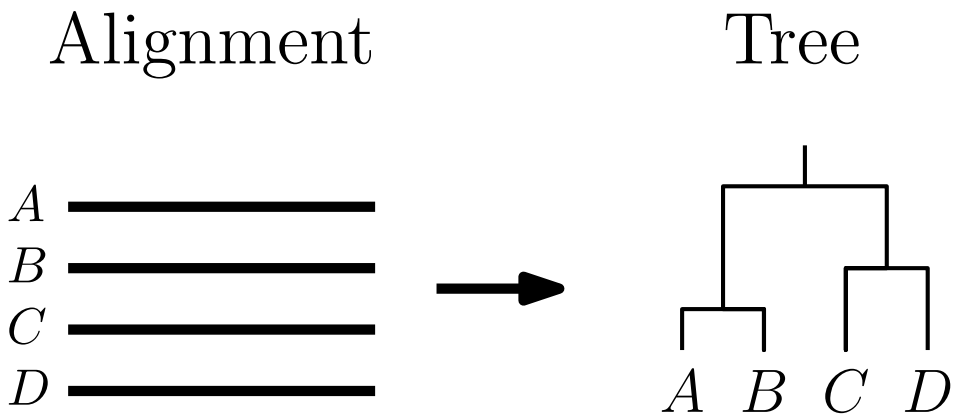
Instability in Phylogenetic Trees after Taxon Addition

