



Dr. April Wright
4.11.2016

The Question(s)

- What is the ancestral state for squamate parity?
- How many times has the alternative state evolved from that?
- Have there been reversals back to the ancestral state?

The Set-Up

- Pyron and Burbrink 2014 (<http://onlinelibrary.wiley.com/doi/10.1111/ele.12168/full>)

The Set-Up

- Pyron and Burbrink 2014
- 8000 species classified by parity mode

The Set-Up

- Pyron and Burbrink 2014
- 8000 species classified by parity mode
- 4161 taxon tree

The Set-Up

- Pyron and Burbrink 2014
- 8000 species classified by parity mode
- 4161 taxon tree
 - Overlap of 4059 taxa

The Set-Up

- Pyron and Burbrink 2014
- 8000 species classified by parity mode
- 4161 taxon tree
 - Overlap of 4059 taxa
 - 12 species were bimodal

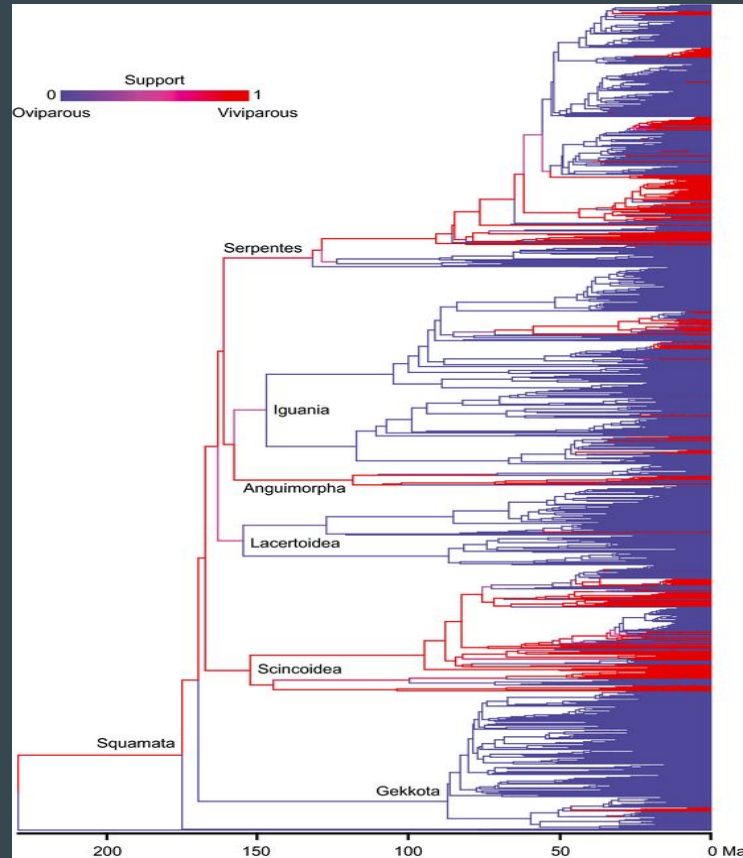
The Set-Up

- Fit a model of speciation and extinction to these data

The Set-Up

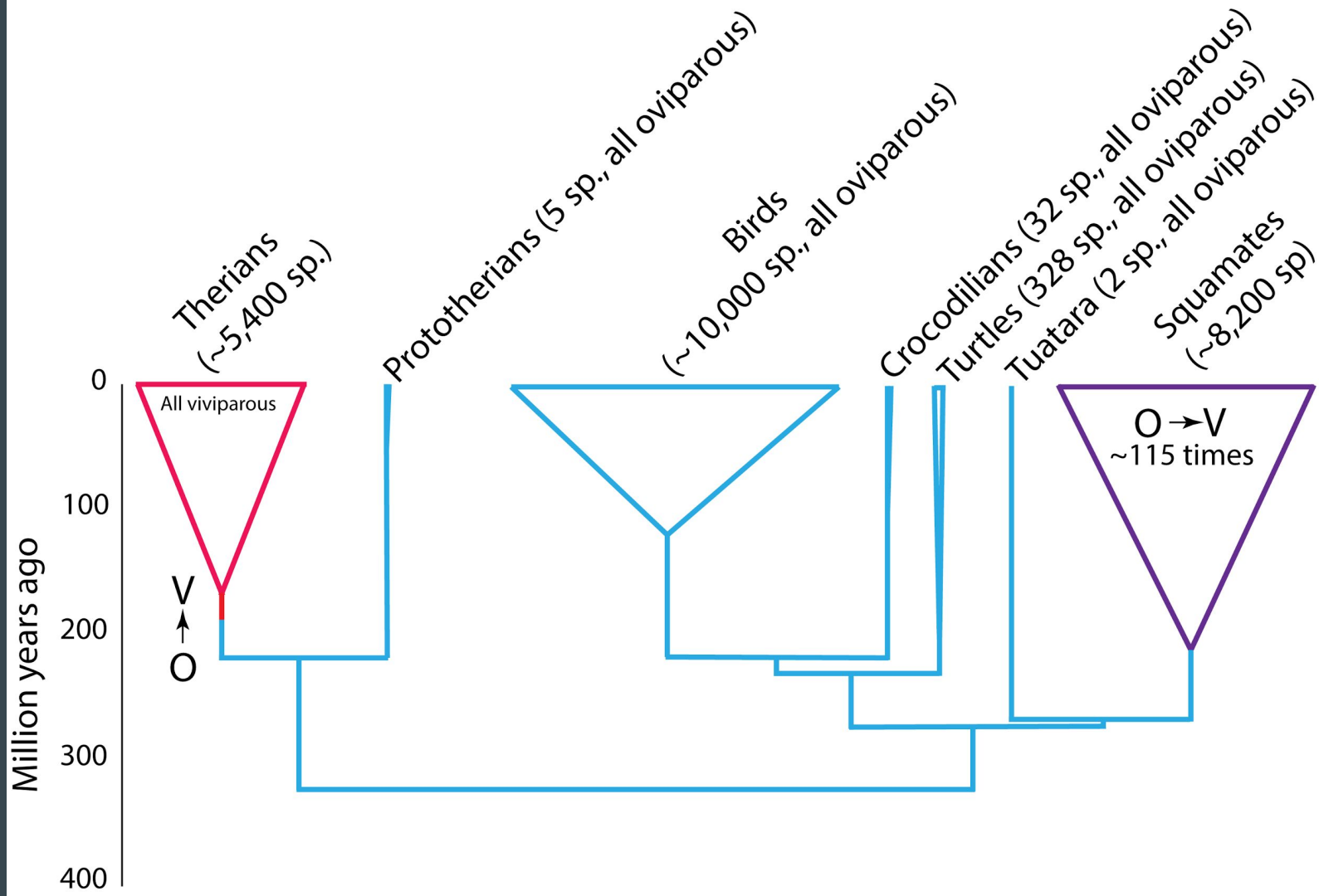
- Fit a model of speciation and extinction to these data
- This model, internally, estimates ancestral states

Early origin of viviparity and multiple reversions to oviparity in squamate reptiles



The Set-Up

- 34 origins of oviparity
- 56 reversals to oviparity



What we'll do today

- Large tree inference
- State-dependent diversification models

http://people.sc.fsu.edu/~dswofford/paup_test/

What do we observe?

