Massive horizontal transfer of mitochondrial genes from diverse land plant donors to the basal angiosperm *Amborella*

Ulfar Bergthorsson[†], Aaron O. Richardson, Gregory J. Young, Leslie R. Goertzen[‡], and Jeffrey D. Palmer[§]

Department of Biology, Indiana University, Bloomington, IN 47405-3700

Contributed by Jeffrey D. Palmer, November 9, 2004

Several plants are known to have acquired a single mitochondrial gene by horizontal gene transfer (HGT), but whether these or any other plants have acquired many foreign genes is entirely unclear. To address this question, we focused on *Amborella trichopoda*, because it was already known to possess one horizontally acquired gene and because it was found in preliminary analyses to contain several more. We comprehensively sequenced the mitochondrial protein gene set of *Amborella*, sequenced a variable number of mitochondrial genes from 28 other diverse land plants, and conducted phylogenetic analyses of these sequences plus those already available, including the five sequenced mitochondrial genomes of angiosperms. Results indicate that *Amborella* has acquired one or more copies of 20 of its 31 known mitochondrial protein genes from other land plants, for a total of 26 foreign genes, whereas no evidence for HGT was found in the five sequenced genomes. Most of the *Amborella* transfers are from other angiosperms (especially eudicots), whereas others are from nonangiosperms, including six striking cases of transfer from (at least three different) moss donors. Most of the transferred genes are intact, consistent with functionality and/or recency of transfer. *Amborella* mtDNA has sustained proportionately more HGT than any other eukaryotic, or perhaps even prokaryotic, genome yet examined.

Genome sequencing has revealed that horizontal gene transfer (HGT), the transfer of genes between nonmating species, is remarkably common and important in bacterial evolution (1). The current picture of HGT in eukaryotes is decidedly mixed. Other than the special case of mobile genetic elements (and plant mitochondrial genomes, see below), HGT is largely unknown in multicellular eukaryotes but is more or less common in diverse groups of unicellular protists, which contain several to many genes derived by HGT from both prokaryotes and other protists (2).

Recent studies indicate that plant mtDNAs are unusually active in HGT relative to all other organellar genomes and nuclear genomes of multicellular eukaryotes. Four papers (3–6) have reported a total of nine cases of mitochondrial HGT within seed plants. Three transfers involve parasitic angiosperms as putative donors or recipients and implicate direct, plant-to-plant transfer of DNA as one mechanism of HGT (5, 6). Each of the nine transfers involves a different set of recipient plants. For this reason, and because only a few mitochondrial genes have been scrutinized for potential HGT in these or any other plants, it is unclear whether these cases are singular exceptions in each genome or whether they are harbingers of perhaps massive mitochondrial HGT in certain plants.

To address this uncertainty, we have assessed the origin and history of the mitochondrial protein gene set of *Amborella trichopoda* and the five angiosperms whose mtDNAs have been sequenced. *Amborella* was chosen because it was already known to contain one foreign gene (3) and because preliminary studies suggested it might be unusually rich in HGT. We show that *Amborella* mtDNA has sustained remarkably massive HGT, whereas the five sequenced mtDNAs show no evidence of HGT.

Materials and Methods

We used primers for conserved regions of angiosperm mitochondrial genes in an attempt to PCR-amplify and sequence all mitochondrial protein genes from *A. trichopoda* (primer sequences available on request). Many *Amborella* reactions produced multiple bands, heterogeneous sequence, or unreadable sequence; these were cloned, and multiple (usually eight) clones were sequenced. This process yielded portions of 27 genes. We then used PCR to amplify and sequence as many of these 27 genes as possible, plus the four genes already sequenced from *Amborella* mtDNA, from 13 other angiosperms (see Fig. 5, which is published as supporting information on the PNAS web site, for taxa and sources) and three gymnosperms. For each of these plants, we carried out 80 PCRs with conserved mitochondrial primers. Selected genes were amplified and sequenced from 12 additional nonangiosperms. PCR was performed under the following conditions: 95°C for 2 min, 35 cycles of 95°C for 30 s, 55° or 52°C for 30 s, 72°C for 2 min, and 72°C for 5 min. PCR products were cleaned by using 2 μl of ExoSAP-IT (United States Biochemical). Sequences were generated by using an ABI 3730 (Applied Biosystems). Sequence traces were assembled and trimmed by using CODONCODE ALIGNER 1.3.2.

Sequences were aligned by using either BIOEDIT or SE-AL V2.0A11 (alignments available on request). Regions containing primers, poor alignment, or only a few taxa, as well as all sites subject to RNA editing in either *Arabidopsis*/*Brassica* or *Oryza*/*Zea*, were excluded from phylogenetic analyses. Analyses used PAUP* 4.0B10 within an automated script (courtesy of D. W. Rice, Indiana University). A starting topology was generated with maximum parsimony, from which the transition/transversion ratio and gamma shape parameter were estimated. A maximum-likelihood (ML) tree was built by using these parameters, the HKY85 model (7), four rate categories, and empirically determined base frequencies. If the ML and parsimony trees differed in topology, a new ML tree was built, using parameters from the preceding ML tree, and this process was repeated until a stable topology was obtained.