Lena Collienne, Mary Barker, Marc Suchard, Erick Matsen

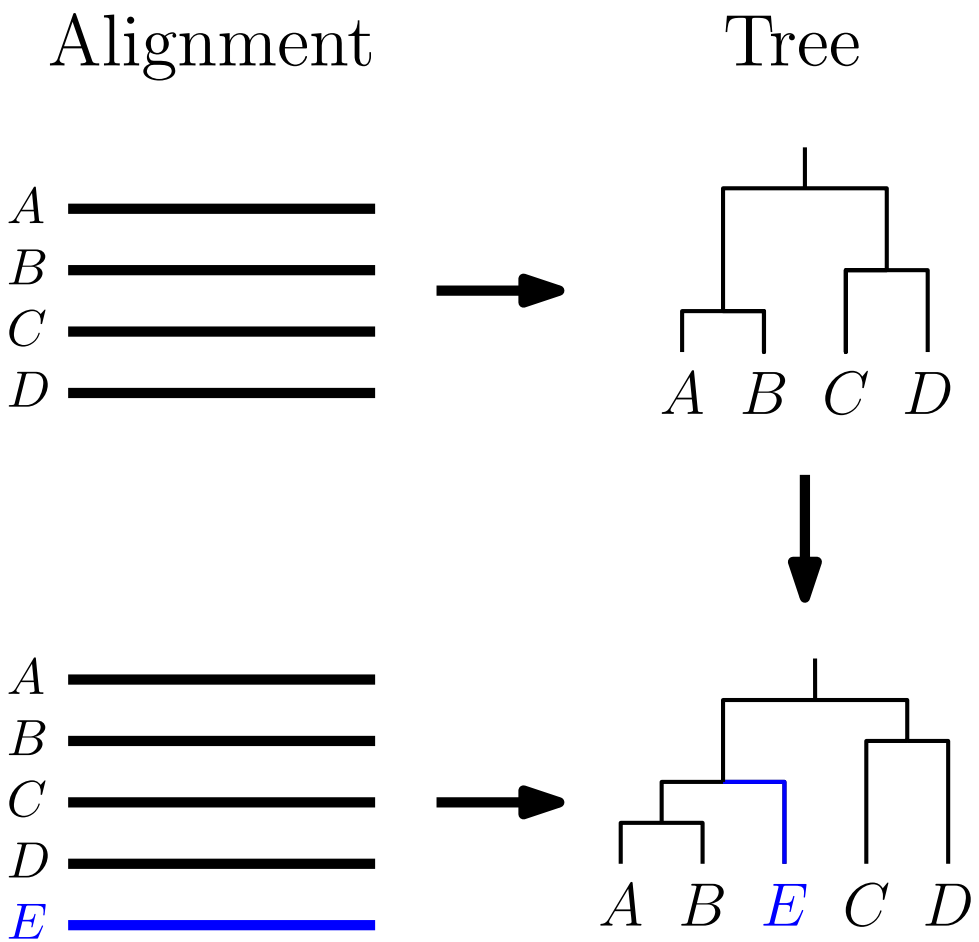
Fred Hutch Cancer Center

July 28, 2024

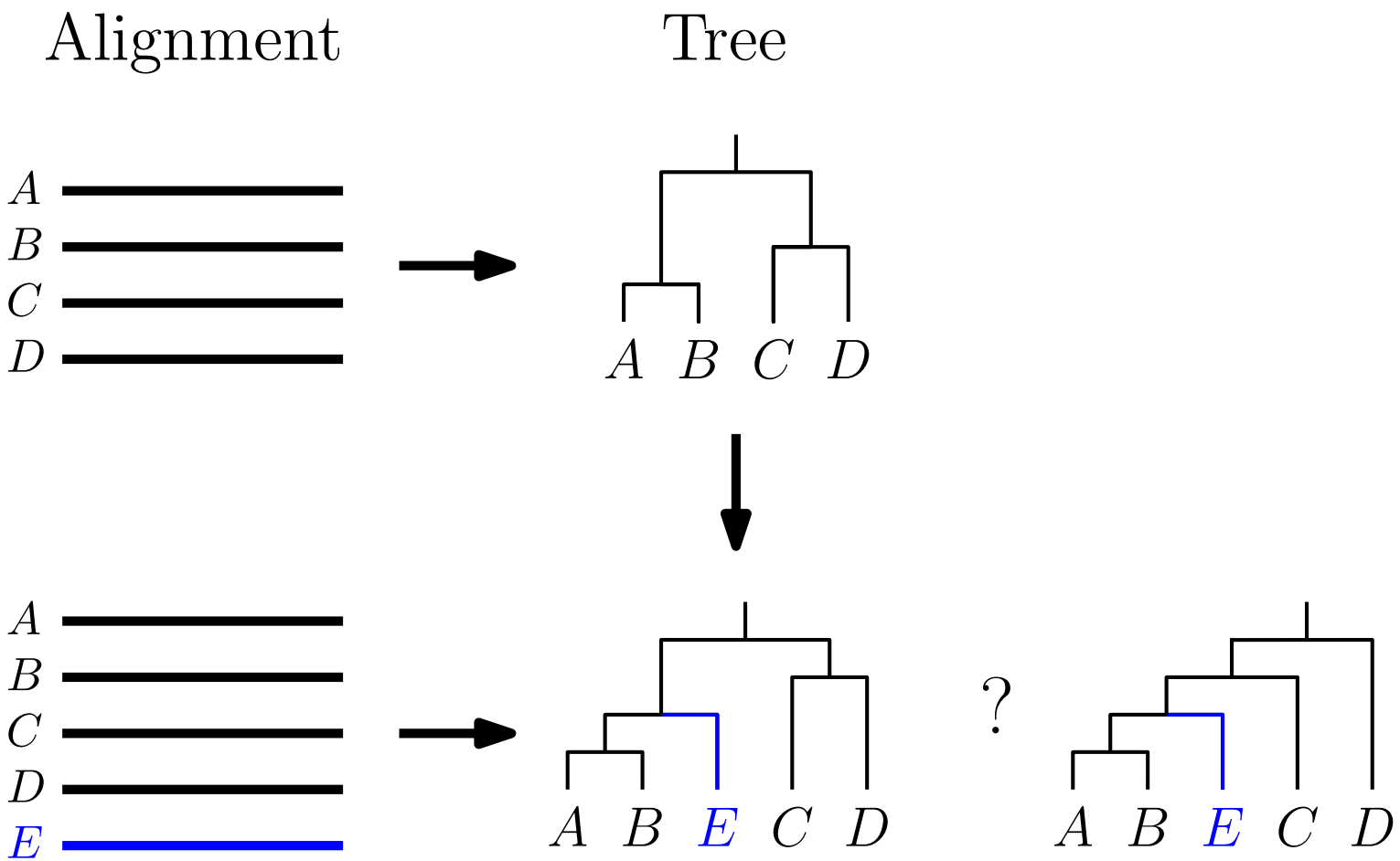
Online Phylogenetic Inference



Online Phylogenetic Inference

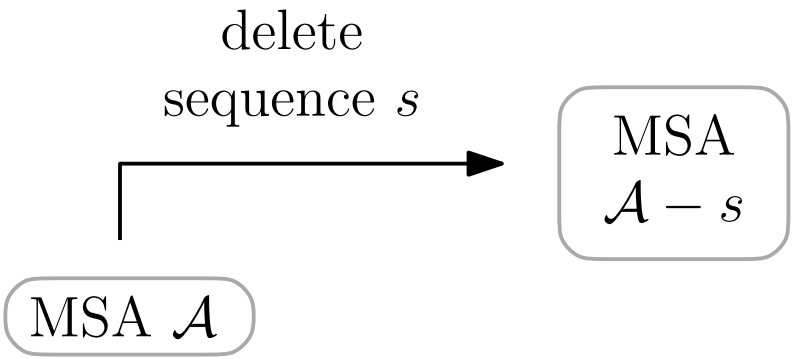


Online Phylogenetic Inference

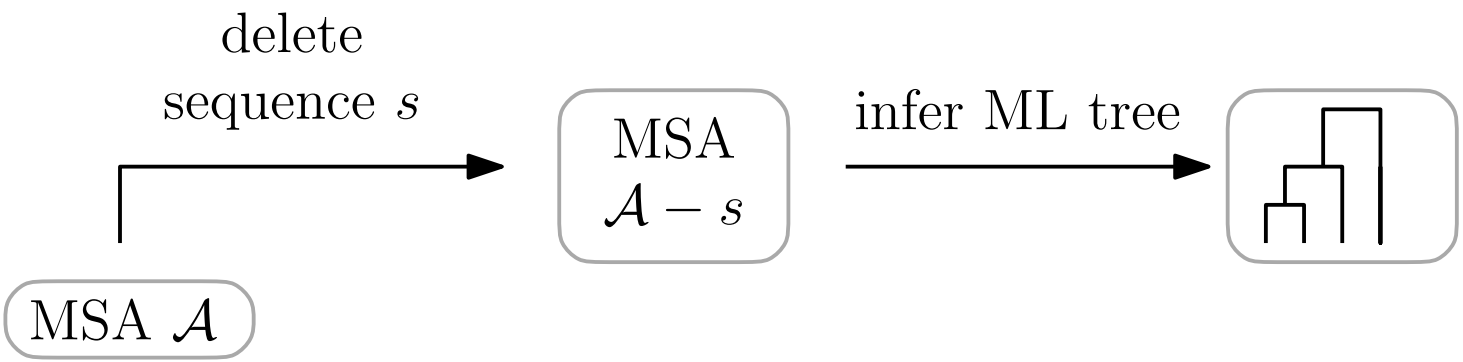