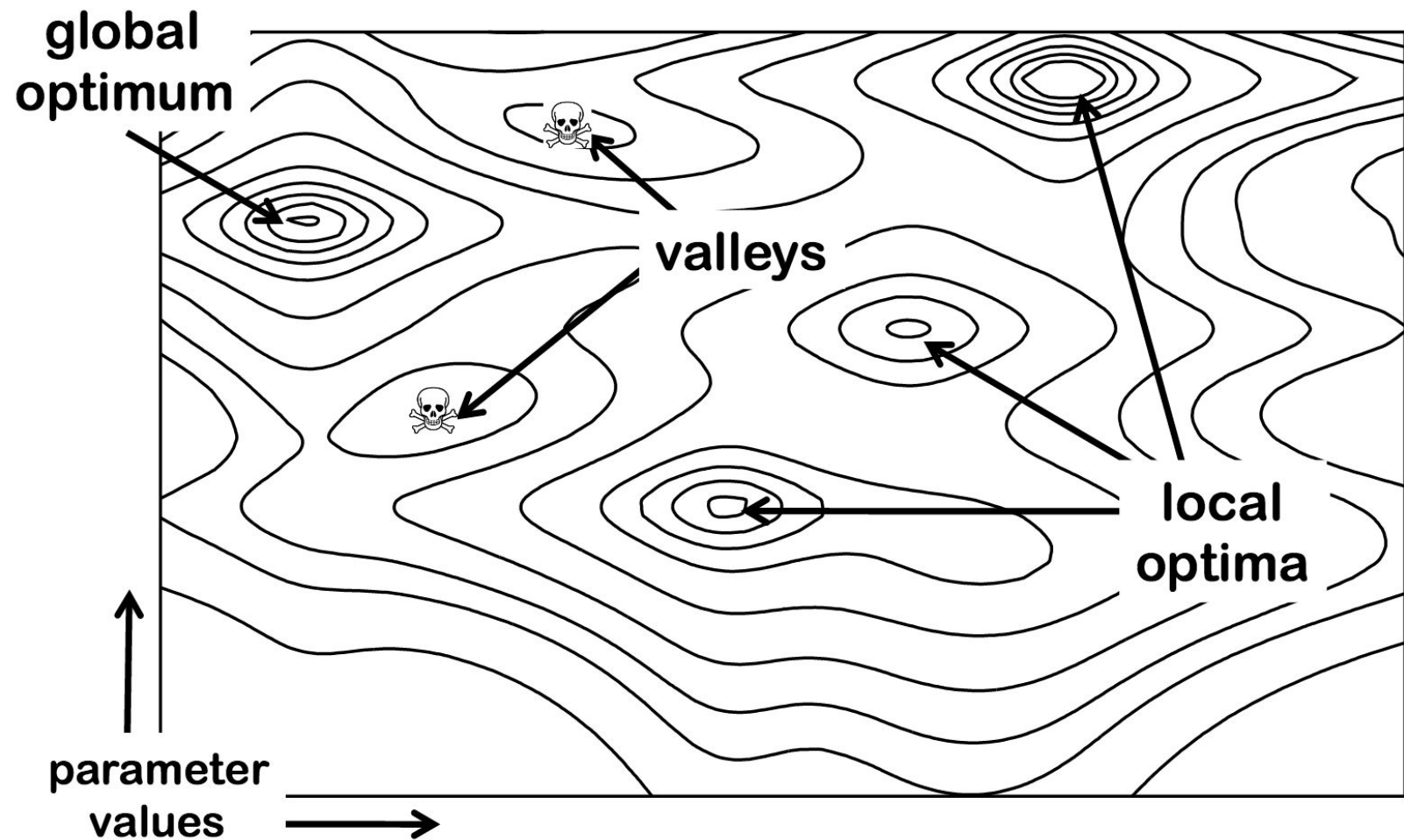
What do we observe?

- Massive amounts of missing data

What do we observe?

- Massive amounts of missing data
- $\frac{1}{8}$ of taxon pairs have no data that can be used to evaluate them with respect to one another

Tips	Number of unrooted (binary) trees	
4	3	
5	15	
6	105	
7	945	
8	10,395	
9	135,135	
10	2,027,025	
11	34,459,425	
12	654,729,075	
13	13,749,310,575	
14	316,234,143,225	
15	7,905,853,580,625	
16	213,458,046,676,875	
17	6,190,283,353,629,375	
18	191,898,783,962,510,625	
19	6,332,659,870,762,850,625	
20	22,164,309,5476,699,771,875	
21	8,200,794,532,637,891,559,375	
22	319,830,986,772,877,770,815,625	
23	13,113,070,457,687,988,603,440,625	> 21 moles of trees
24	563,862,029,680,583,509,947,946,875	



Missing data and topology estimation

- Random missing data are generally not a problem, if the sample size is large

Missing data and topology estimation

- Random missing data are generally not a problem, if the sample size is large
 - = reducing sample size

Missing data and topology estimation

- Random missing data are generally not a problem, if the sample size is large
 - = reducing sample size
- Biased missing data tend to cause a big problem