The Shimodaira–Hasegawa (SH) test (8) was used to assess whether phylogenetically anomalous gene placements suggestive of HGT are significantly favored over the hypothesis of strictly vertical transmission. This test assigns a *P* value to the difference in likelihood between the best ML tree found (as shown in all of our figures) and that ML tree, based on the same data set, in which the *Amborella* gene in question has been constrained to fit

Abbreviations: HGT, horizontal gene transfer; ML, maximum likelihood; SH, Shimodaira–Hasegawa.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AY831968–AY832318).

[†]Present address: Department of Biology, University of New Mexico, Albuquerque, NM 87131.

[‡]Present address: Department of Biological Sciences, Auburn University, Auburn, AL 36849.

[§]To whom correspondence should be addressed. E-mail: jpalmer@bio.indiana.edu.

Table 1. Horizontally acquired mitochondrial genes in *Amborella*

Gene	No. of copies		HGT donor	SH test	Gene length	Gene integrity
	Total	HGT				
cox2	4	3	Moss	<0.001	266	I
			Eudicot	NS	266	I
			Eudicot	NS	311	I
nad2	2	2	Moss	<0.001	433	I
			Eudicot	NS	686	Ψ
nad3	2	1	Moss	<0.001	341	I
nad4	3	2	Moss	<0.001	537	I
			Eudicot	NS	358	I
nad5	3	2	Moss	<0.001	1062	I
			Angiosperm	0.025	601	I
nad6	2	1	Bryophyte	<0.001	539	I
nad7	3	2	Moss	<0.001	1090	Ψ
			Eudicot	NS	1080	I
atp1	2	1	Eudicot	0.001	1254	I
atp4	2	1	Eudicot	NS	473	I
atp6	2	1	Eudicot	NS	389	Ψ
atp8	2	1	Eudicot	0.008	416	I
atp9	2	1	Angiosperm	NS	181	I
ccmB	2	1	Eudicot	NS	622	Ψ
ccmC	2	1	Eudicot	0.03	670	Ψ
$ccmF_{N1}$	2	1	Eudicot	0.004	142	I
cox3	2	1	Angiosperm	NS	393	I
nad1	2	1	Eudicot	<0.001	1285	I
rpl16	2	1	Eudicot	NS	467	Ψ
rps19	2	1	Eudicot	0.003	223	Ψ
sdh4	2	1	Eudicot	NS	439	Ψ

Protein genes present in only one, putatively vertical copy in *Amborella* are $ccmF_{N2}$, *cob*, *cox1*, *matR*, *nad9*, *rpl2*, *rps1*, *rps2*, *rps4*, *rps7*, and *rps13* (Fig. 6, which is published as supporting information on the PNAS web site). Protein genes present ancestrally in angiosperm mtDNA (11), but not recovered from *Amborella* are *mtt2*, *nad4L*, *rpl5*, *rps3*, *rps10*, *rps11*, *rps12*, *rps14*, and *sdh3*. $P <$ 0.05 are given for passing the SH test (8) for origin via HGT (see *Materials and Methods*), with NS indicating not significant ($P > 0.05$). Gene length in nucleotides is given for the *Amborella* gene region used in phylogenetic analyses. I indicates an intact ORF, and Ψ indicates a pseudogene.

a vertical scenario of paralogy (duplication) by being placed as sister to its putatively vertically transmitted homolog.

All cases of suspected *Amborella* HGT from bryophyte donors and most cases from angiosperm donors were confirmed by obtaining the same sequence from multiple (3–5) independent preparations of *Amborella* DNA. These DNAs originated from material sent from four different sources. Two shipments of fresh leaves, received and DNA-extracted 18 months apart, came from the University of Santa Cruz Arboretum courtesy of Brett Hall. Silica-dried leaves were obtained from Doug Soltis (University of Florida, Gainesville), fresh leaves were obtained from the University of Massachusetts Greenhouse, Amherst, courtesy of Teddi Bloniarz, and *Amborella* DNA was received from Yin-Long Qiu (University of Michigan, Ann Arbor). Leaves were inspected carefully for any signs of epiphytic growth and other potential sources of biological contamination, in some cases under a dissecting microscope, and were thoroughly washed before DNA extraction. All attempts to confirm HGT by using alternative sources of *Amborella* DNA were successful, with PCR product ratios constant among DNA preps for those primers giving size-heterogeneous products. Further evidence against contamination or sample mix-up came from the pseudogene nature of nine cases of putative HGT (Table 1), i.e., contamination or mix-up is much more likely to result in artefactual isolation of intact, functional copies of a gene. Further verification was obtained for two cases of HGT by showing that cDNA sequences are identical to genomic sequences except for a few sites of RNA editing (ref. 3 and unpublished data).

