Comparing new measures of phylogenomic support

Genevieve G. Mount[1], Jeremy M. Brown[1]

[1]Department of Biological Sciences and Museum of Natural Science, Louisiana State University, Baton Rouge, LA

@Lizardeve
gmount1@lsu.edu

Discordant topologies, each with strong support values

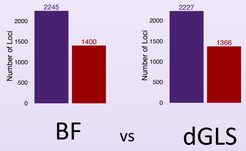
- Currently a problem with many controversial phylogenetic relationships. E.g. Birds, turtles, metazoans.
- Bootstrap values & posterior probabilities − not perfect.
- Maximized by large data sets.
- Measures stochastic error and sampling variance.
- No information about conflict within data set.

Current approaches

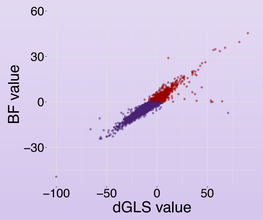
- Assess relative support for one topology over another.
- Quantify variation in support across topologies and loci.
- Agnostic to cause of discordance.
- **Bayes Factors (BF)** - Marginal likelihood ratio. Relative fit of model to data.
- **Difference in gene log-likelihood scores (dGLS)** - Sum of site specific likelihoods across locus. Maximum likelihood given model.

Do the majority of loci support the same topology across methods?

- Calculate number of loci with “strong” support.
- BF > 10 dGLS > 0.5



Do locus support values correlate across methods?



- Compare BF and dGLS values for each locus.

Purple and **red** represent an example of two conflicting topologies.

Topological hypotheses in Squamates

(snakes and lizards)

Scleroglossa | Toxicofera | Anguimorph Iguania | Snake Anguimorph | Iguania Snake

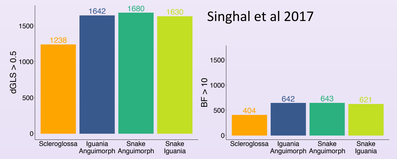
- Scleroglossa supported by morphology.
- Toxicofera supported by genetic data.
- Relationships within Toxicofera conflicting and/or have low support across studies.

Preliminary Results

Do the majority of loci support the same topology across methods?

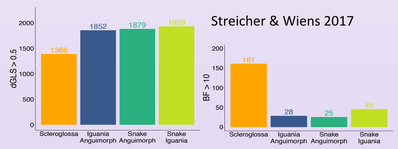
Yes!

The proportion of loci strongly supporting each topology is similar between BF and dGLS



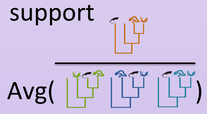
No?

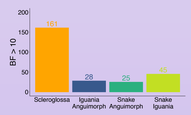
More loci strongly supported scleroglossa when using BF.



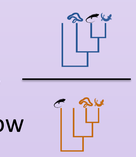
- Results varied between data sets, requiring further investigation…
- We found that preferred topology may change depending on how support values are calculated for some data sets

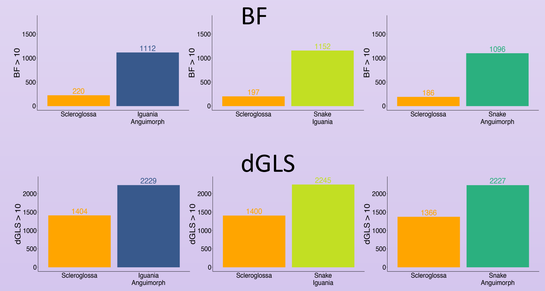
- **Compare against average for all other topologies**
  - Unequal variation in support for topologies obscures loci counts of support

Avg()



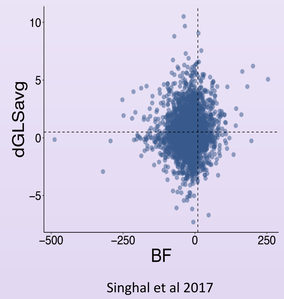
- **Pairwise comparison for each topology**
  - Pairwise support values calculated for each topology
  - Now, the majority of loci show strong support for the same topology across methods

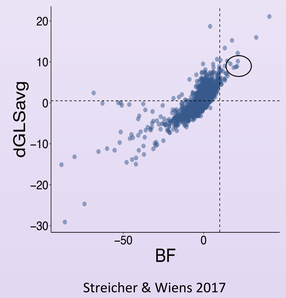




Do locus support values correlate across methods?

- In some cases large BF also have large dGLS values
- Loci with conflicting support values may provide further insight





Discussion, implications, and take homes

- Some correlation of support values across methods.
- Loci with support values that disagree may provide insight into molecular evolution, phylogenetic models.
- Majority of loci support same topology for pairwise comparisons.
- Number of loci supporting each topology may change depending on how topological hypotheses are compared.
- What do you think?

Looking for post-doc position!!

Computational and/or empirical research

Available Summer/Fall/Winter 2020

Acknowledgements: This work was supported by NSF grant DBI-1262571 and DEB-1355071 to JMB. Center for Computing and Technology for computational resources.