

Phylogenetic Analysis of Molecular Data (Botany 563)

Computer Lab 04: Congruence and hypothesis tests

Learning objectives:

- Learn to assess data partition homogeneity using PAUP*
- Learn to use and implement Templeton tests

Study case: Figs and wasps: a tale of coevolution? (based on Weiblen 2001. Syst.Biol.50(2):243–267)

Figs (*Ficus* spp., Moraceae) and their pollinating wasps (Agaonidae, Chalcidoidea) constitute one of the most tightly integrated pollination mutualisms known. Figs are characterized by extraordinarily high global and local species diversity.

Functionally dioecious fig species have two kinds of plants: “gall figs” with staminate and pistillate flowers with short styles, and “seed figs”, with pistillate flowers only, which have long styles. Female wasps push through the fig opening, lay eggs and pollinate pistillate flowers. In seed figs the wasp larvae will fail to reach resources to survive because styles are long and wasp ovipositors tend to be short. Functionally dioecious fig pollination seems therefore paradoxical because wasps pollinating seed figs are denied their only opportunity to reproduce, and gall fig pollinators would be more fit. Avoidance of seed figs has not been observed. The existence of at least 350 functionally dioecious species suggests that this mutualism is evolutionarily stable.

With regard to pollinating fig wasps, it has been argued that morphology is perhaps more indicative of the functional constraints imposed by host associations than of phylogenetic relationships. Wasp ovipositors that are shorter than the wasp abdomen tend to be present in wasps that pollinate dioecious but not monoecious figs (dioecious figs do not form a clade). We will assess if wasps that pollinate functionally dioecious figs form a clade.

Results from morphological studies and the notion of one-to-one species specificity of the interaction led to the proposal of strict-sense coevolution and tight cospeciation between the two groups. Such dynamics would lead to phylogenies that “mirror” each other. Assuming that the fig section *Sycocarpus*2 is a well supported clade, under the one-to-one strict speciation model, the wasps that pollinate this group should also form a clade.

We want to address the following two main points:

1. It has been suggested that wasp morphology associated with functionally dioecious pollination (traits like ovipositor length) is indicative of phylogeny. Under such model, wasps that pollinate functionally dioecious figs would form a clade even if dioecious figs do not form a clade. Do the data support/reject that wasps pollinating functionally dioecious figs are a monophyletic group?
2. Lets focus on the fig group *Sycocarpus*2 and its wasp pollinators. Assuming that the fig section *Sycocarpus*2 is a well supported clade, under the one-to-one strict cospeciation model, the wasps that pollinate this group should also form a clade. Do the data support/reject a strict coevolutionary mode of speciation?

Dataset: **wasps.nex**

How many named character sets does this data have? _____

Which are they? _____

Task 1. Is there evidence that the data partitions are non-homogeneous? Do an ILD test.

a. What is the null hypothesis you are testing?

b. Edit the PAUP block to insert a character partition by gene using the character sets already defined, and the following format.

```
charpartition threegenes = gene1:gene1, gene2:gene2, gene3:gene3;
```

Save and execute the file.

c. Data on both ends (beginning and end of dataset, as well as an insertion) were very hard to align, and there the assessment of homology is questionable. Exclude these characters.

a. What characters did you exclude: _____

Remember to re-exclude these characters if you re-execute your dataset!!!

d. After executing the file, select ANALYSIS>PARTITION HOMOGENEITY TEST. Choose the character partition you just named (e.g., “threegenes”) and then select the heuristic search option and 100 replicates. I suggest doing SIMPLE, HOLD=1; maxtrees=100, TBR searches under parsimony.

e. Look at the output.

How many random partitions had a higher combined tree length than the original data matrix? _____

What is the p-value for the hypothesis that the partitions are drawn from the same underlying character population? _____

Do you think the data conflict?

These three partitions correspond to three mitochondrial genes. How do explain your findings in the light of this information?

Optional Is there evidence that the 1st, 2nd and 3rd codon positions represent different populations of characters for this dataset?

Report what you did including relevant p-value(s).

Task 2. Do a thorough heuristic search on the combined data. Report the following:

Your search parameters: _____

Results (number of MPTs, score): _____

Save your tree to a file.

Do your results suggest that wasps pollinating functional dioecious figs are a monophyletic group? Scoring of fig groups and their breeding systems for the wasp taxa we are working with is available in the dataset **wasps.morpho.sm.nex**. You might find it useful to look at in MacClade, with the tree(s) you found. *Apocryptophagus spinitarus* is an appropriate outgroup.

Can you **reject** the hypothesis that wasps pollinating functional dioecious figs are a monophyletic group based on your analysis so far? Why or why not?

Task 3. Is there evidence to statistically support/reject the hypothesis of monophyly of wasps that pollinate functional dioecious figs? Assess this using a Templeton test.

- a. To compare competing hypothesis you will need to do a search under the constraint of monophyly (or non-monophyly, depending on your results in the previous section). Report the constraint you used and explain why you used it.

Report the number and score of trees you found: _____

- b. Get both the optimal overall tree and the best tree under the constrain into memory, using TREES>GET TREES. To get the second set of trees in you need to select OPTIONS and ask it to retain trees already in memory and add the trees from file (make the entire Venn diagram dark).
- c. Go to TREES>SCORES>PARSIMONY and select “show only characters whose length varies” and “length” (you can deselect “show overall tree lengths...”). Hit okay and look at the output. Are there any characters that require more steps under the optimal tree than under the constrained tree?

- d. Do a Templeton and winning-sites tests: Go to TREES>SCORES>PARSIMONY again and select “show overall tree lengths...,” character output = none, and “nonparametric.” Look at the resulting table and report the following p-values:
Templeton test _____ Winning-sites _____
- e. Can you reject the hypothesis of monophyly of wasps that pollinate functional dioecious figs?

Task 4. Is there evidence to statistically support/reject that wasp pollinators of the *Sycocarpus*2 fig group form a clade or is it necessary to invoke host shifts among wasp species to explain the data? Do a Templeton test. Now you know what to do.

Report the constraint(s) you used.

Why did you use this constrain(s)?

What are the relevant p-values?

Can you reject host shifts among wasp species? What do your results say about the one-to-one strict cospeciation model in the fig-wasp association?