Results

We took advantage of the generally very low substitution rate in plant mitochondrial genes (9, 10) and used a PCR approach to assess the extent of HGT in plant mtDNAs. A set of ≈100 pairs of primers was designed to PCR-amplify the entire set of 40 angiosperm mitochondrial protein genes (including introns) that were present in the last common ancestor of angiosperms (11). Pilot amplifications to assess primer efficacy involved three test plants. Plant mitochondrial genes are generally present once per genome, and rice and *Arabidopsis* routinely give a single PCR product of the expected size based on their known genome sequences (12, 13). But to our surprise, with many primer pairs *Amborella* gave either two distinct bands or a single broad band. Three of these mixed products were examined and found to consist of vertically and horizontally transmitted genes, similar to the *atp1* case already described for *Amborella* (3). Finding so much HGT among so few examined *Amborella* genes led us to focus on *Amborella*. We amplified and sequenced all readily isolated *Amborella* mitochondrial protein genes, taking care to sequence multiple clones for each *Amborella* gene whose PCR products showed either size or sequence heterogeneity. For most genes, too few homologs were available to enable meaningful phylogenetic analysis. We therefore chose 13 diverse angiosperms and three gymnosperms (as outgroups) and sequenced their genes from PCR products, setting aside complicated cases (of potential HGT) involving size or sequence heterogeneity. Where appropriate, we also sequenced selected genes from a few nonseed plants. Phylogenetic analyses included all of these genes, all relevant genes from the five sequenced angiosperm mitochondrial genomes (12–16), and selected other available sequences.

Of the 40 protein genes present in the ancestral angiosperm mitochondrial genome (11), 31 were recovered from *Amborella* (Table 1). Of these 31 genes, 20 showed what we interpret as reasonable to compelling evidence for one or more cases of HGT. The strongest evidence for HGT comes from seven genes for which *Amborella* possess a bryophyte-like copy (Table 1 and Fig. 1). Six of these seven bryophyte-like genes are far more similar in sequence to homologs from mosses than to angiosperm homologs, and in phylogenetic analyses these all group with mosses with convincing support (Fig. 1 and data not shown). No moss sequences are available for *nad6*, which is more similar to the only bryophyte sequence available (from the liverwort *Marchantia*) than to angiosperm homologs (Fig. 1). Three of the six moss-derived genes (*cox2*, *nad5*, and *nad7*) probably were acquired from different lineages of moss donors (Fig. 2). For the other three genes, there is insufficient sampling of mosses (Fig. 1) to address this issue.

For five of the seven genes (Table 1) for which it contains a bryophyte-derived copy, *Amborella* also possesses a second (or in one case, a third) divergent copy that we interpret as being the product of HGT from other angiosperms. All of these putatively angiosperm-acquired genes group with eudicots, albeit with low bootstrap support in what are largely poorly resolved trees within angiosperms (Fig. 1). The eudicot-nested *Amborella nad5* gene is complicated because it actually groups as sister to monocots; this is the only one of the 31 genes for which monocots are placed within eudicots. This complexity notwithstanding, we emphasize that the SH test (see *Materials and Methods*) significantly favors ($P = 0.025$) a horizontal origin of this *Amborella* eudicot-like *nad5* gene.

A total of 13 *Amborella* genes show evidence of HGT from angiosperm donors only (Table 1). In each case, a pair of divergent gene copies was isolated, one of which is putatively

