Heredity (2002) 89, 394–400

www.nature.com/hdy

Host-associated allozyme variation in tree cambium miners, *Phytobia* spp. (Diptera: Agromyzidae)

T Nyman[1,3], T Ylioja[2,4] and H Roininen[1,2]
[1]*Department of Biology, University of Joensuu, P.O. Box 111, FIN-80101 Joensuu, Finland;* [2]*Punkaharju Research Station, Finnish Forest Research Institute, FIN-58450 Punkaharju, Finland*

The larvae of the agromyzid flies that belong to the genus *Phytobia* Lioy feed by mining in the differentiating xylem just below the cambium of growing forest trees. The genus, which is apparently one of the most primitive groups in the Agromyzidae, comprises over 50 currently recognized species. Most of the species are mono- or oligophagous, and the host plants belong to numerous genera in about 60 families. Thus, *Phytobia* is an attractive candidate for studies on the evolution of insect-plant relationships. In spite of this, the taxonomy of *Phytobia* is currently poorly understood, mainly because the morphological differences between species are small. We used allozyme electrophoresis to investigate whether molecular markers could be used to separate and identify species in *Phytobia*, and to study the patterns of host use in the group. For this, we collected *Phytobia* larvae from eight host tree species occurring in southern Finland. An analysis of 10 variable allozyme loci showed that there are probably five species of *Phytobia* that feed on the hosts included in our study: one occurs on birches (*Betula pubescens* Ehrh. and *B. pendula* Roth) and alders (*Alnus incana* (L.) Moench and *A. glutinosa* (L.) Gaertn.), one on rowan (*Sorbus aucuparia* L.), and three species with overlapping feeding ranges on aspen (*Populus tremula* L.) and two willow species (*Salix phylicifolia* L. and *S. caprea* L.). Because birches and alders belong to the plant family Betulaceae, rowan to Rosaceae, and aspen and willows to Salicaceae, the host associations of the individual fly species can be explained by the taxonomic affinities of the hosts. However, our results also show that on a larger scale the evolution of host-plant associations in *Phytobia* cannot be explained by strict parallel cladogenesis (cospeciation) between the flies and their hosts.
Heredity (2002) **89,** 394–400. doi:10.1038/sj.hdy.6800158

Keywords: Agromyzidae; allozymes; coevolution; host specificity; insect-plant interactions; parallel cladogenesis; *Phytobia*; speciation

Introduction

Most plant-feeding insects feed on only one or a few host species (Jermy, 1984; Bernays and Graham, 1988), but the exact proportion of true specialists is difficult to determine. In fact, it is very likely that the number of monophagous insect species is currently underestimated, because speciation and ecological specialization in insects is frequently accompanied by little or no change in morphology (Mayr, 1963; Diehl and Bush, 1984; Drès and Mallet, 2002). Thus, an adaptive radiation to multiple hosts may result in a complex of morphologically indistinguishable host races or sibling species, which are then erroneously interpreted as one polyphagous species (Diehl and Bush, 1984; Drès and Mallet, 2002). However, cryptic species or host races can sometimes be distinguished by using molecular methods, such as allozyme electrophoresis (Feder *et al*, 1988; McPheron *et al*, 1988; Roininen *et al*, 1993; Emelianov *et al*, 1995; Condon and Steck, 1997; Nyman, 2002; Aguin-Pombo, 2002), random amplified polymorphic DNA (De Barro *et al*, 1995; Pornkulwat *et al*, 1998), or DNA sequencing (Scheffer, 2000; Scheffer and Wiegmann, 2000).