Measuring Stability

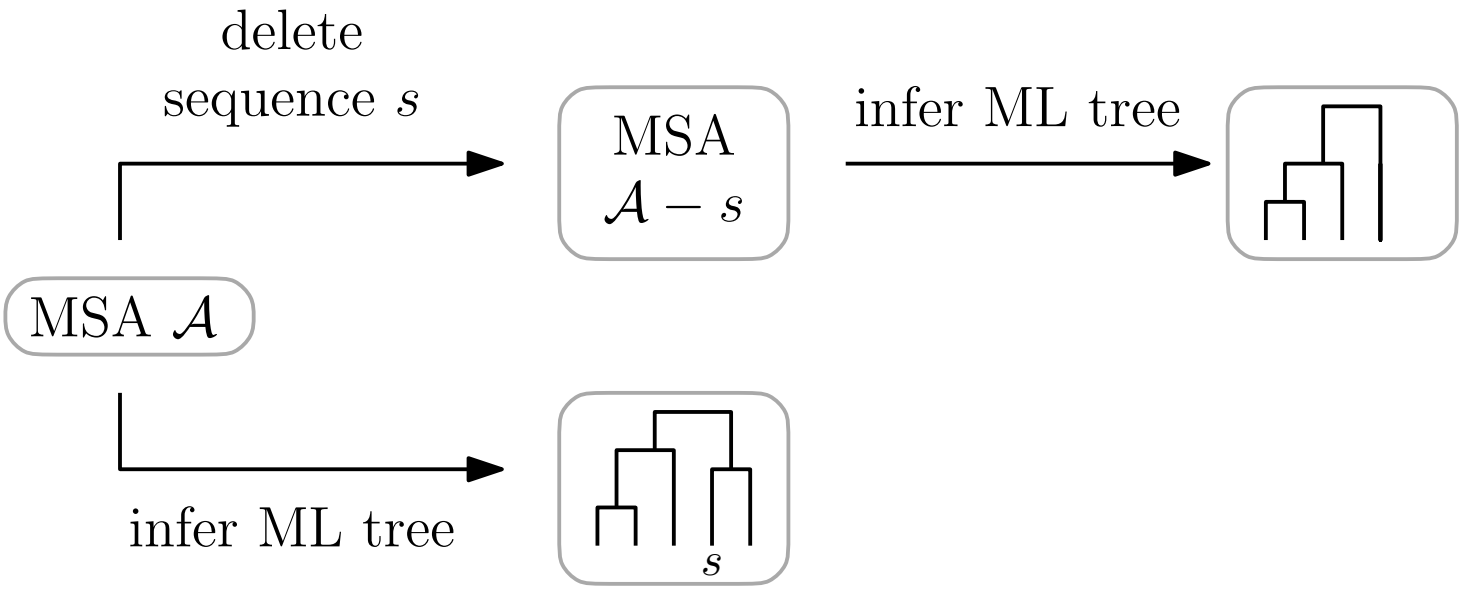
Measuring Stability



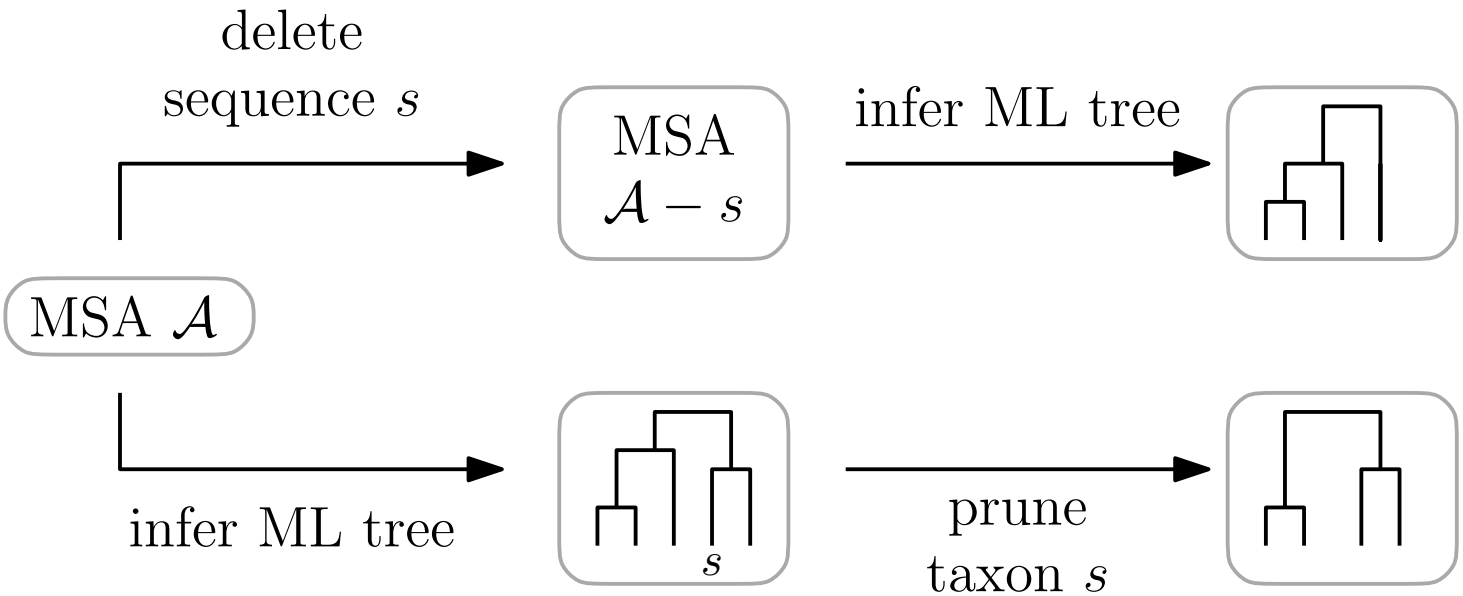
Measuring Stability



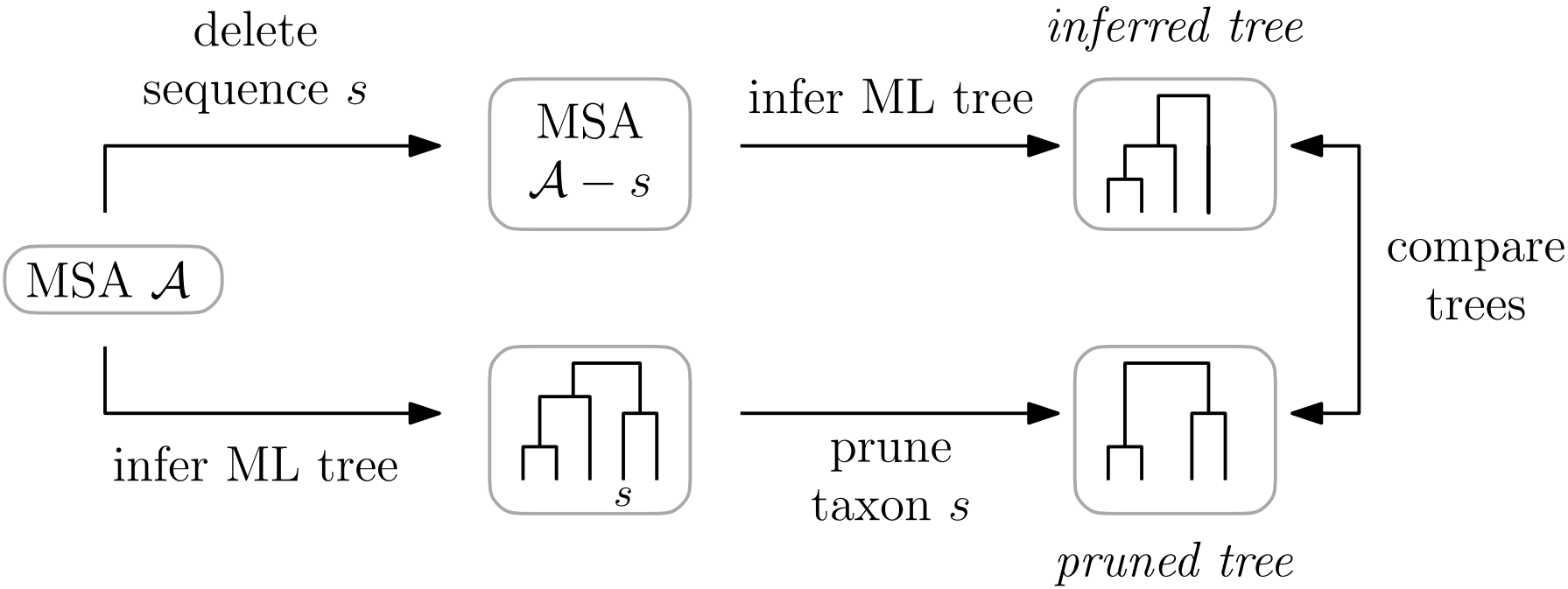
Measuring Stability



Measuring Stability

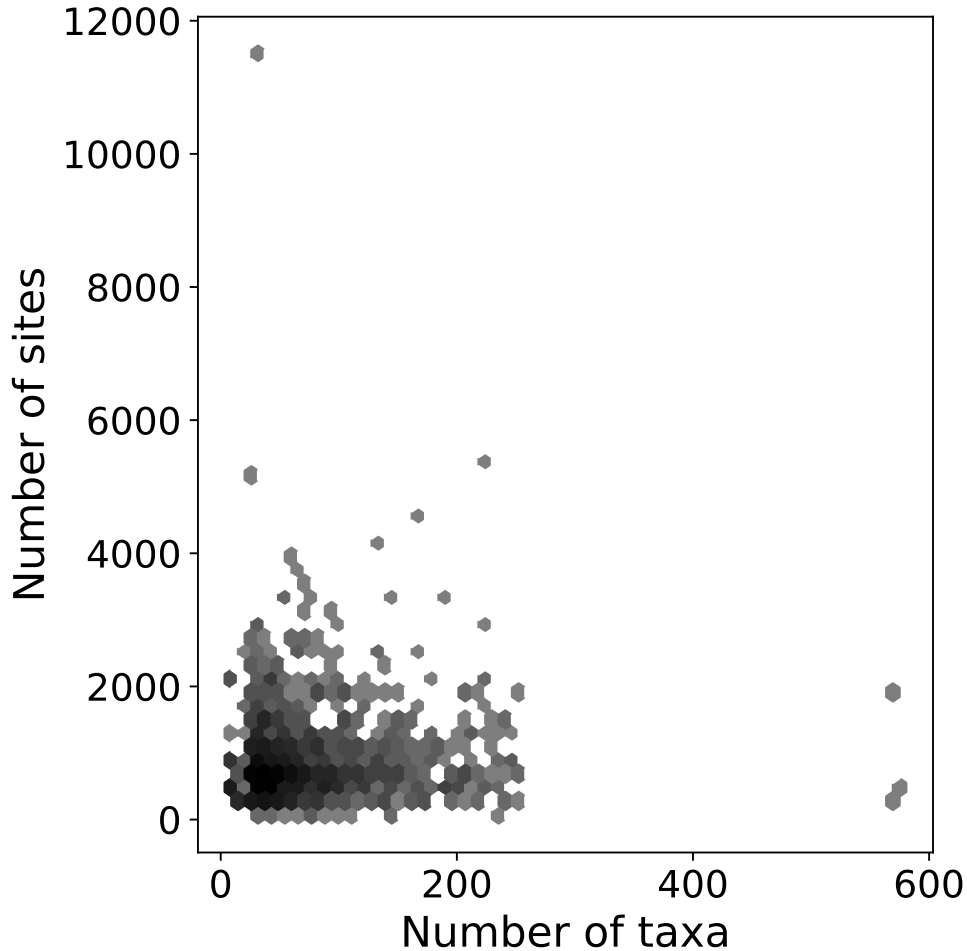


Measuring Stability



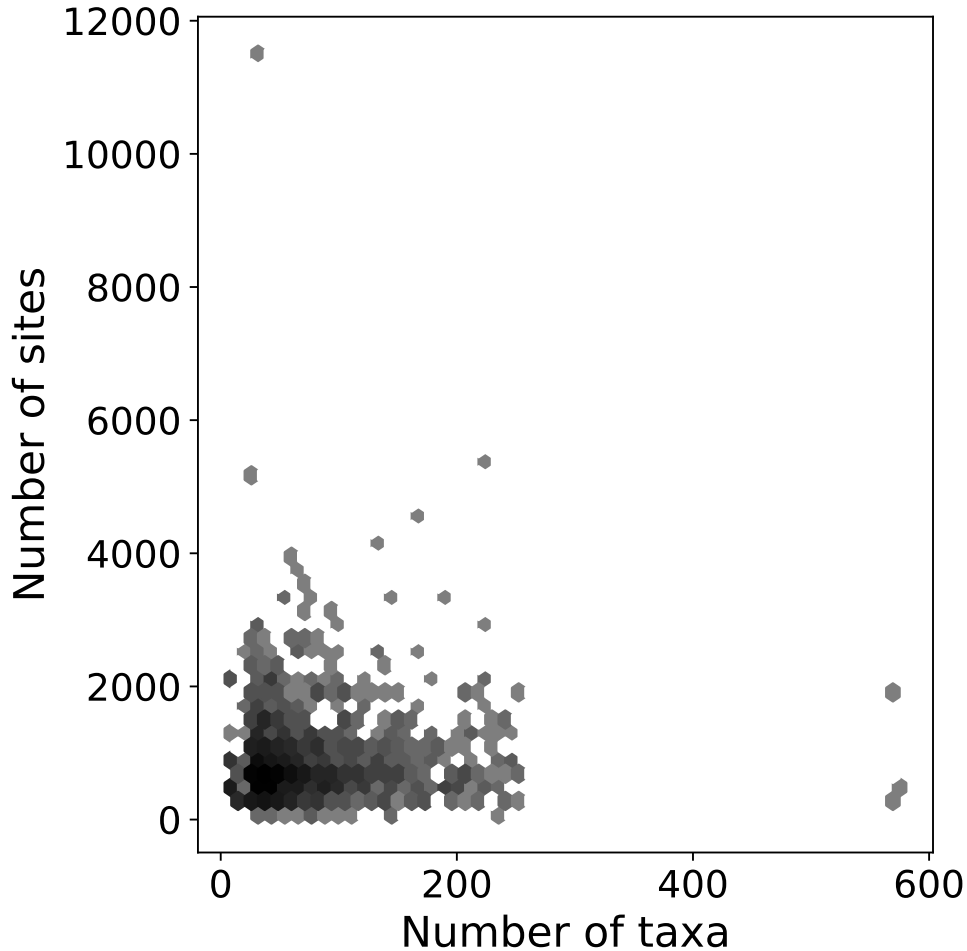
Data

Harrington et al. (2021)