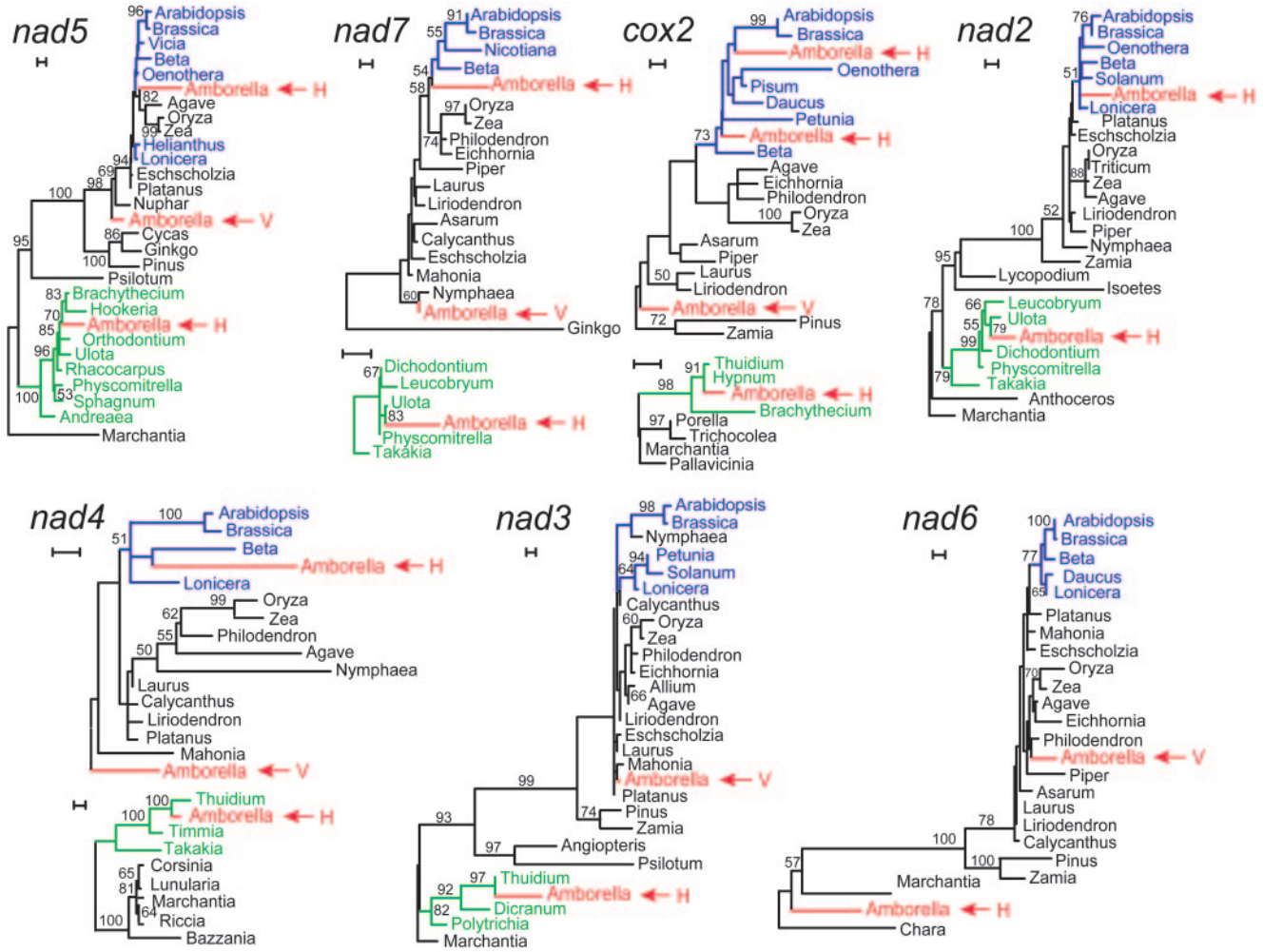

Fig. 1. Phylogenetic evidence for horizontal acquisition of genes from mosses and angiosperms in *Amborella*. Shown are ML trees. Bootstrap values (100 ML replicates) >50% are shown. H and V indicate *Amborella* genes of putatively horizontal and vertical transmission, respectively. *Amborella* genes are in red, core eudicot genes are in blue (basal eudicots commonly included are *Platanus*, *Eschscholzia*, and *Mahonia*), and moss genes are in green. Note that for *nad7*, *cox2*, and *nad4*, seed and nonseed plants were analyzed separately. Scale bars correspond to 0.01 substitutions per site.

vertical (see below) and the other putatively horizontal in origin. All but two of the 13 donors appear to be eudicots, albeit with varying levels of support. At one extreme are genes such as *nad1*, *atp1*, and *ccmF*$_{N1}$, with strong bootstrap support for being of eudicot origin (Fig. 3) and which also pass the SH test for being the product of HGT (Table 1). Five other genes (*ccmB*, *ccmC*, *atp8*, *atp4*, and *rps19*) show moderately good bootstrap support (70–80%) for being derived from eudicots (Fig. 3) and/or pass the SH test (Table 1). The other five genes show weak or no support for being eudicot-derived and fail the SH test (Fig. 3 and Table 1).

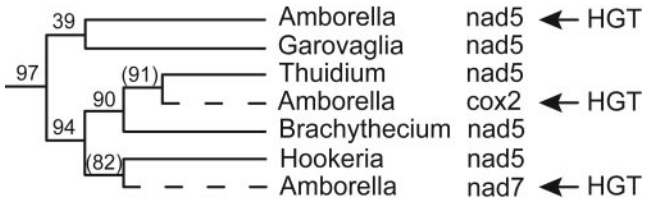

Fig. 2. *Amborella* acquired three genes from different moss donors. The solid parts of the cladogram and nonparenthetical bootstrap values are from the *nad5* intron phylogeny of Fig. 6. The dashed lines and other bootstrap values indicate the relationship to the indicated mosses of the moss-derived *cox2* and *nad7* genes of *Amborella*, as per the *cox2* gene tree of Fig. 1 and the *nad7* intron tree of Fig. 6.

The remaining 11 genes isolated from *Amborella* are, based on current data, present in one copy only and probably of vertical descent (Table 1 and Fig. 7, which is published as supporting information on the PNAS web site). Numerous phylogenetic studies (refs. 17–22 and references therein), in aggregate using many chloroplast genes (up to 61), several nuclear genes, and several mitochondrial genes that seem unafflicted by HGT, position *Amborella* as sister to all other angiosperms, either by itself or together with Nymphaeaceae. Each of the 11 *Amborella* singleton genes falls in a basal or near-basal position more or less consistent with organismal phylogeny under the hypothesis of strictly vertical descent, although in several cases the absence of any nonangiosperm outgroup limits the force of this conclusion (Fig. 7). Overall, there is no good reason to suspect a horizontal origin for any of these, so-far single-copy genes, but at the same time HGT cannot be ruled out either. Similarly, we conclude that *Amborella* probably contains a vertically transmitted copy of all but one of the 20 genes for which one or more cases of HGT have been invoked (Table 1 and Figs. 1 and 3). The exception is *nad2*, whose lone angiosperm-like copy falls within eudicots, albeit