The aim of this study was to use allozyme electrophoretic data to clarify species identification in the agromyzid genus *Phytobia* Lioy. The genus consists of species that are separated from the rest of the Agromyzidae not only morphologically but, more importantly, based on the unique feeding habit of their larvae (von Tschirnhaus, 1971; Spencer, 1976; Spencer and Steyskal, 1986): the growing larvae mine within the cambial cylinder of growing trees and bushes, tunnelling several meters downwards from tree crowns towards the base (Barnes, 1933; Kangas, 1935; Hanson and Benjamin, 1967; Ylioja *et al*, 1998). Host trees fill the larval tunnels with parenchyma tissue that keep the tunnels visible within the annual rings of trees (Greene, 1914; Kangas, 1935; Gregory and Wallner, 1979; Ylioja *et al*, 1998). In trees with high commercial value used for furnishing and other specialty purposes, the dark larval tunnels are often considered defects, and can decrease the value of raw material considerably (Kangas, 1935; Hanson and Benjamin, 1967; Spencer, 1976; Spencer and Steyskal, 1986; Ylioja *et al*, 1998).

T Nyman, Department of Biology, University of Joensuu, PO Box 111, FIN-80101 Joensuu, Finland. E-mail: Tommi.Nyman@joensuu.fi
[3]*Current address: Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA.*
[4]*Current address: Department of Biological Sciences, Dartmouth College, Hanover, NH 03755, USA.*
Received 28 January 2002; accepted 30 June 2002

In addition to being a pest for the forestry industry, *Phytobia* is a promising group for evolutionary studies. It is apparently one of the most primitive genera in the Agromyzidae (Nowakowski, 1962; Spencer, 1990), and larval tunnels have been found in fossil wood dating back to the Eocene, ca. 37–58 million years ago (Süss and Müller-Stoll, 1980; but see Spencer and Steyskal, 1986, for a cautionary note). The globally distributed genus comprises over 50 currently recognized species, all of which are apparently mono- or oligophagous (Spencer and Steyskal, 1986); the host plant species belong to numerous genera in about 60 families (Süss and Müller-Stoll, 1980; Spencer, 1990). Thus, *Phytobia* can be used to study how the utilization of different host species has changed during the evolutionary history of insects.

Studying the evolution of host-plant associations in *Phytobia* would require knowledge on the phylogenetic relationships between the species within the genus, but even the basic taxonomy of the group is unclear. Species descriptions in *Phytobia* are based on morphological characters of adults, which can be very similar between some species (Spencer, 1976; Spencer and Steyskal, 1986). In many cases, only small differences in male genitalia may be used to differentiate the species (Spencer, 1976; Spencer and Steyskal, 1986). The adults are difficult to catch and rear, and the white, slender larvae, which can be easily collected from young trees, offer few useful morphological characters for taxonomic studies (Barnes, 1933; Kangas, 1935; Spencer, 1976; Süss and Müller-Stoll, 1980). Consequently, it is not surprising that the taxonomy of the genus has been changing constantly, and the host ranges of many species are unknown (Spencer, 1976, 1990; Spencer and Steyskal, 1986).

The taxonomy of *Phytobia* has been most extensively studied in Europe and North America, but even in these areas there is considerable confusion about the number of known species. During the first half of the last century, Kangas (1935, 1949) collected and described cambium miners in southern Finland. The first two new species that he described were *P. betulae* Kangas on *Betula* spp., and *P. aucupariae* Kangas on *Sorbus aucuparia* L. Kangas (1949) considered that European Salicaceae hosts support three *Phytobia* species: *P. cambii* Hendel and *P. barnesi* Hendel on *Populus* and *Salix* species, and *P. tremulae* Kangas on *Populus tremula* L. However, these three Salicaceae-feeding species have later been synonymized under *P. cambii* (Spencer, 1976, 1990; Martinez *et al*, 1985). Recently, even *P. betulae* has been synonymized with *P. cambii* (von Tschirnhaus, 2000).

The purpose of our study was to investigate whether allozyme electrophoretic markers can be used to differentiate between *Phytobia* species occurring on eight host tree species belonging to five different genera in three plant families. An allozyme-based identification method would circumvent the need to rear the larvae to adults, facilitate rapid identification, and provide tools for research aiming at management of *Phytobia*.

Furthermore, we wanted to study whether the host ranges of the species can be explained by the phylogenetic relationships of the host plants, because it has been suggested that endophagous (mining) larval habits may enhance the probability of parallel cladogenesis (= cospeciation) between insects and their hosts (Farrell and Mitter, 1990, 1998).