Tree estimation

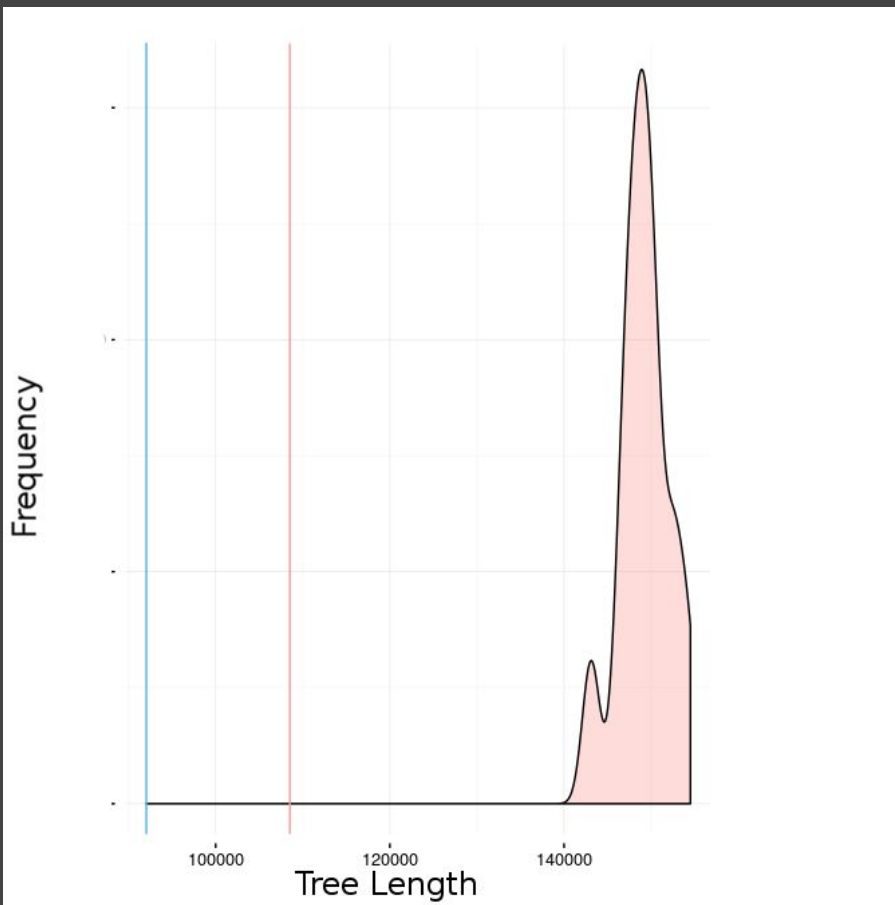
- Maximum likelihood

Tips	Number of unrooted (binary) trees	
4	3	
5	15	
6	105	
7	945	
8	10,395	
9	135,135	
10	2,027,025	
11	34,459,425	
12	654,729,075	
13	13,749,310,575	
14	316,234,143,225	
15	7,905,853,580,625	
16	213,458,046,676,875	
17	6,190,283,353,629,375	
18	191,898,783,962,510,625	
19	6,332,659,870,762,850,625	
20	22,164,309,5476,699,771,875	
21	8,200,794,532,637,891,559,375	
22	319,830,986,772,877,770,815,625	
23	13,113,070,457,687,988,603,440,625	> 21 moles of trees
24	563,862,029,680,583,509,947,946,875	