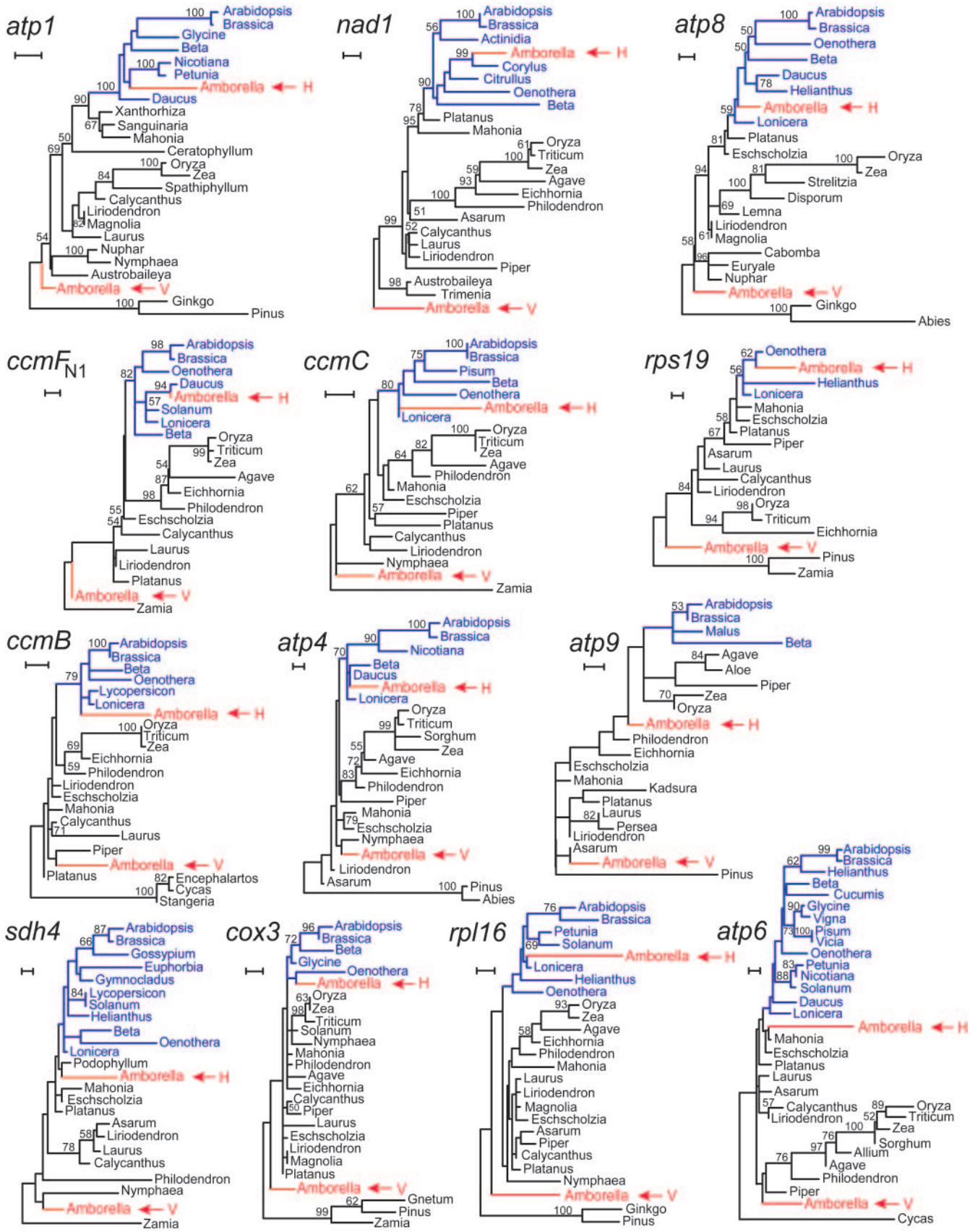

Fig. 3. Phylogenetic evidence for horizontal acquisition of 13 genes from angiosperms (mostly eudicots) in *Amborella*. Shown are ML trees. Bootstrap values (100 ML replicates) >50% are shown. H and V indicate *Amborella* genes of putatively horizontal and vertical transmission, respectively. *Amborella* genes are in red, and core eudicot genes are in blue (see Fig. 1 for basal eudicots). Scale bars correspond to 0.01 substitutions per site.

with only 51% bootstrap support, and which we interpret as most likely the product of HGT.

Where do the 26 transferred genes reside within *Amborella*? The chloroplast genome can be ruled out because it has been sequenced in *Amborella* (23) and does not contain any of the transferred genes (data not shown). The nuclear genome can be ruled out and a mitochondrial location be assigned with confidence for those two transferred genes (*atp1* and *atp8*) that are transcribed and subject to mitochondrial-characteristic RNA editing (ref. 3 and unpublished data). We favor a mitochondrial location for the 24 other horizontally transferred genes for three reasons. First, all six HGT cases (including the two above from *Amborella*) whose genomic provenance has been established are indeed located in mtDNA (ref. 3 and unpublished data). Second, nucleotide substitution rates are almost always far higher in the nucleus than in the mitochondrion in plants (9, 10), such that even relatively recent cases of functionally transferred mitochondrial genes present in the plant nucleus have extremely long branches in gene trees compared to their mitochondrial counterparts (e.g., refs. 10 and 24). Only the most divergent transferred genes in *Amborella* (e.g., *atp6* and *rpl16*) even approach this level of divergence, and most transferred genes show conventional mitochondrial-like branch lengths (Figs. 1 and 3). Third, the transferred genes amplify by PCR to roughly the same abundance as vertically transmitted *Amborella* homologs. Because mtDNA is typically present in hundreds of copies per cell, this result, even though the PCR was not carried out in a quantitative manner, suggests that the putatively horizontal and vertical copies reside in the same genome, that of the mitochondrion [conuclear localization is highly improbable, given that survey of 280 diverse angiosperms (11) showed that genes corresponding to 23 of the 26 transfers have never, or in one case very rarely, been lost from mtDNA].