Data

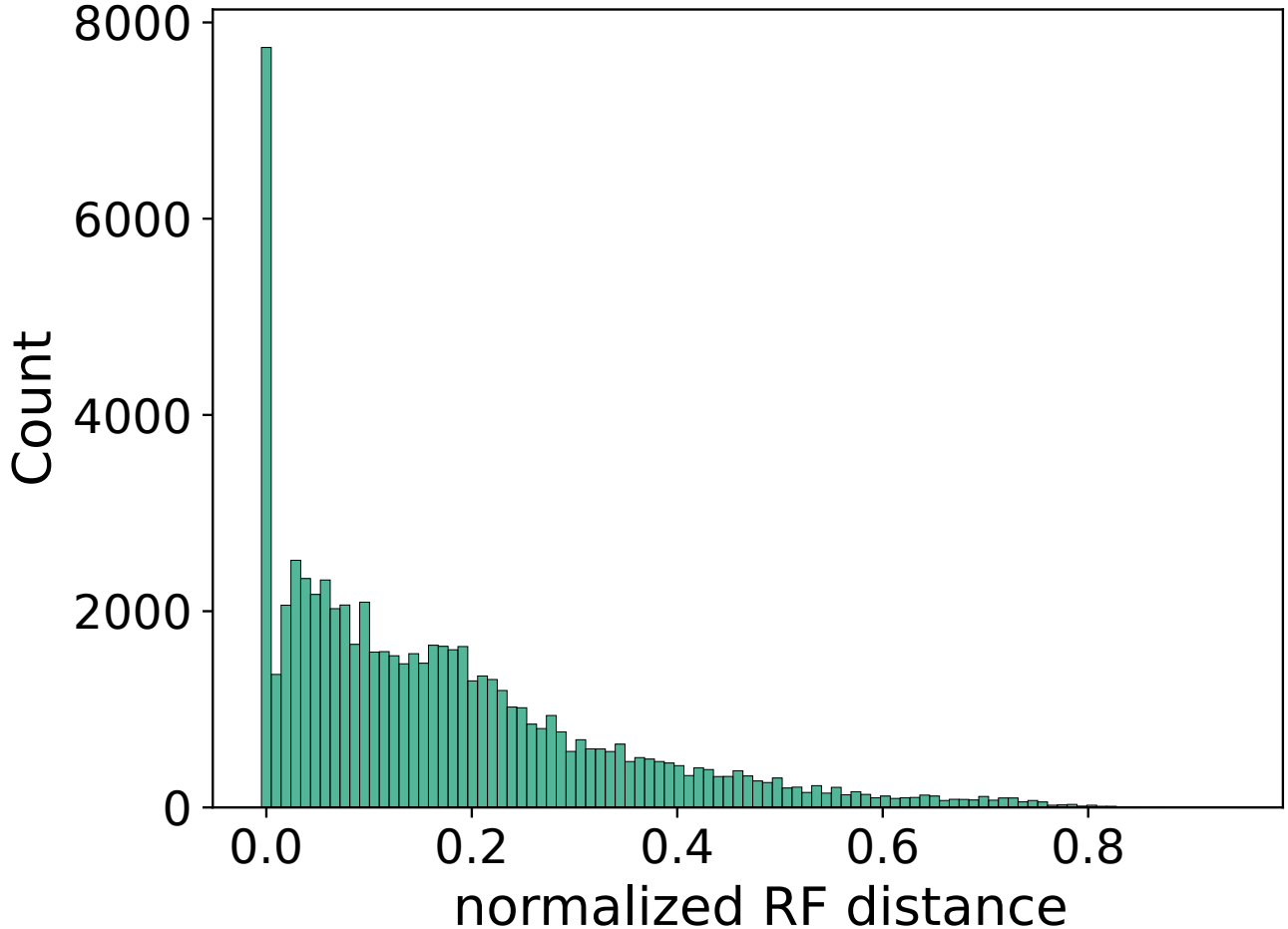
Harrington et al. (2021)



