

## Host-to-Parasite Gene Transfer in Flowering Plants: Phylogenetic Evidence from Malpighiales

Charles C. Davis<sup>1\*</sup> and Kenneth J. Wurdack<sup>2</sup>

<sup>1</sup>Ecology and Evolutionary Biology, University of Michigan Herbarium, 3600 Varsity Drive, Ann Arbor, MI 48108–2287, USA.

<sup>2</sup>Department of Botany and Laboratories of Analytical Biology, Smithsonian Institution, 4210 Silver Hill Road, Suitland, MD 20746, USA.

\*To whom correspondence should be addressed. E-mail: chdavis@umich.edu.

**Horizontal gene transfer (HGT) between sexually unrelated species has recently been documented for higher plants, but mechanistic explanations for HGTs have remained speculative. We show that a parasitic relationship may facilitate HGT between flowering plants. The endophytic parasites Rafflesiaceae are placed in the diverse order Malpighiales. Our multigene phylogenetic analyses of Malpighiales show that mitochondrial (*matR*) and nuclear loci (18S rDNA and *PHYC*) place Rafflesiaceae in Malpighiales, perhaps near Ochnaceae/Clusiaceae. Mitochondrial *nad1B-C*, however, groups them within Vitaceae, sister to their obligate host, *Tetrastigma*. These discordant phylogenetic hypotheses strongly suggest that part of the mitochondrial genome in Rafflesiaceae was acquired via HGT from their hosts.**

Malpighiales are one of the most diverse clades of flowering plants uncovered in recent phylogenetic analyses. The order comprises 27 families (*sensu lato*) previously assigned to 13 different orders (2), including more than 16,000 species spanning tremendous morphological and ecological diversity (3). A recent surprising addition to Malpighiales are the endophytic holoparasites Rafflesiaceae (4), which lack leaves, stems, and roots, and rely entirely on their host plants, species of *Tetrastigma* (Vitaceae), for their nutrition. Despite their extreme vegetative reduction, they are unmistakable in flower, producing the largest flowers in the world, which mimic rotting flesh—an enticement to the carrion flies that pollinate them (5).

Barkman *et al.* (4) used mitochondrial (mt) *matR* sequences to place Rafflesiaceae firmly with Malpighiales (100% bootstrap percentage; BP). Their use of an (single) mt gene was appropriate in a family that has resisted placement with standard genetic loci. To further examine this placement, we obtained sequences representing all families of Malpighiales, all genera of Rafflesiaceae, and numerous basal eudicots for four loci from the mt and nuclear genomes (6). Low-copy nuclear genes are an underutilized resource for resolving the placement of problematic taxa, and

phytochrome C (*PHYC*), as used here, has been useful for revealing relationships within Malpighiales (7).

Our phylogenetic analyses are summarized in Fig. 1 (8). The tree from the *matR* and nuclear loci firmly (100% BP) place Rafflesiaceae within Malpighiales. In contrast, the mitochondrial locus *nad1B-C*, suggests that Rafflesiaceae are not members of Malpighiales, but belong (100% BP) in Vitaceae as sister to their host *Tetrastigma*. Each of these mutually exclusive hypotheses cannot be attributable to contamination (9) and receive strong support from parsimony analyses and from alternative topology tests.

Which of these conflicting hypotheses reflect the true species affinities of Rafflesiaceae? Vitaceae possess several synapomorphies that are rare among angiosperms, including sieve-tube plastids with starch and protein inclusions, pearl glands, stamens opposite the petals, and seeds with a cordlike raphe. If Rafflesiaceae were embedded in Vitaceae, as suggested by *nad1B-C*, we would expect species to possess at least some of these characters, but they do not (2, 3). A definitive malpighialean sister-group for Rafflesiaceae is unclear given our data. However, among the closest relatives (*i.e.*, Ochnaceae and Clusiaceae *sensu lato*) suggested in the combined analysis (10) they share tenuinucellate ovules (among mostly crassinucellate relatives), and staminal fusion with Rafflesiaceae (2, 3).

The position of Rafflesiaceae based on *nad1B-C* provides a new example of horizontal gene transfer. If *nad1B-C* were vertically transmitted, as we believe to be the case for the other loci, we would expect Rafflesiaceae to group with Malpighiales. Instead, phylogenetic evidence from *nad1B-C* suggests that part of the mitochondrial genome in Rafflesiaceae originated from their hosts, *Tetrastigma* (either stem or crown group members), and was horizontally transferred to these obligate parasites. A similar HGT of *nad1B-C* was recently reported (11) in seed plants involving a transfer from an asterid to *Gnetum*. And, Bergthorsson *et al.* (12) documented several instances of mitochondrial HGT between distantly related angiosperm groups.

The underlying mechanism for HGT between sexually unrelated plants, however, has been speculative, citing various pathogens as primary vector agents (11, 12). Our study documents a case in which there is no need to propose an intermediary vector for HGT. In these plants the transfer appears to have been facilitated by the intimate nature of the association between the host and the endophytic parasite, which lives its whole vegetative life as "an almost mycelial haustorial system," "ramifying and anastomosing throughout the [tissues of the] host" (13). This pattern may be an important mechanism by which parasites assemble their genetic architecture, and additional cases of HGT should be sought among other endophytic parasites and their hosts. It will also not be surprising if reciprocal genetic transfers are found to have occurred, from parasite to host.