Outside of *Amborella*, there is no convincing evidence for HGT in the 31 mitochondrial gene trees. Between 24 and 30 of the 31 genes are present, as intact, single-copy genes (except for identical duplications; see *Discussion*), in the five sequenced angiosperm mtDNAs (12–16). None of these genomes shows any evidence of HGT. The two grasses (*Zea* and *Oryza*) always either pair together or form part of a larger clade of grasses, the grasses in turn always group with at least one to all three of the most commonly sampled other monocots (*Eichhornia*, *Agave*, and *Philodendron*), and monocots are almost always monophyletic. Likewise, the two crucifers (*Arabidopsis* and *Brassica*) always pair as sisters, and these, together with *Beta*, always tree within core eudicots, as expected. Core eudicots are almost always monophyletic. Although there are numerous phylogenetic anomalies involving basal eudicots (usually *Platanus*, *Eschscholzia*, and *Mahonia*) and the magnoliids (usually *Piper*, *Laurus*, *Asarum*, *Calycanthus*, and *Liriodendron*), these are only poorly supported and are best attributed to poor resolution of generally slowly diverging sequences in these parts of angiosperm phylogeny.

Discussion

Massive HGT in *Amborella*. *A. trichopoda* is extraordinarily rich in horizontally acquired mitochondrial genes, possessing some 26 of them. In marked contrast, the five sequenced mtDNAs of angiosperms show no evidence for HGT. An important caveat, though, is that increased taxon sampling within the huge groups to which they belong (there are ≈70,000 species of monocots and 175,000 eudicots) may reveal phylogenetically local cases invisible to our sampling. The 13 other angiosperms sampled extensively in this study also show little evidence for HGT, but with still further caveat. For these plants, we deliberately ignored complicated PCR results that might, as with *Amborella*, reflect a mixture of vertical and horizontal products, focusing instead on clean PCR sequences to boost phylogenetic coverage. At the same time, none of these plants showed nearly as many complex, heterogeneous PCR products as *Amborella*, and thus none could approach it in extent of HGT.

Clearly, then, the incidence of HGT varies markedly from one plant to another. One wonders how many other *Amborella*-type situations exist among the ≈255,000 species of flowering plants. Are the eight cases of thus-far singleton HGT identified in other seed plants (3–6) exceptional for these genomes, or are some of these genomes also replete with HGT?

Why is *Amborella* so extraordinarily rich in HGT? *Amborella* is a monotypic genus (and family) of shrubs endemic to New Caledonia, where it grows in midelevation (600–900 m high) montane tropical rain forests (25). Epiphytic and parasitic plants are common in this environment, and *Amborella* leaves and stems are often covered with diverse epiphytes, including mosses and other bryophytes (e.g., Fig. 4). This could readily promote direct, plant-to-plant HGT, especially given the potential for herbivory to introduce epiphytic tissue and exudates within wounded *Amborella* tissue. Evidence for direct plant-to-plant HGT has recently been reported in the context of parasitism, to account for three well supported cases of transfer of mitochondrial genes from parasitic angiosperms to their hosts (6) or vice versa (5). Epiphytism may offer similar opportunities for HGT.

The New Caledonian flora is one of the most bizarre and endemic in the world, with endemism approaching 80% for the ≈3,400 vascular plants native to the island (26). Molecular examination of the flora growing on and in the general habitat of *Amborella* should prove crucial in efforts to (*i*) elucidate the factors promoting such extensive HGT, (*ii*) uncover other cases of extensive HGT, (*iii*) pinpoint donor identities, (*iv*) estimate the timing of transfer, and (*v*) estimate the number of transfers. This last issue relates to the fact that plant mitochondria frequently fuse (27), with their genomes recombining (28), which makes it easy to imagine multiple mitochondrial genes being acquired in a single event involving whole-mitochondrial transfer. Three of the moss transfers are evidently from different donor lineages (Fig. 2), as are two of the eudicot transfers (*nad1* and *ccmF*$_{N1}$; Fig. 3), but assessing the overall balance between fewer multigene transfers or more numerous single-gene transfers will require far more extensive sampling of genes and plants.

Limits and Logical Bases of Inferring HGT in Plant Mitochondrial Genomes. The PCR approach we used to census the *Amborella* mitochondrial protein gene set will clearly miss important components of the mitochondrial genome. These include mitochondrial rRNA and tRNA genes (the latter are also too small for a meaningful PCR approach), chloroplast-derived sequences (which are commonly found in plant mtDNAs; refs. 12–16), and intergenic DNA (which makes up most of a typical plant

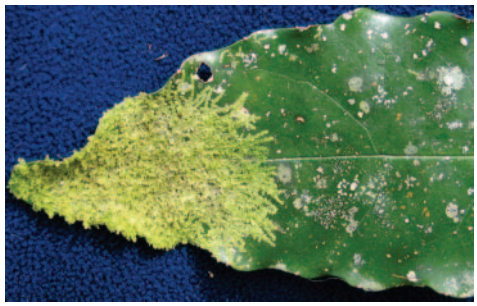

Fig. 4. *A. trichopoda* leaf from a cloud forest at Massif de l'Aoupinié (Province Nord in New Caledonia) at 801 m altitude. Note the greenish bryophyte (liverwort) growth covering the leaf tip, and the small spots of lichens and other epiphytes elsewhere on the leaf. Photograph courtesy of Sean Graham, Centre for Plant Research, University of British Columbia, Vancouver.

mitochondrial genome; refs. 12–16). Furthermore, mitochondrial protein genes from nonland plants will be so divergent (land plant mtDNAs have exceptionally low rates of sequence evolution; refs. 9 and 10) as to be strongly disfavored by PCR when faced with competition from vertically retained homologs. Only by sequencing the *Amborella* mitochondrial genome can we census its population of horizontally acquired DNA in a comprehensive and phylogenetically unbiased manner.