$\Rightarrow 67,709$ taxa

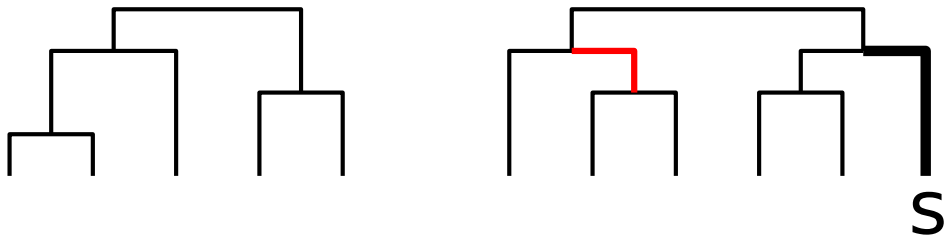
Data

Robinson-Foulds Distance



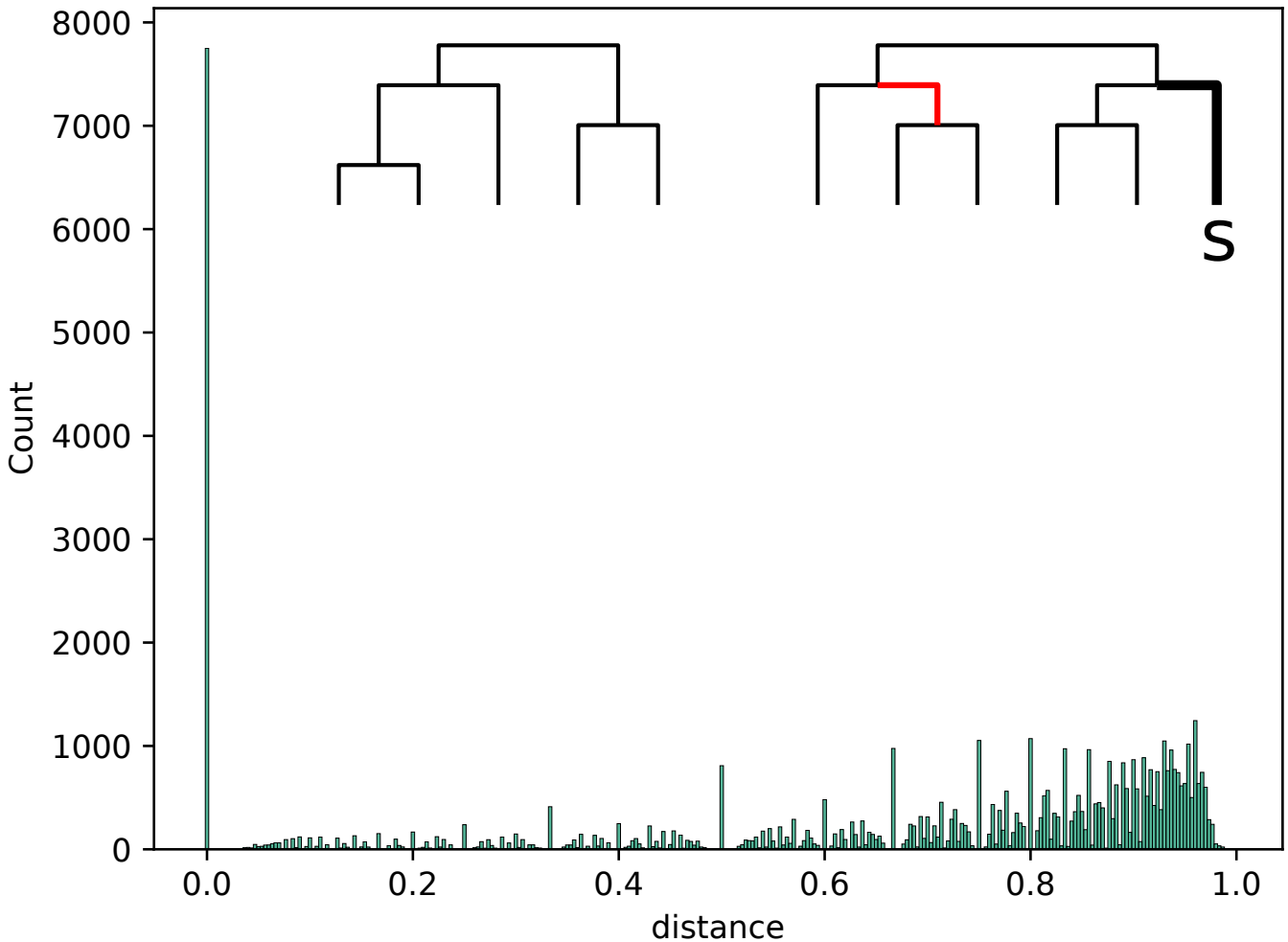
Regions of Instability

Distance to added taxon



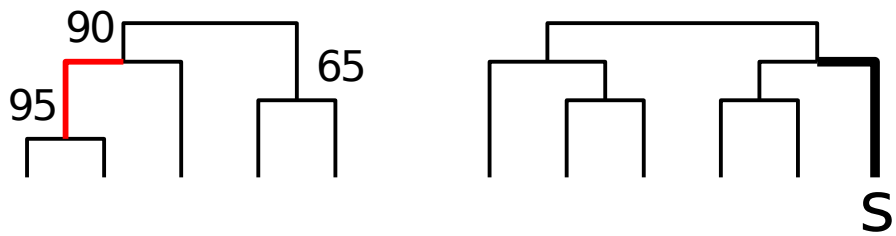
Regions of Instability

Distance to added taxon



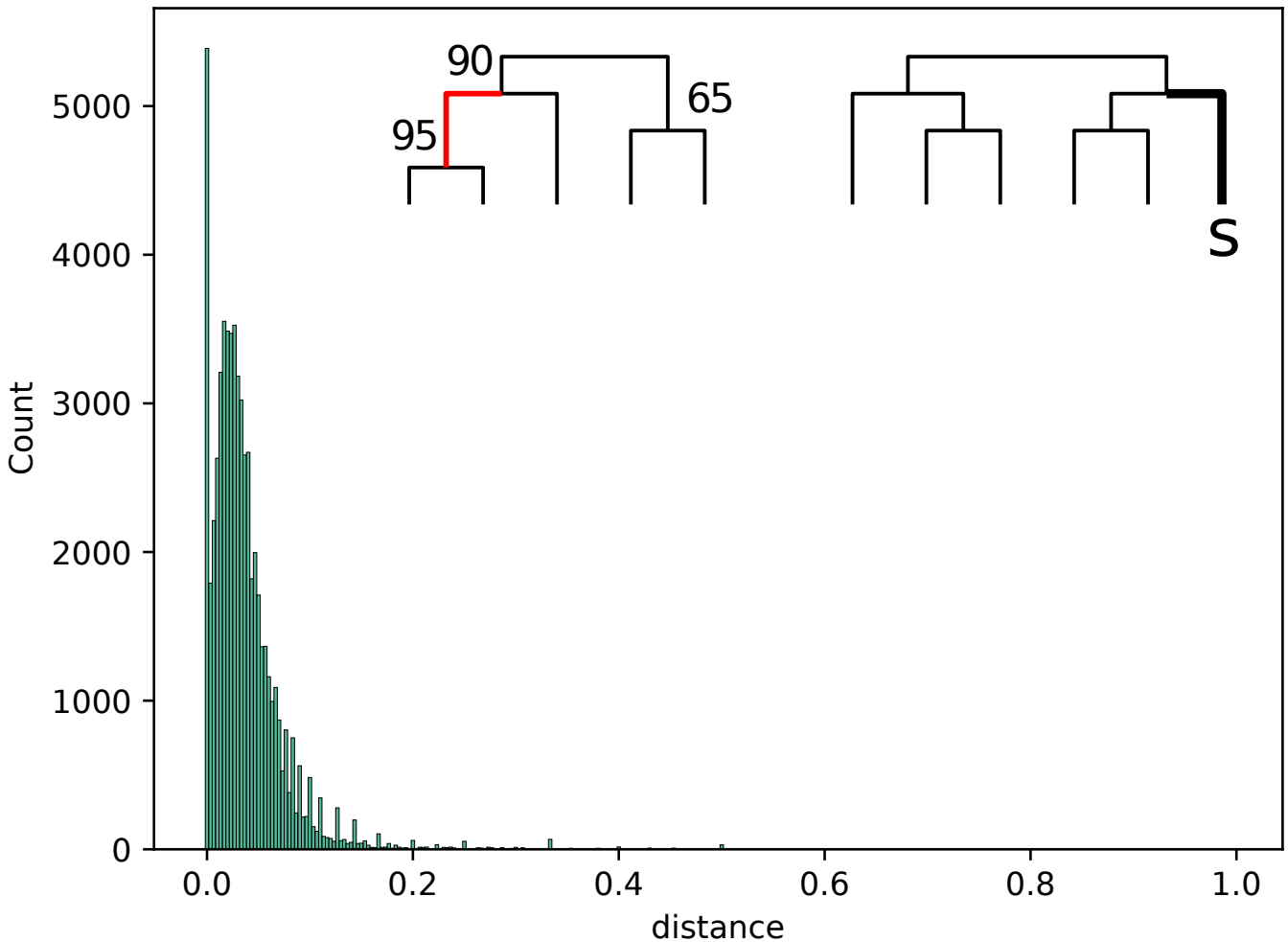
Regions of Instability

Distance to low bootstrap edges

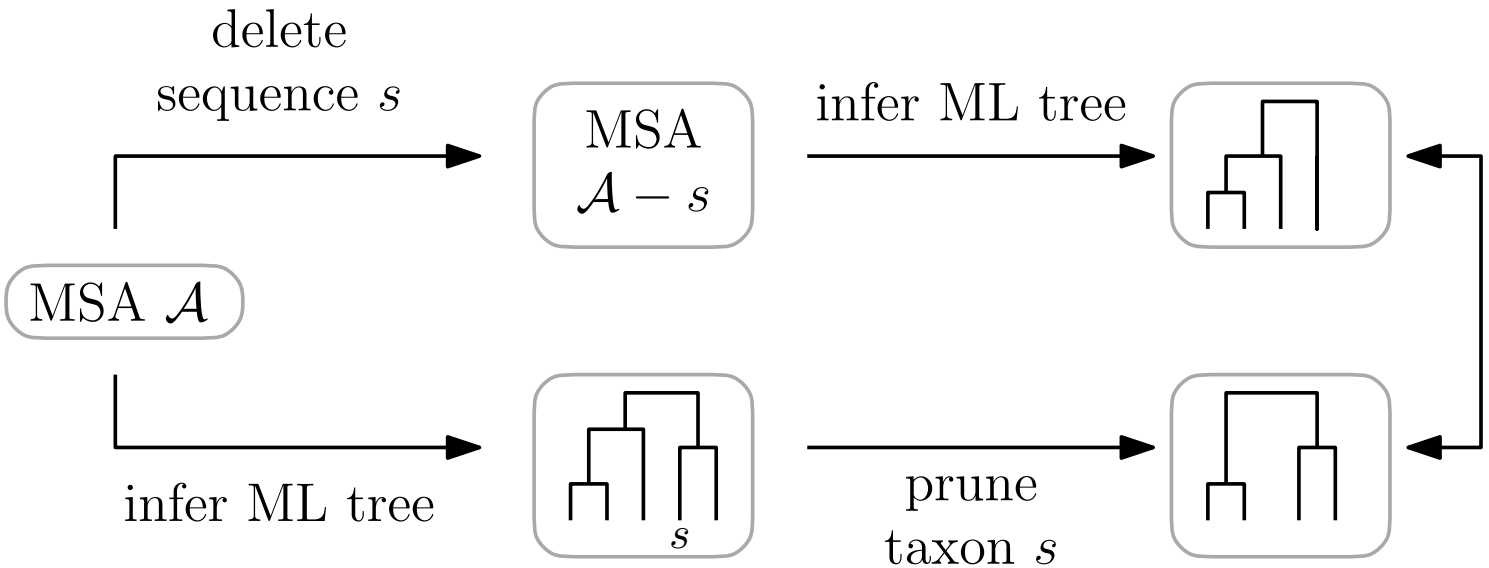


Regions of Instability

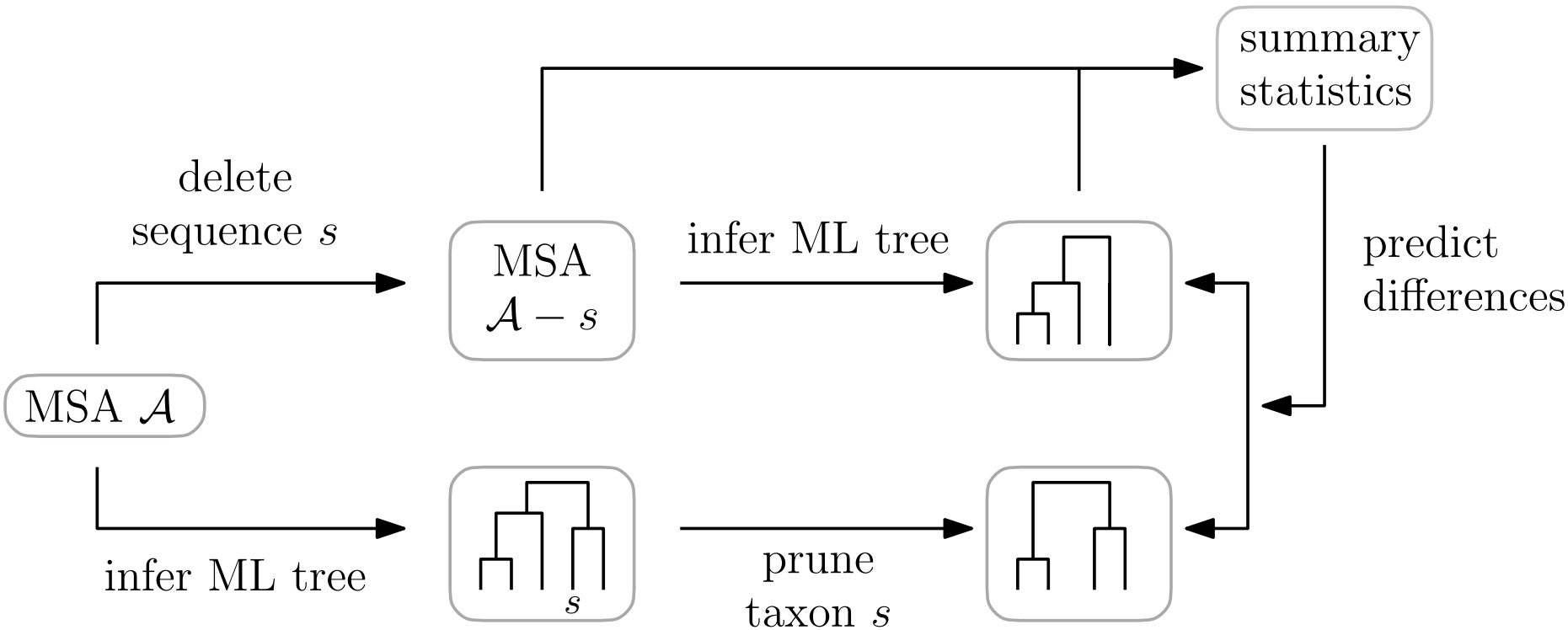
Distance to low bootstrap edges



