parameterization of model may lead to poor uncertainty estimates
e.g.,

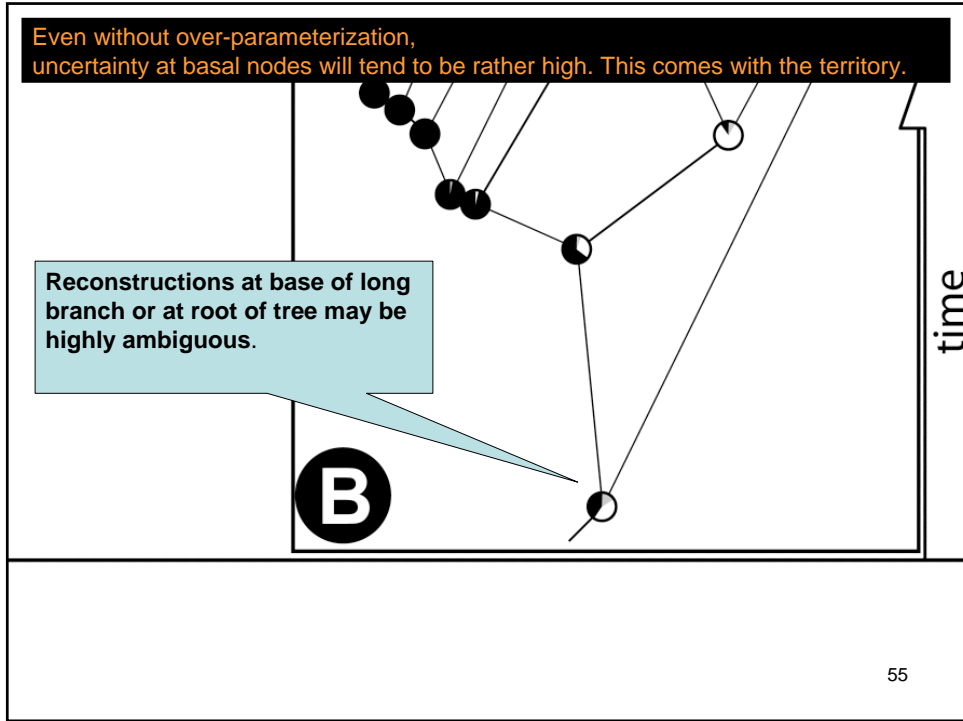
1 parameter
(forward-reverse rates the same)

vs.
Multi-parameter
(forward & reverse rates estimated independently)



ML (symmetric-gradual)





Just a few programs account for uncertainty as fully as it should be.

Two sources of uncertainty:

1. [Reconstruction uncertainty at particular nodes in any particular tree.](#)
2. [Tree uncertainty \(best done within Bayesian framework\).](#)

Table 1 Some programs available for ancestor reconstruction and the basic methods they implement. Abbreviations are as follows: ML, maximum likelihood; GLS, generalized least squares; binary, binary characters; multistate, multistate characters

	Discrete characters		Continuous characters	
	Parsimony	ML	Parsimony	ML/GLS
PAUP* 4.0 (Swofford, 2002)	Fitch & Wagner	Yes (binary and multistate)	Linear (Wagner) & squared-change	Yes
MACCLADE 4.0 (Maddison & Maddison, 1992)	Fitch & Wagner	—	Linear (Wagner) & squared-change	—
WINCLADA 1.0 (Nixon, 2002)	Fitch & Wagner	—	—	—
MESQUITE 1.05 (Maddison & Maddison, 2003)	Fitch & Wagner	Yes (binary and multistate)	Linear (Wagner) & squared-change	—
TNT 1.0 (Goloboff <i>et al.</i> , 2004)	Fitch & Wagner	—	Linear (Wagner)	—
BAYESMULTISTATE and DISCRETE (Pagel <i>et al.</i> , 2004)	—	Yes (binary and multistate)	—	—
DISCRETE (Pagel, 1997)	—	Yes (binary)	—	—
MULTISTATE (Pagel, 1997)	—	Yes (binary and multistate)	—	—
ANCML (Schluter, 1997)	—	—	—	Yes
CONTINUOUS* (Pagel, 1997)	—	Yes (binary)	—	Yes
COMPARE 4.6 (Martins, 2004) ?	—	—	—	Yes
SIMMAP 1.0 (Bollback, 2006)	—	Yes (binary and multistate)	—	Yes
BAYESTRAITS 1.0 (Pagel & Meade, 2004 onwards)	—	Yes (binary and multistate)	—	Yes

*CONTINUOUS actually only provides reconstructions for the root node: the program is really only designed for testing for correlated evolution between traits, adaptive evolution and differential rates or trends in character evolution.

Summary:

1. *Ancestor reconstruction one of many items in our toolkit.*
2. *There are many programs that allow reconstructions, not all the same.*
3. *Parsimony methods complement model-based methods, as a 'second opinion'.*
4. *We should exploit the recent theoretical and computational advances in modeling*
 - Factor in uncertainty.
 - Adequately explore and test assumptions about relationship between molecular-derived branch lengths and morphological character evolution.
 - Use parameter estimates to learn about character evolution.
5. *Someone should write a module for Mesquite that fully integrates multivariate approaches with ancestor reconstruction, correlation tests, etc. (and provide a slick visualization front-end to it).*