Tree estimation

- Maximum likelihood
 - RAxML
 - Bootstraps

Branch Lengths

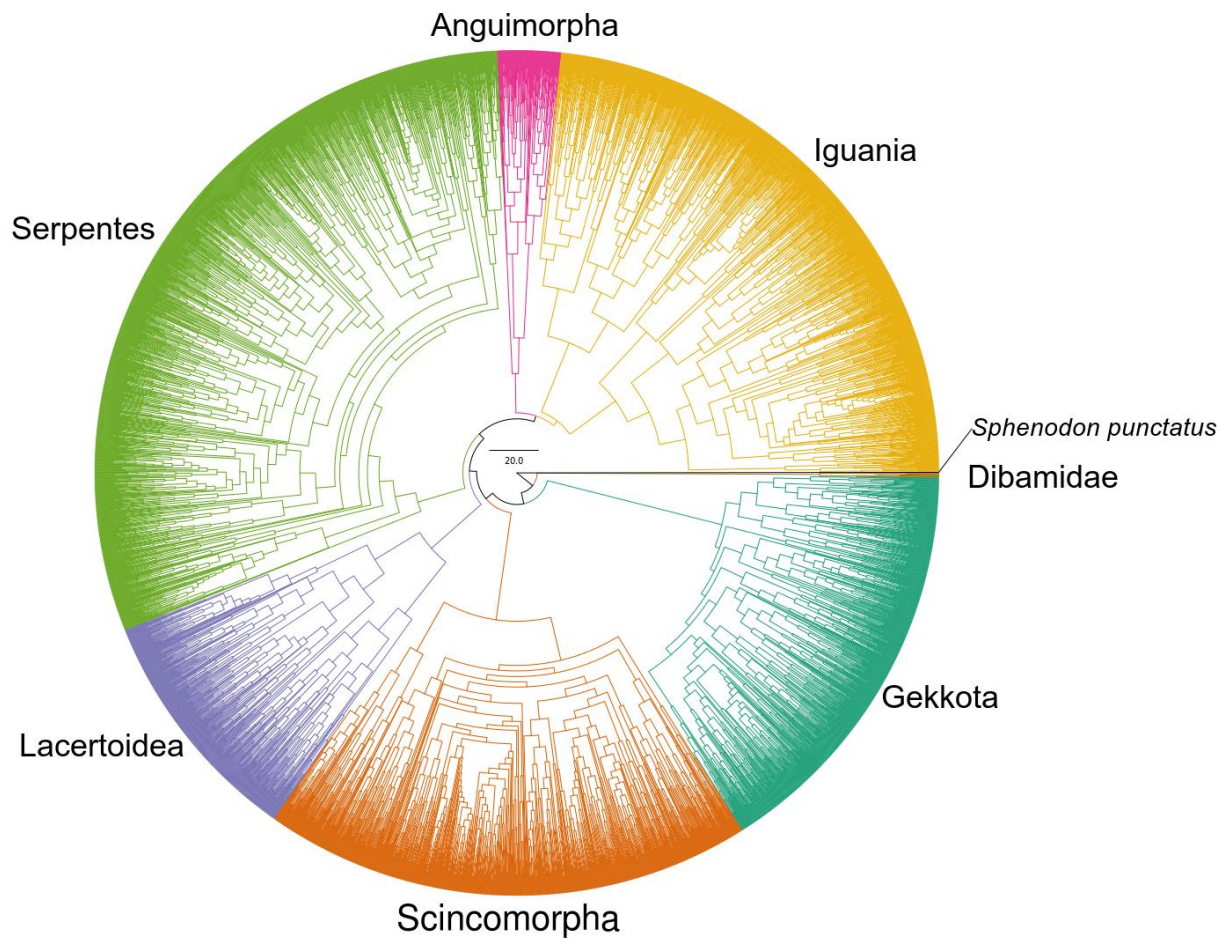


Tree estimation

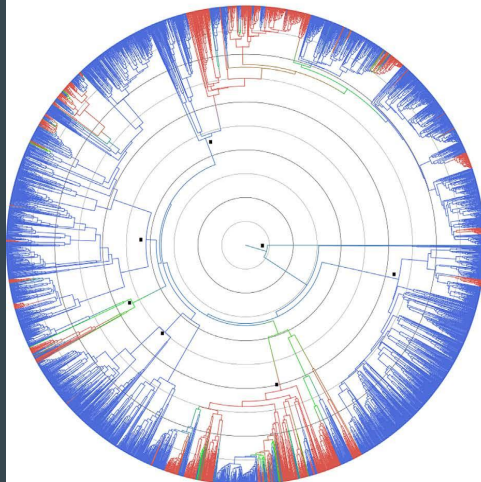
- So we compromised

Tree estimation

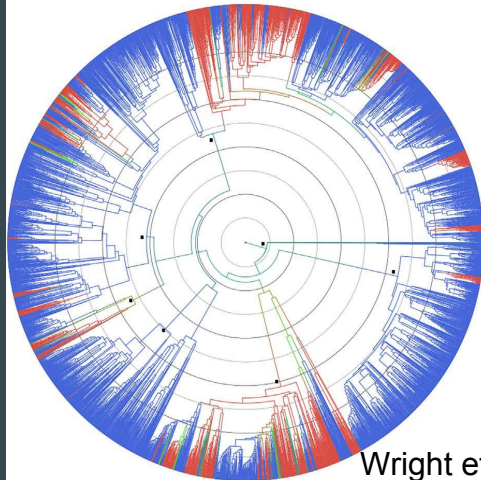
- So we compromised
 - First pass estimation in RAxML
 - Optimization in Garli
 - For just branch lengths
 - For BL + topology