Predicting Instability

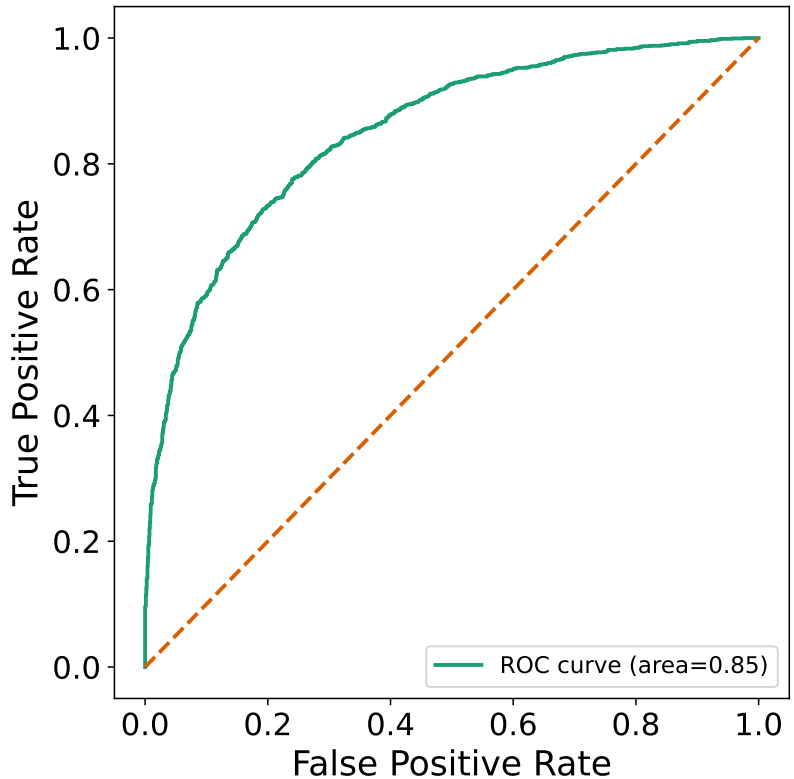


Predicting Instability



Predicting Instability

Random Forest Classification

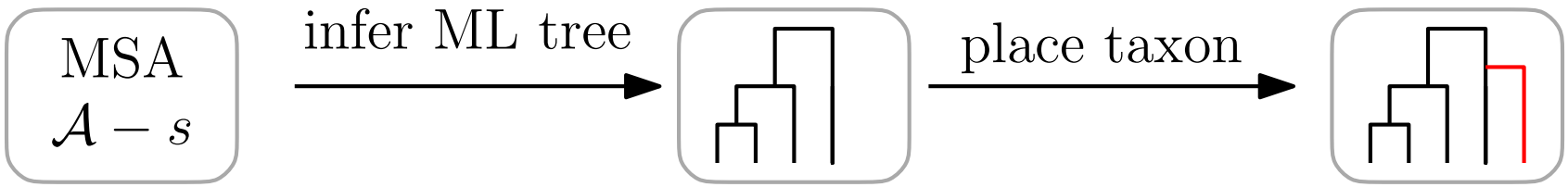


Predicting Instability

Predicting Instability

High importance:

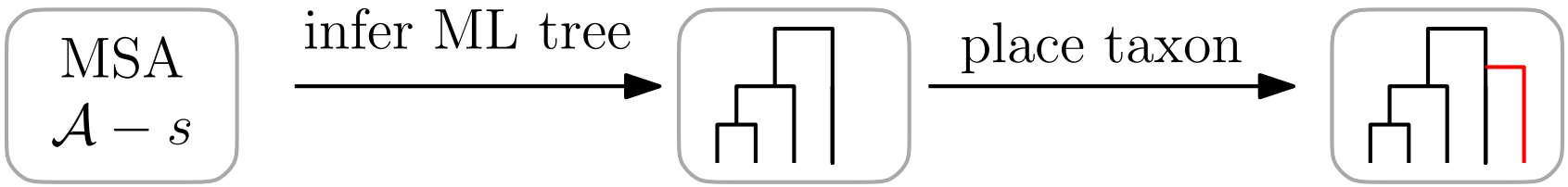
- Distance of best taxon placement to low bootstrap support edge



Predicting Instability

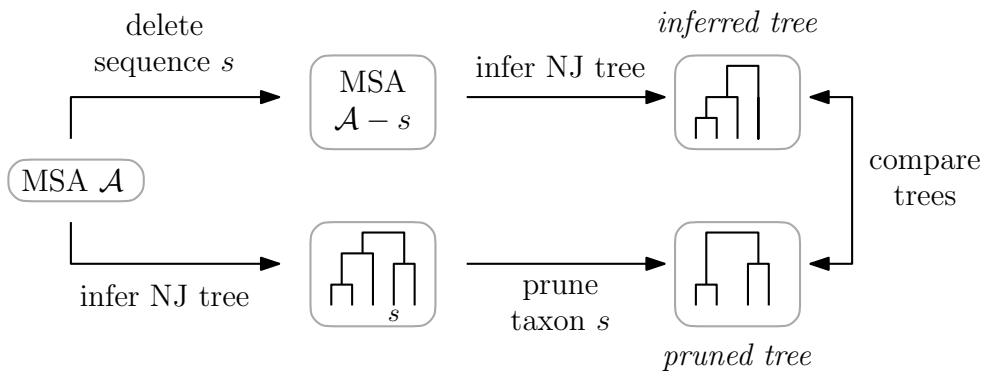
High importance:

- Distance of best taxon placement to low bootstrap support edge



Low importance:

- NJ stability

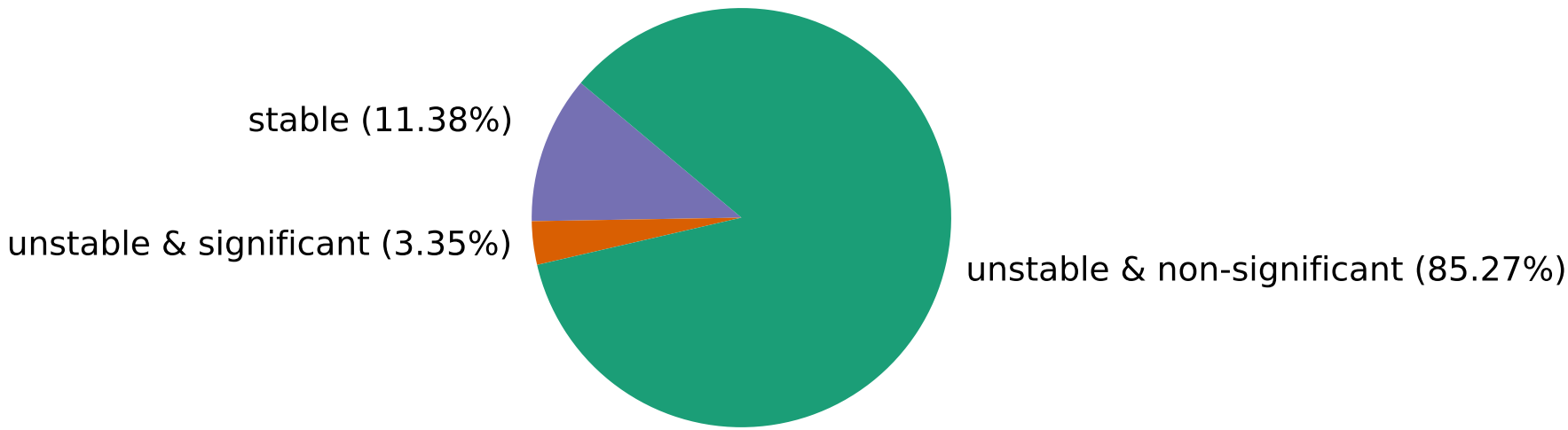


Acknowledgements

- Mary Barker
- Marc Suchard
- Erick Matsen

Supplement

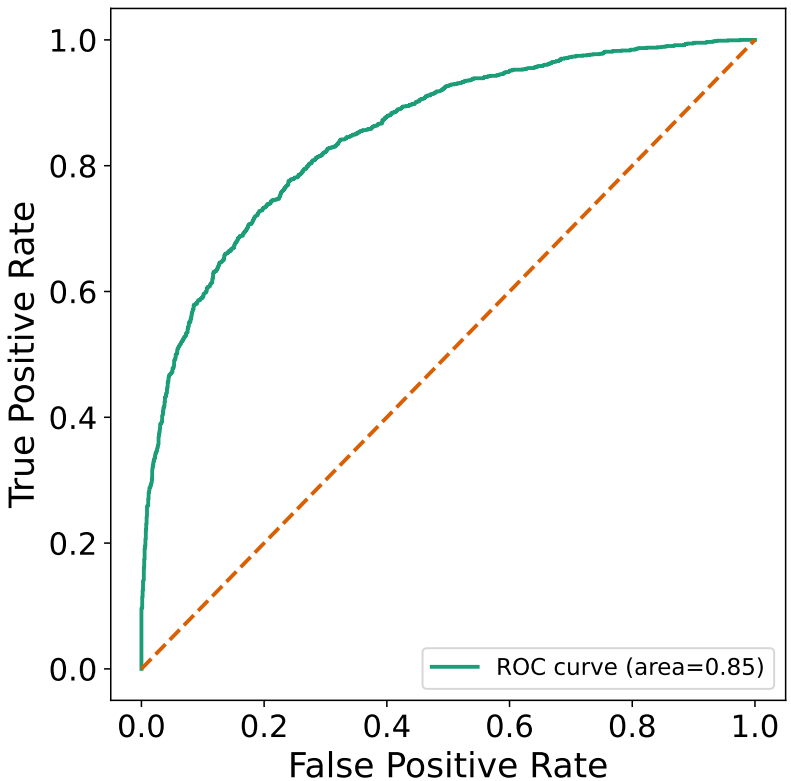
AU-test



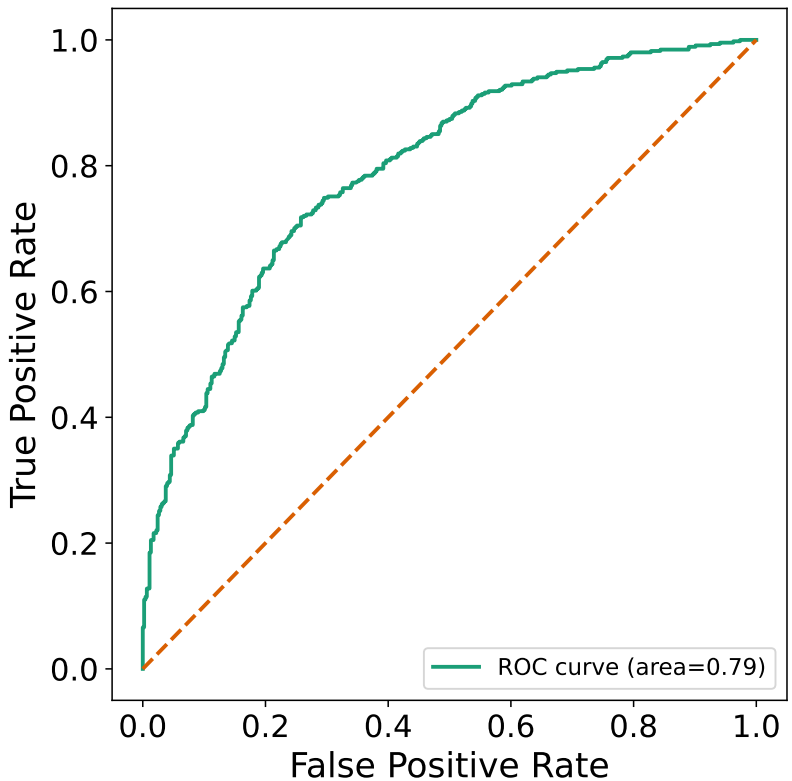
Supplement

Random Forest Classification

Tree Topology Difference



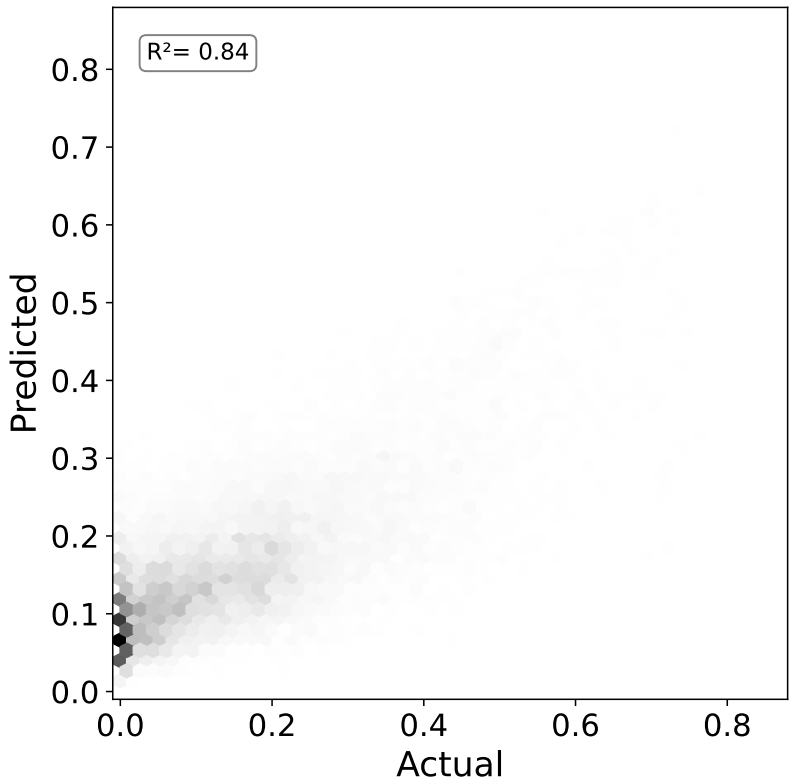
AU-test



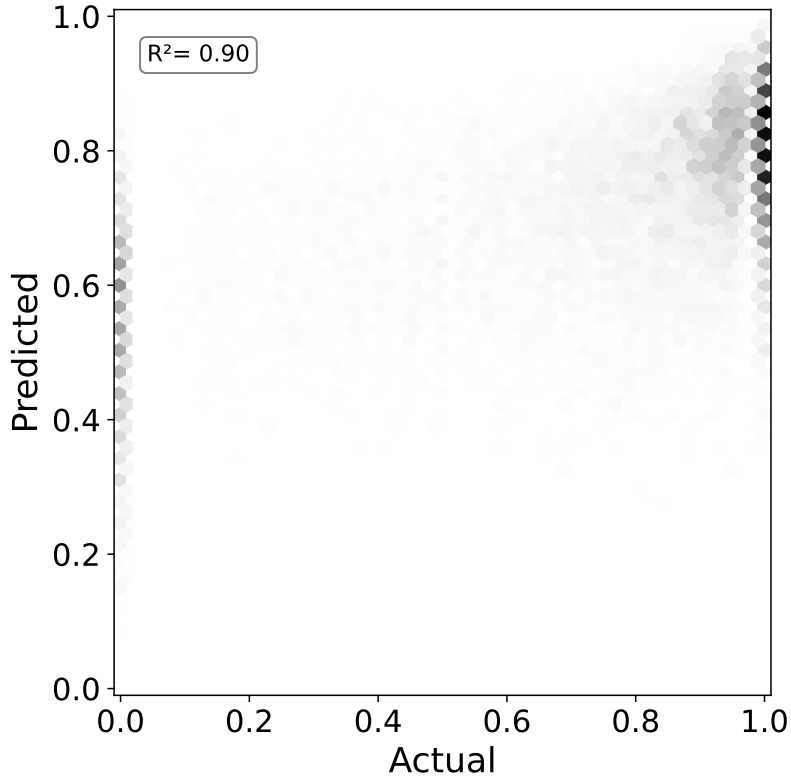
Supplement

Regression

RF distance

