2.3.2 Calculating the CI and RI in Mesquite

Calculating the CI and RI in Mesquite

Now let's work through an example using the free program Mesquite (Maddison and Maddison 2006).

First, download the example file:

RI_CI_Example2.3.nex

. This file includes the characters and trees from Lycett et al. (2010). This is a dataset of chimpanzee behavioral variations that many researchers think are cultural, which in this case means the behaviors diffuse between individuals through a social learning process. The cultural hypothesis implies that variation in the behaviors are not determined by genetic inheritance. "Absent" behaviors are scored as state 0, "rare" behaviors as state 1, "regular" behaviors as state 2, and missing data as question marks ("?") for three communities in this dataset. Lycett et al. (2010) used the RI to test the hypothesis that the behavioral variations are consistent with genetic inheritance. We will work through their analyses and interpretations while learning how to calculate the CI and RI.

To start, open the file in Mesquite ("Open File ..." in the File menu in the top drop down bar). Mesquite accepts standard NEXUS files. Click on the 'Character Matrix' link to the left of the window to view the data.

The authors first inferred a parsimony tree from the behavioral data. Click on the 'Tree Window' in Mesquite. You should see a tree in this window with 11 taxa named "MP all taxa." Click on the arrow in the upper left of this window labeled "Tree #" to see tree 2. It contains 8 taxa and is named "MP central-East taxa" (the tree name is seen in the lower left of the screen).

These are the most parsimonious trees that Lycett et al. (2010) inferred from the behavioral data in the character matrix. (The tree for all taxa is actually only one of 3 equally most parsimonious trees, but that does not affect this example because the RI will be identical on all trees that are equally parsimonious.) The smaller tree with 8 taxa includes social groups of chimpanzees that are thought to belong to the same subspecies, and thus are expected to exhibit greater genetic similarity. Lycett et al. (2010) reasoned that if the behaviors were inherited genetically, then the RI should be higher for the larger tree that includes social groups from different subspecies because greater genetic structuring should create more tree-like signal in genetically inherited traits (in this case, behaviors).

Select the tree with 8 taxa in your tree window. Now click on "Analysis" in the main drop down bar. From there, select "Values for Current Tree ...". A new window should open. In this new menu, under "Tree value using character matrix", select "Retention Index for Matrix" and click "OK." A legend outlined in blue should appear within your tree window. If you click and drag the border you can move this legend to a convenient place on your screen. The legend shows the ensemble RI under Mesquite's default model for parsimony characters, which allows 0 <--> 2 changes to be equivalent to 0 <--> 1 changes.

Because the states in this case are ordinal - they indicate the prevalence of a character - Lycett et al. (2010) made an assumption common in parsimony analyses to 'order' the character states. This makes changes between 0 and 2 cost 2 parsimony steps while changes between adjacent states only cost one step. This effectively renders state 1 an evolutionary intermediate between 0 and 2. You can implement the ordered parsimony model by clicking the "Legend" option from the main drop down bar. Then select "Source of parsimony models" --> "Stored Parsimony Models". A new window should pop up entitled "Choose model." Pick the "Ordered" option and click "OK."

You can see the same ensemble RI value that Lycett et al. (2010) obtained for the 8-taxon tree, 0.68. Use the blue arrows to switch to the 11-taxon tree. Notice that the ensemble RI in the legend changes to reflect the value for the tree that is currently displayed in the tree window. The 11-taxon tree has a lower RI of 0.56.

Does this result mean that trait variation is inconsistent with genetic inheritance? Maybe. While the result does indicate less tree-like signal for the tree with greater genetic differentiation, we do not know how much of a change in RI is "significant." In other words, we do not know how much these statistics change just due to sampling different species, because we do not know their null distributions. Thus, one might say that the RI alone does not reject or support any particular hypothesis; it just measures the extent to which the characters violate the parsimony principle, which is why Lycett et al. (2010) also included a number of other more statistical tests of their hypothesis. Additionally, as pointed out by Nunn et al. (2010), the RI might go down with the inclusion of more divergent taxa if the traits are genetically inherited but evolve rapidly. This could occur if the traits are tied to the frequencies of just a few alleles - but it is probably an unlikely scenario, given other knowledge of chimpanzee behavioral development.

Using the same procedures, you can also calculate the ensemble CI for these data. After selecting "Analysis" and "Values for Current Tree," simply select "Consistency Index for Matrix."

It is also possible to calculate the RI or CI for individual characters. In the box, simply choose "Retention Index for Character," and the program will provide you with a list of characters. Choose a character that interests you, and repeat for the consistency index.

References

Lycett, S. J., M. Collard, and W. C. McGrew. 2010. Are behavioral differences among wild chimpanzee communities genetic or cultural? An assessment using tool-use data and phylogenetic methods. American Journal of Physical Anthropology. *142:461-467.*

Maddison, W. P., and D.R. Maddison. 2011. Mesquite: a modular system for evolutionary analysis. Version 2.75, [http://mesquite project.org](http://mesquiteproject.org).

Nunn, C. L., C. Arnold, L. J. Matthews, and M. Borgerhoff Mulder. 2010. Simulating Trait Evolution for Cross-Cultural Comparison. Philosophical Transactions of the Royal Society. *365:3807-3819.*

Contributed by Luke Matthews, Charles Nunn and Christian Arnold