Tree
Ln Likelihood

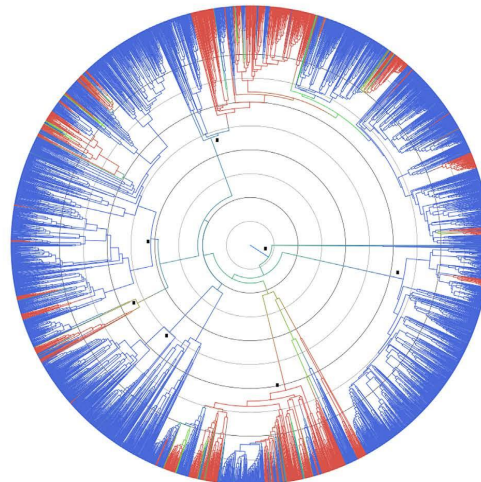


Tree
Ln Likelihood

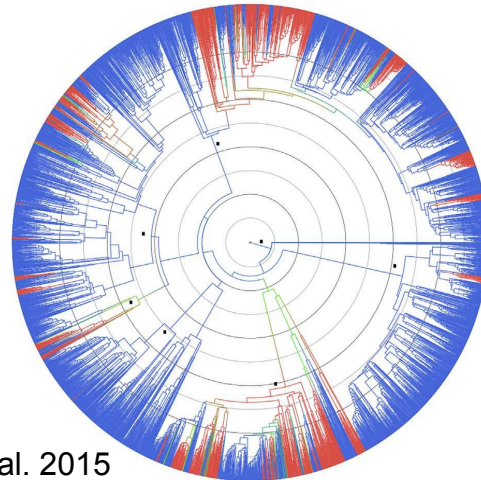


Branch Length Improved
-2,595,843

ExaML Best-Scoring
-2,598,386



Branch Length and Topology Improved
-2,592,703



iPython Notebook

RF_Checking

iPython Notebook

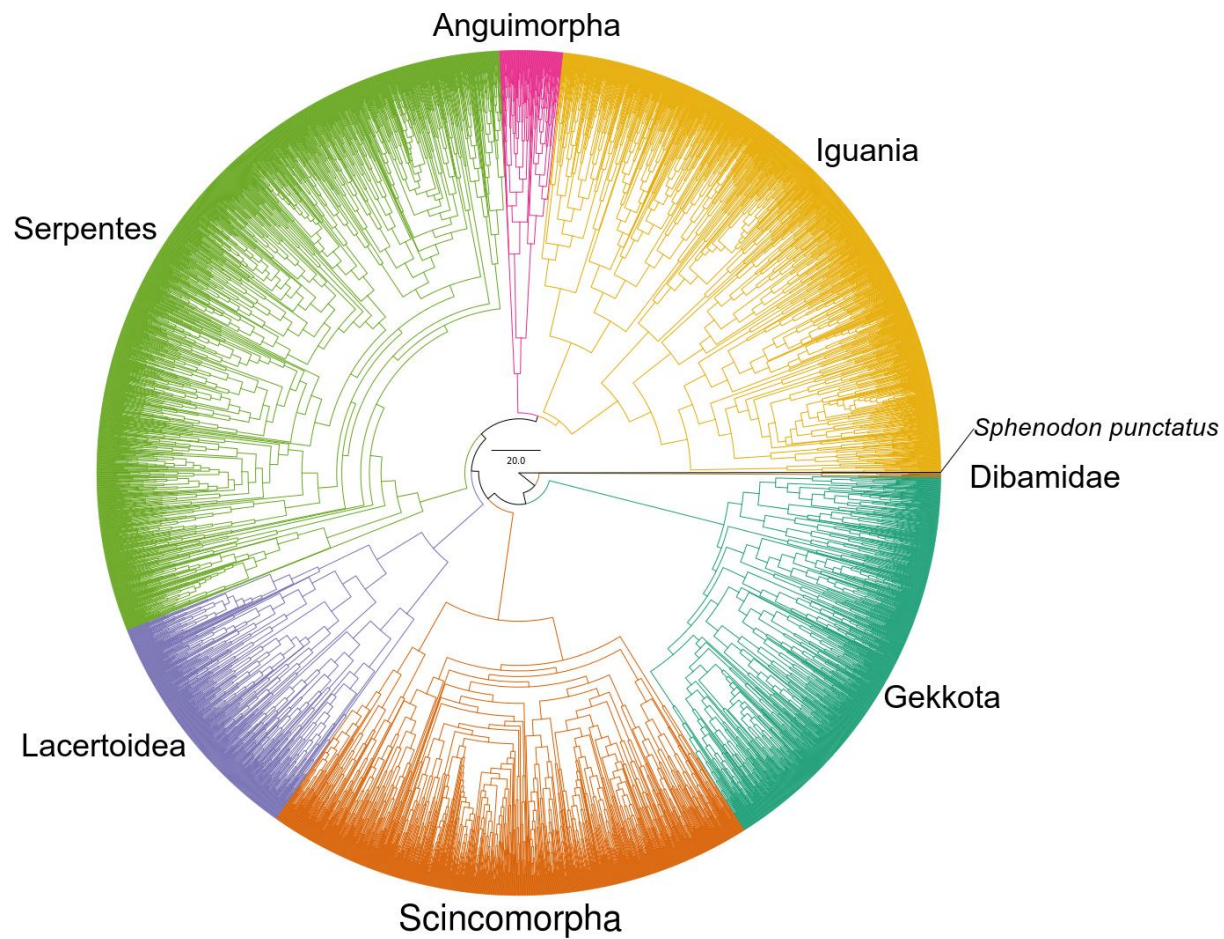
RF_Checking

- What I want you to note is that there are 854 differences between the Pyron and Burbrink tree and our best one.

iPython Notebook

RF_Checking

- What I want you to note is that there are 854 differences between the Pyron and Burbrink tree and our best one.
 - Total possible differences 8118
 - About 10% of the tree is different



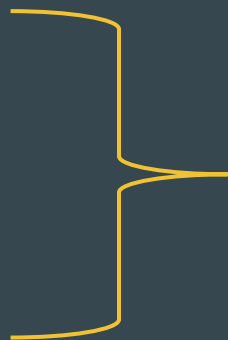
State-dependent diversification models

State-dependent diversification models

- Full BiSSE model