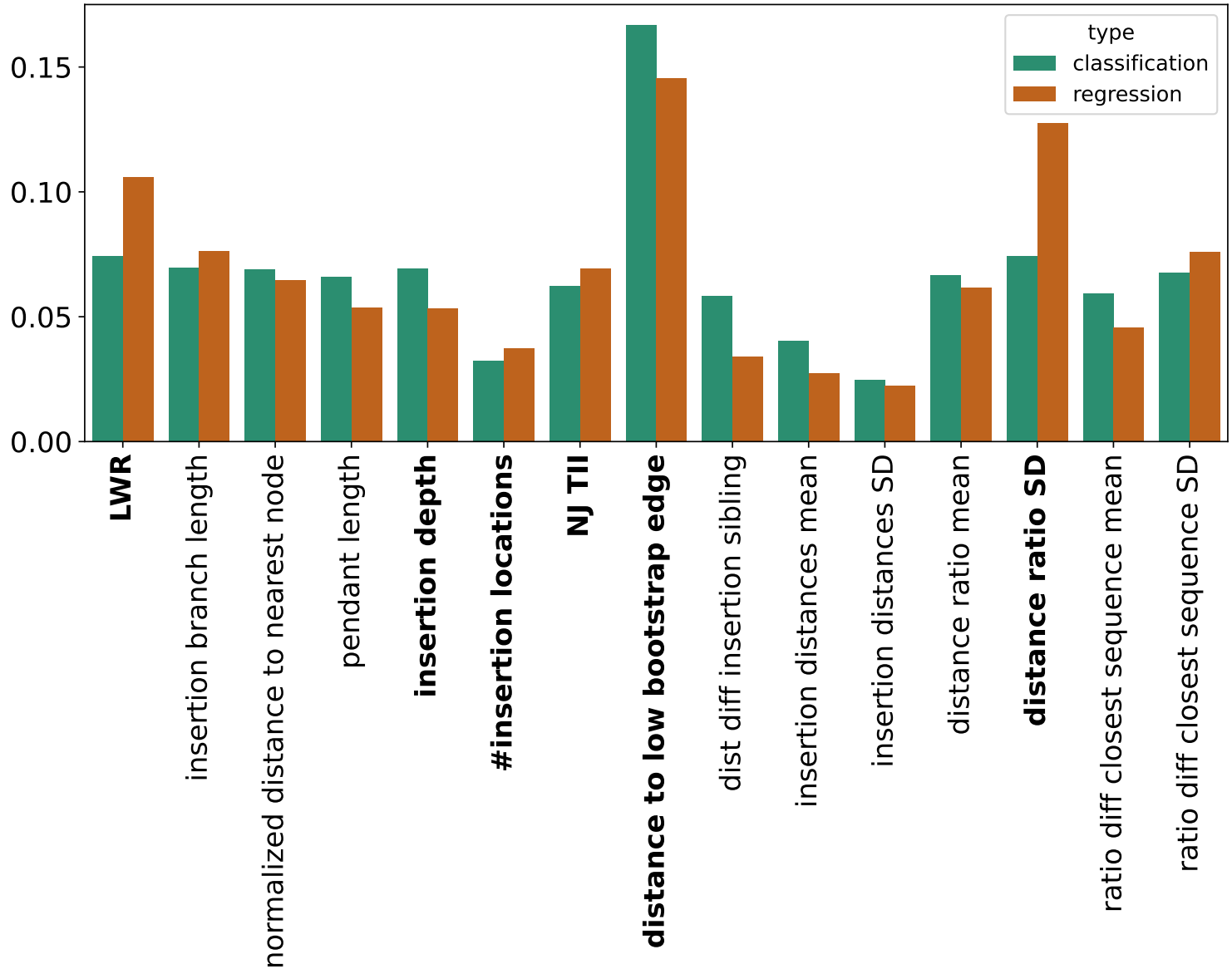


RF radius

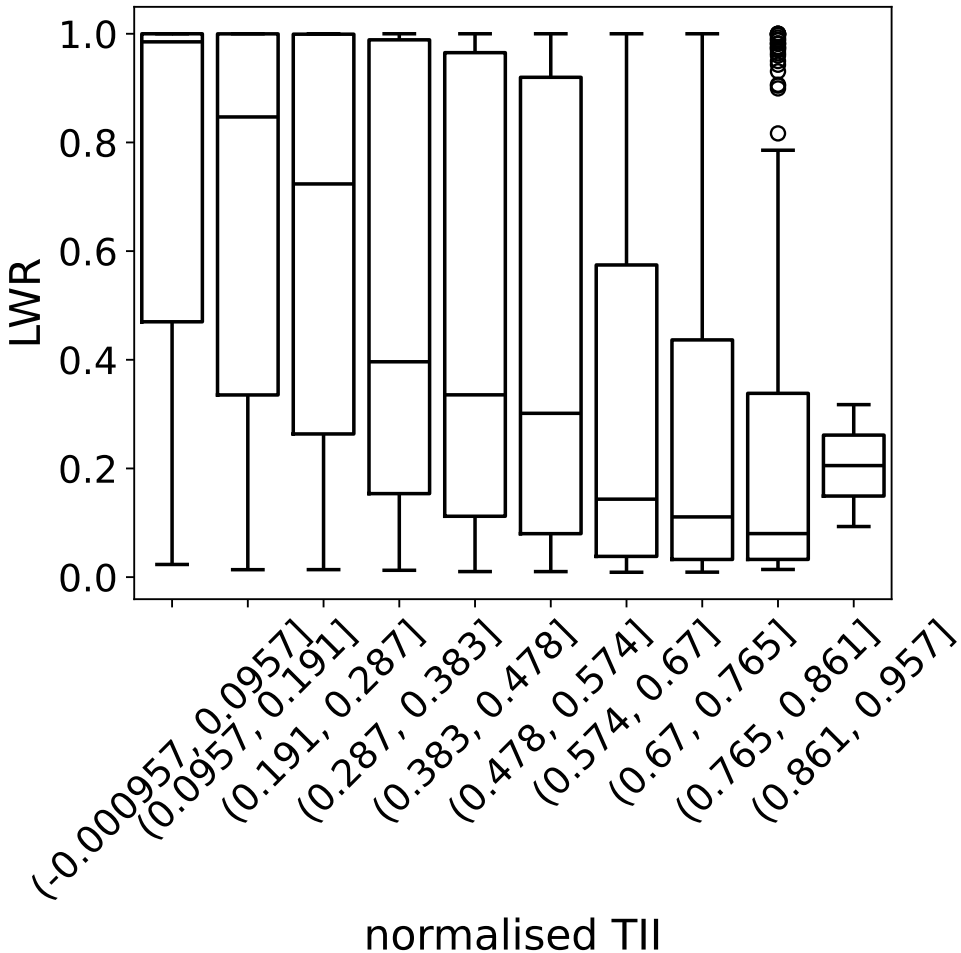


Supplement

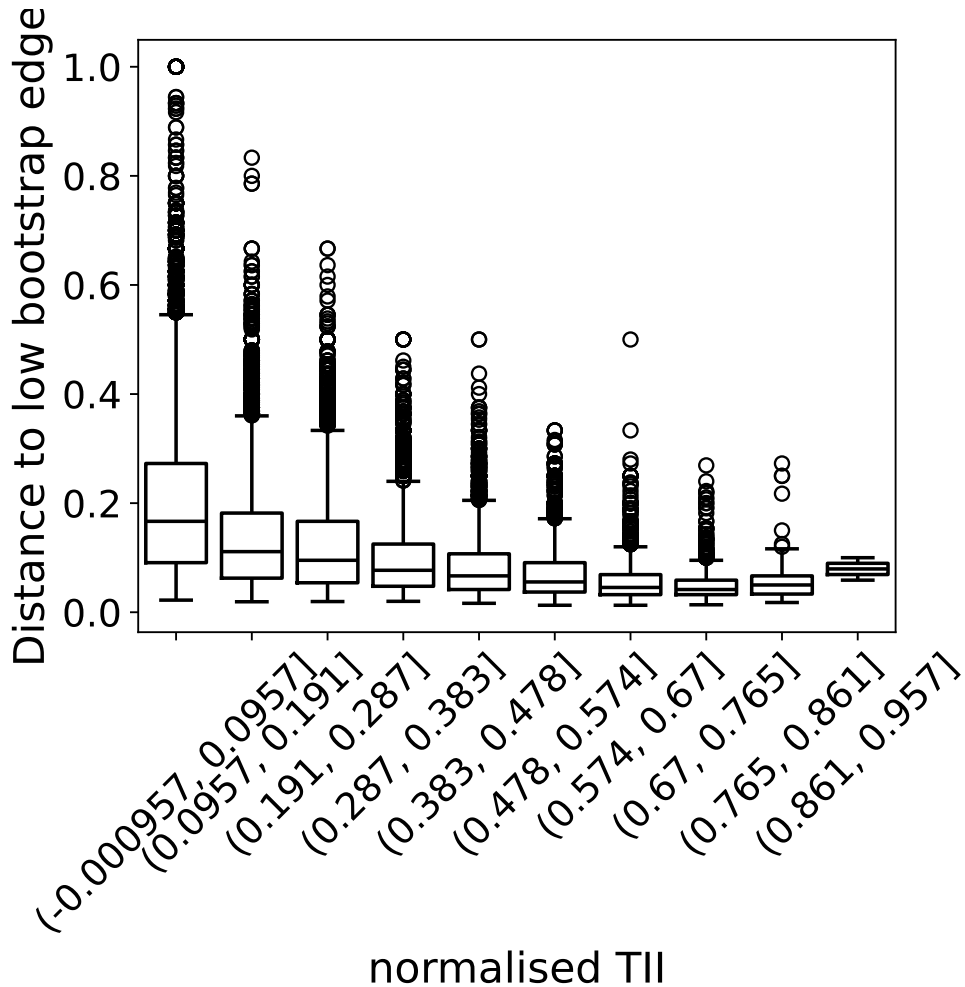
TII Feature importances



LWR vs TII


LWR
1.0
0.8
0.6
0.4
0.2
0.0
(-0.000957, 0.0957]
(0.0957, 0.191]
(0.191, 0.287]
(0.287, 0.383]
(0.383, 0.478]
(0.478, 0.574]
(0.574, 0.67]
(0.67, 0.765]
(0.765, 0.861]
(0.861, 0.957]
normalised TII

Distance to low bootstrap vs TII



Tree distance ratio vs TII