### References and Notes

1. APGII, *Bot. J. Linn. Soc.* **141**, 399 (2003).
2. A. Cronquist, *An Integrated System of Classification of Flowering Plants* (Columbia Univ. Press, New York, 1981), p. 1152.
3. P. F. Stevens, *Angiosperm Phylogeny Website*. ed. 4 (2001 onwards), <http://www.mobot.org/MOBOT/research/APweb>.
4. T. J. Barkman, S.-K. Lim, K. Mat Salleh, J. Nais, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 787 (2004).
5. R. S. Beaman, P. J. Decker, J. H. Beaman, *Am. J. Bot.* **75**, 1148 (1988).
6. For primer and GenBank information see Supporting Online Materials (SOM).
7. C. C. Davis, M. W. Chase, *Am. J. Bot.* **91**, 262 (2004).
8. For information on data sets, phylogenetic analyses, and complete annotated figures see SOM. We were unable to sample *PHYC* from all families of Malpighiales because it is probably absent from some clades (7). Hence, we analyzed these data in combination with 18S to ensure that all families were represented for the nuclear genome. We also examined the placement of Rafflesiaceae with 18S in two ways. First, we excluded the most divergent domains from the analysis, V2 and V4 [D. L. Nickrent, E. M. Starr, *J. Mol. Evol.* **39**, 62 (1994); (322 of 1813 base pairs)], across all taxa. Second, we treated these domains as missing data only for Rafflesiaceae. Both approaches yielded congruent topologies, with the latter (i.e., shown here) being much better resolved. The spurious placement of Rafflesiaceae as sister to all angiosperms using 18S (4) may be attributable to high divergence in these small domains, and suggests that 18S may not be as useless for placing problematic taxa as previously suggested (4).
9. We have taken several precautions to avoid/detect contamination. First, our results were independently corroborated in the laboratories of each author. Second, Rafflesiaceae data were acquired before starting any work

on Vitaceae. Third, if our DNA were cross-contaminated we would not expect such strongly conflicting results regarding the placement of Rafflesiaceae given the same DNA. Nor would we expect such high degree of sequence divergence of *nad1B-C* between Rafflesiaceae and *Tetrastigma* (or other Vitaceae we sampled). If contamination occurred, we would expect sequences to be nearly identical to other sampled Vitaceae, especially, given the relatively low amount of sequence divergence between all accessions of Vitaceae for *nad1B-C* (0.51–0.95% sequence divergence). Not only is *nad1B-C* divergent within Rafflesiaceae (6.2%), but they are also highly divergent from other, phylogenetically diverse Vitaceae [M. J. Ingrouille *et al.*, *Bot. J. Linn. Soc.* **138**, 421 (2002)] included here (2.5–3.3%).

10. Taxa in the combined analysis of 18S, *PHYC*, and *matR* were included if they were sampled for *matR* plus at least one of the two nuclear loci. We believe this analysis to represent the best estimate of Malpighiales phylogeny. Similar approaches combining multiple genes have provided powerful insights into angiosperm phylogeny where single gene studies have failed [e.g., Y.-L. Qiu *et al.*, *Nature* **402**, 404 (1999)].
11. H. Won, S. S. Renner, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 10824 (2003).
12. U. Bergthorsson, K. L. Adams, B. Thomason, J. D. Palmer, *Nature* **424**, 197 (2003).
13. J. Kuijt, *The Biology of Parasitic Flowering Plants* (Univ. of California Press, Berkeley & Los Angeles, 1969), pp. 246.
14. We thank W. Anderson, D. Boufford, S. Mathews, Y.-L. Qiu, and P. Stevens. C. Davis was supported by a grant from the University of Michigan, and by the Michigan Society of Fellows. This paper is dedicated to V. Morrison.

### Supporting Online Material

[www.sciencemag.org/cgi/content/full/1100671/DC1](http://www.sciencemag.org/cgi/content/full/1100671/DC1)

Materials and Methods

Figs. S1 to S5

References

25 May 2004; accepted 6 July 2004

Published online 15 July 2004; 10.1126/science.1100671

Include this information when citing this paper.

**Fig. 1.** Two conflicting hypotheses for the phylogenetic placement of Rafflesiaceae. (A) The strict consensus of 136 angiosperms for combined mt *matR* and nuclear (*PHYC* and ribosomal 18S) data showing a well-supported (100% BP) Malpighiales clade (in blue), which includes all members of the order sensu APG II (1) plus Rafflesiaceae (in red; *Rafflesia*, *Rhizanthus*, and *Sapria*). (B) The strict consensus

of 147 angiosperms for mt *nad1B-C* (i.e., the *nad1* intron 2 and part of the adjacent exons b and c) showing a well-supported (100% BP) Malpighiales clade, which includes all members of the order except Rafflesiaceae. Rafflesiaceae (*Rafflesia* and *Sapria*) are strongly placed (100% BP) in the basal eudicot family Vitaceae (in yellow) as sister to their host genus, *Tetrastigma*. Dashed line is hypothesized host-parasite horizontal gene transfer.

Scienceexpress

**A** *matR*, *PHYC*, *18S*

**B** *nad1B-C*