A major limitation in our ability to detect HGT in plant mtDNA is the often poor resolution of individual gene trees (Figs. 1, 3, and 7), which is largely a consequence of the very low rate of nucleotide substitutions in most plant mtDNAs (9, 10) and the short length of most gene regions used in our phylogenetic analyses (Table 1). Importantly, though, some of the weakly supported conflicts between mitochondrial gene trees and organismal phylogeny are most likely, given the growing evidence for HGT as an ongoing and moderately frequent process in plant mitochondrial evolution, the residue of horizontal transfer occurring within poorly resolved portions of the gene trees. We are lucky that, of all angiosperms, *Amborella* happens to be so rich in HGT, because its distinctive position at the base of angiosperms makes it relatively easy to detect with reasonable confidence transfers from other angiosperms, even with the scanty taxon sampling of this study.

Even so, a number of the putative *Amborella* transfers are admittedly not well supported by purely phylogenetic criteria. The SH test is a stringent test, and those 14 transfers that passed it (Table 1) should therefore be regarded as well supported. Some of the 12 other angiosperm cases appear to be good candidates for HGT based solely on visual inspection of phylogenetic trees, but others are less compelling (Figs. 1 and 3). Importantly, there is a second, independent criterion that we hereby invoke, namely, the very existence of divergent copies of a gene within a plant mitochondrial genome. With one possible exception (29), we are unaware of any examples of divergent duplicate genes in plant mtDNAs that are paralogs, i.e., that trace back phylogenetically to duplication events within a mitochondrial lineage. Instead, all divergent duplicates behave phylogenetically as xenologs, as the products of horizontal evolution. Moreover, plant mitochondria possess evolutionary mechanisms that tend to prohibit paralogs from diverging with time: repeated elements larger than ≈500 bp in plant mtDNAs are subject to frequent concerted evolution such that they generally remain identical to one another (12–16). HGT may be the only mechanism plant mitochondria possess to establish divergent copies of a gene. Therefore, the presence of divergent duplicates in plant mtDNA (especially when they are distantly related by phylogeny, as here) can be taken as *prima facie* evidence for HGT.

Functionality of Transferred Genes in *Amborella*. Whereas all but one of the vertically transmitted genes in *Amborella* have intact ORFs, 8 of the 26 transferred genes are pseudogenes (Table 1). Whether any of the 18 intact transferred genes are functional and under selection is an open question. Both transferred genes (*atp1* and *atp8*) whose expression has been assayed are transcribed and RNA-edited (ref. 8 and unpublished data); however, transcribed and RNA-edited pseudogenes are known to occur in plant mitochondria (30, 31). Although some of these transferred genes may be functional in *Amborella* mitochondria, we suspect that most are not, and that with time an increasing proportion will evolve into obvious pseudogenes. The time frame and dynamics of HGT in *Amborella* mitochondria may well be similar to those described in bacterial systems, where HGT regularly supplies the genome with foreign genes, most of which soon decay as pseudogenes (32).

HGT in Different Plant Genomes. These results highlight the disparity between plant mitochondrial and chloroplast genomes in their propensity to take up foreign DNA. Despite vastly more chloroplast than mitochondrial sequencing in plants, HGT is now well established for the latter but unknown for the former. Of greatest relevance, the sequenced chloroplast genome of *Amborella* (23) shows no evidence of HGT (D. W. Rice and J.D.P., unpublished work). This disparity in frequency of HGT is in keeping with other features that distinguish the two genomes. Plant mtDNAs contain much more noncoding DNA than compact chloroplast DNAs and are renowned for their frequent incorporation of chloroplast and nuclear DNA sequences, whereas chloroplasts show no evidence of intracellular gene transfer (12–16, 33). Plant nuclear genomes, on the other hand, have a loose, fluid organization (mostly noncoding DNA, many gene duplications) that would seem to accommodate HGT, are known to frequently take up DNA from organelle genomes via intracellular transfer (33), and offer one clear example of recent multigene HGT (from bacteria; ref. 34). Given this evidence and how rich it is in mitochondrial HGT, we predict that substantial levels of nuclear HGT will be found in *Amborella*.

We thank Brett Hall and the University of California Santa Cruz Arboretum, Teddi Bloniarz and the University of Massachusetts Greenhouse, Doug Soltis, and Yin-Long Qiu for supplying *Amborella* leaves, plants, and/or DNAs; Danny Rice for helpful discussion, analytical assistance, and searching for transferred genes in *Amborella* chloroplast DNA; Sean Graham for helpful discussion and Fig. 4; Lawrence Washington for expert operation of the Indiana Molecular Biology Institute DNA sequencing facility; and Jeff Doyle for critical reading of the manuscript. This research was supported by National Institutes of Health Grant R01-GM-35087 (to J.D.P.) and a National Science Foundation graduate fellowship (to A.O.R.)