 - Two speciation parameters
 - Two extinction parameters
 - Two transition rates

State-dependent diversification models

- Full BiSSe model
 - Two speciation parameters
 - Two extinction parameters
 - Two transition rates



One of each
parameter for
oviparity and
viviparity

State-dependent diversification models

- R CMD BATCH bisse.R

Let's look inside the script

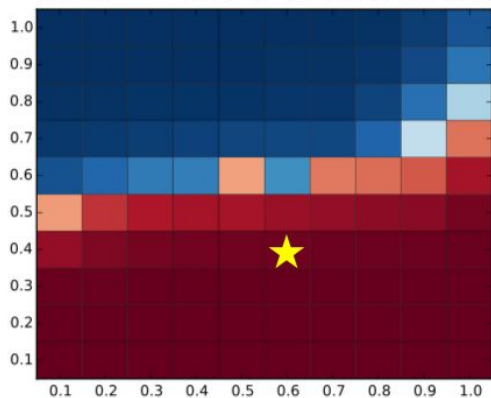
Heatmaps

Pyron and Burbrink

Tree

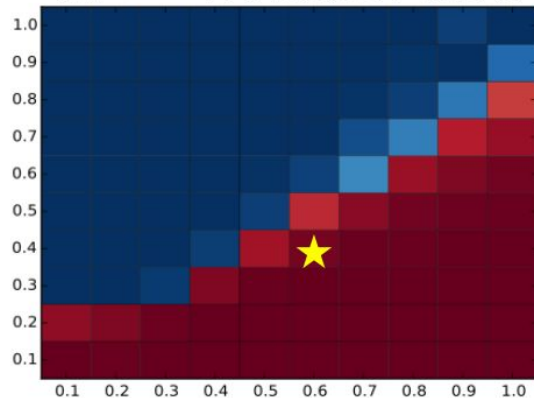
$\ln L = -2676499.387$

Proportion of Viviparous
Taxa Sampled



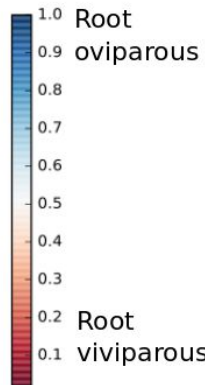
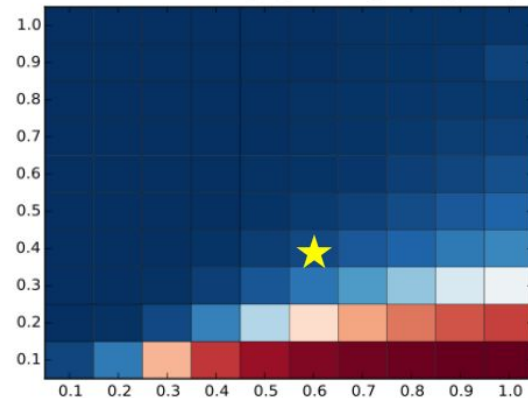
ExaML Tree

$\ln L = -2605097.638$



Garli-Optimized Tree

$\ln L = -2592702.940$



Proportion of Viviparous Taxa Sampled

Show that 4-tree figure again, zoom in on one.