1. Boucher, Y., Douady. C. J., Papke, R. T., Walsh, D. A., Boudreau, M. E., Nesbo, C. L., Case, R. J. & Doolittle, W. F. (2003) *Annu. Rev. Genet.* **37,** 283–328.
2. Richards, T. A., Hirt, R. P., Williams, B. A. & Embley, T. M. (2003) *Protist* **1,** 17–32.
3. Bergthorsson, U., Adams, K. L., Thomason, B. & Palmer, J. D. (2003) *Nature* **424,** 197–201.
4. Won, H. & Renner, S. S. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 10824–10829.
5. Davis, C. C. & Wurdack, K. J. (2004) *Science* **305,** 676–678.
6. Mower, J. P., Stefanović, S., Young, G. J. & Palmer, J. D. (2004) *Nature* **432,** 165–166.
7. Hasegawa, M., Kishino, H. & Yano, T. (1985) *J. Mol. Evol.* **22,** 160–174.
8. Shimodaira, H. & Hasegawa, M. (1999) *Mol. Biol. Evol.* **16,** 1114–1116.
9. Wolfe, K. H., Li, W.-H. & Sharp, P. M. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 9054–9058.
10. Laroche, J., Li, P., Maggia, L. & Bousquet, J. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 5722–5727.
11. Adams, K. L., Qiu, Y.-L., Stoutemyer, M. & Palmer, J. D. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 9905–9912.
12. Unseld, M., Marienfeld, J. R., Brandt, P. & Brennicke, A. (1997) *Nat. Genet.* **15,** 57–61.
13. Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., Hirai, A. & Kadowaki, K. (2002) *Mol. Genet. Genomics* **268,** 434–445.
14. Kubo, T., Nishizawa, S., Sugawara, A., Itchoda, N., Estiati, A. & Mikami, T. (2000) *Nucleic Acids Res.* **28,** 2571–2576.
15. Handa, H. (2003) *Nucleic Acids Res.* **31,** 5907–5916.
16. Clifton, S. W., Minx, P., Fauron, C., Gibson, M., Allen, J. O., Sun, H., Thompson, M., Barbazuk, B., Kanuganti, S., Tayloe, C., *et al*. (2004) *Plant Phys.* **136,** 3486–3503.
17. Qiu, Y.-L., Bernasconi-Quadroni, F., Soltis. D. E., Soltis, P. S., Zanis, M. J., Zimmer, E. A., Chen, Z., Savolainen, V. & Chase, M. W. (1999) *Nature* **402,** 404–407.
18. Mathews, S. & Donoghue, M. J. (1999) *Science* **286,** 947–950.
19. Parkinson, C. L., Adams, K. L. & Palmer, J. D. (1999) *Curr. Biol* **9,** 1485–1488.
20. Barkman, T. J., Ghenery, G., McNeal, J. R., Lyons-Weiler, J. & dePamphilis, C. W. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 13166–13171.
21. Graham, S. W. & Olmstead, R. G. (2000) *Am. J. Bot.* **87,** 1712–1730.
22. Stefanović, S., Rice, D. W. & Palmer, J. D. (2004) *BMC Evol. Biol.* **4,** 35.
23. Goremykin, V. V., Hirsch-Ernst, K. I., Wolfl, S. & Hellwig, F. H. (2003) *Mol. Biol. Evol.* **20,** 1499–1505.
24. Adams, K. L., Daley, D. O., Qiu, Y.-L., Whelan, J. & Palmer, J. D. (2000) *Nature* **408,** 354–357.
25. Feild, T. S., Zweiniecki, M. A., Brodribb, T., Jaffré, T., Donoghue, M. J. & Holbrook, N. M. (2000) *Int. J. Plant Sci.* **161,** 705–712.
26. Morat, P. (1993) *Biodiv. Lett.* **1,** 72–81.
27. Arimura, S., Yamamoto, J., Aida, G., Nakazono, M. & Tsutsumi, N. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 7805–7808.
28. Boeshore, M. L., Lifshitz, I., Hanson, M. R. & Izhar, S. (1983) *Mol. Gen. Genet.* **190,** 459–467.
29. Perrotta, G., Grienenberger, J. M. & Gualberto, J. M. (2002) *Plant Mol. Biol.* **50,** 523–533.
30. Brandt, P., Unseld, M., Eckert-Ossenkopp, U. & Brennicke, A. (1993) *Curr. Genet.* **24,** 330–336.
31. Subramanian, S., Gallahi, M. & Bonen, L. (2001) *Curr. Genet.* **39,** 264–272.
32. Liu, Y., Harrison, P. M., Kunin, V. & Gerstein, M. (2004) *Genome Biol.* **5,** R64.
33. Timmis, J. N., Ayliffe, M. A., Huang, C. Y. & Martin, W. (2004) *Nat. Rev. Genet.* **5,** 123–135.
34. Intrieri, M. C. & Buiatti, M. (2001) *Mol. Phylogenet. Evol.* **20,** 100–110.