Show that 4-tree figure again, zoom in on one.

- yes ma'am.

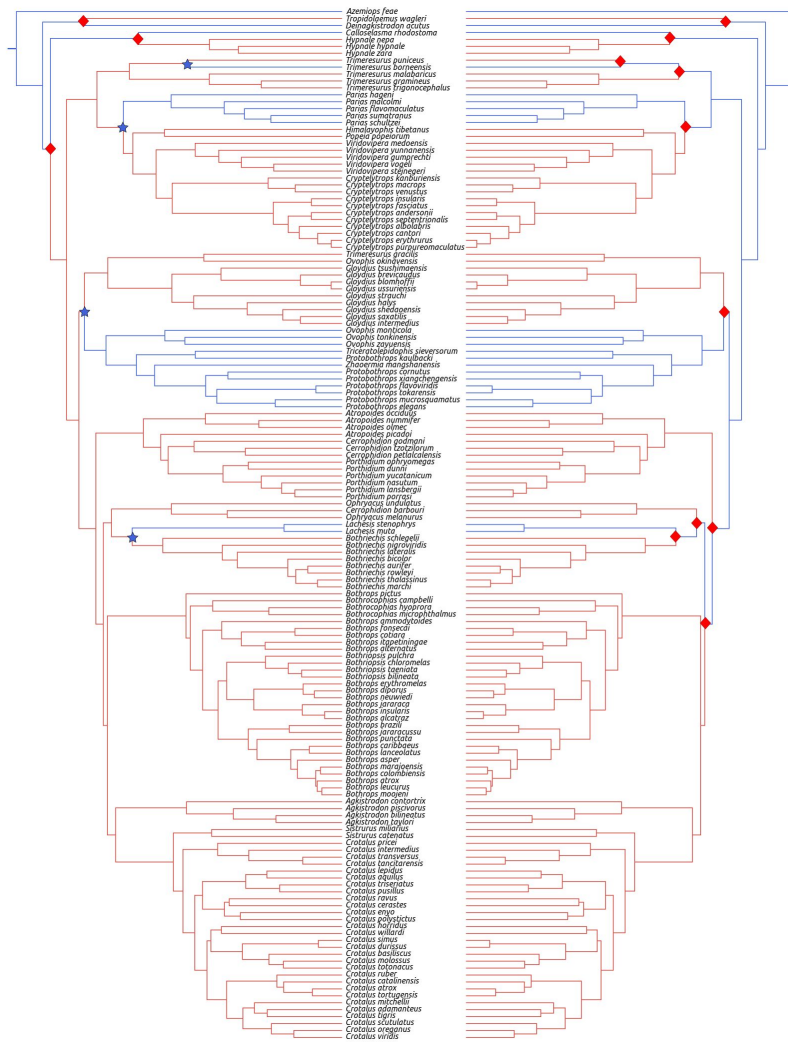


Table 2. Number of changes to viviparity and oviparity, and support for the root state, as estimated across four trees and three methods of taxon-sampling correction.

Tree	Origins of viviparity	Minimum number of reversals to oviparity				Root state (% support for oviparity)		
		1X	2X	3X	4X	No sampling correction	One term correction	Two term correction
Pyron and Burbrink tree	106–129	13	2	1	0	80.8	7.2	7.4
ExaML tree	103–126	12	3	1	0	75.0	87.9	11.2
Branch-length improved tree	103–124	12	3	1	0	90.4	97.5	24.4
Best tree	98–129	18	3	0	0	88.6	94.5	14.3

The ranges for origins of viviparity indicate the number of origins under the equal weighted model to the number of origins assuming no reversals to oviparity. Reversals to oviparity are shown for four models of character change: 1X = equal weighting of changes between viviparity and oviparity; 2X = twofold greater penalty of reversals to oviparity; 3X = threefold greater penalty of reversals to oviparity; 4X = fourfold greater penalty of reversals to oviparity.

Conclusions

The story we tell about this trait depends on:

Conclusions

The story we tell about this trait depends on:

- The tree we use

Conclusions

The story we tell about this trait depends on:

- The tree we use
- The model we use

Conclusions

The story we tell about this trait depends on:

- The tree we use
- The model we use
- How we parse the model output

Future Directions

- Threshold model