Materials and methods

Life history of *Phytobia*

All *Phytobia* species that have been studied are univoltine (Kangas, 1935, 1949; Spencer, 1976). In the spring, the ca. 0.5 cm long, black females oviposit into growing shoots of their host trees (Barnes, 1933; Kangas, 1949; Ylioja *et al*, 1998; Ylioja, 2000). The larva starts mining down the shoot, and when it reaches the trunk, it continues downward, mining in the differentiating xylem just below the cambium (Gregory and Wallner, 1979; Ylioja *et al*, 1998). The feeding tunnel can be seen as a brownish 1–2 mm wide streak under the bark of the host tree (Spencer, 1976; Ylioja *et al*, 1998). The larvae may reverse their direction and mine up and down if they reach the base of the tree (Greene, 1914; Grossenbacher, 1915; Kangas, 1935; Hanson and Benjamin, 1967; Ylioja *et al*, 1998). In autumn, the ca. 2 cm long larvae of most *Phytobia* species exit the tree via a vertical slit that they cut in the bark; pupation and overwintering occurs in the ground, and the adults emerge the next spring (Barnes, 1933; Kangas, 1935, 1949; Hanson and Benjamin, 1967; Spencer, 1976). However, in some species (eg, *P. aucupariae*), the larvae overwinter under the bark and resume feeding the next spring (Grossenbacher, 1915; Kangas, 1949, 1955; Spencer, 1976; Spencer and Steyskal, 1986). In these species, the larvae exit the tree and pupate on the ground in late spring, and the adults emerge, mate, and oviposit later in the summer (Grossenbacher, 1915; Kangas, 1949, 1955; Spencer, 1976).

Sample collection and electrophoresis

Following Kangas (1935, 1949), we collected larvae of *Phytobia* from the surroundings of the Punkaharju Research Station in southeastern Finland (61°48′N, 29°20′E). Larvae were collected by peeling the bark off young 2–5 m tall trees representing eight host species: *Betula pendula* Roth (40 larvae from 12 trees), *B. pubescens* Ehrh. (60 larvae from 19 trees), *Alnus incana* (L.) Moench (25 larvae from 21 trees), *A. glutinosa* (L.) Gaertn. (20 larvae from 10 trees), *Populus tremula* L. (15 larvae from 11 trees), *Salix phylicifolia* L. (35 larvae from 25 trees), *S. caprea* L. (30 larvae from 19 trees), and *Sorbus aucuparia* L. (30 larvae from 23 trees). Of these, *Betula* and *Alnus* belong to the plant family Betulaceae, *Populus* and *Salix* to Salicaceae, and *Sorbus* to Rosaceae. Most collections were made between 28 July and 20 August in 1997; the larvae infesting *Sorbus aucuparia* were, however, collected from the same location in July 1998, because the larvae collected in the previous autumn were too small for electrophoresis. All larvae were stored in Eppendorf tubes at –80°C.

In all, 255 larvae were used for electrophoresis (Table 1). The samples were homogenized in 60 μl sample buffer (Ferguson, 1980), after which the tubes were centrifuged at 10 000 rpm for 4 min; 45 μl of the supernatant was used in standard starch gel electrophoresis according to the protocol outlined in Vuorinen (1984). The horizontal 12% starch gels were cut into 2 mm thick slices and stained for specific enzymes. Buffer systems were as in Roininen *et al* (1993) and Nyman *et al* (1998), except that *Fumh*-1 and *Fumh*-2 were scored from buffer B gels (see Nyman *et al*, 1998).

Ten variable enzyme loci were scored: phosphoglucomutase (*Pgm*, EC 5.4.2.2.), malic enzyme (*Me*, EC 1.1.1.40),

Table 1 Allele frequencies and sample sizes (*N*) at ten polymorphic loci in *Phytobia* larvae collected from eight different host species

Locus	*Allele*	*Host*										
		B. pen	*B. pub*	*A. inc*	*A. glu*	*P. tre*	*S. phy*	*S. cap*	*S. auc*	*'Sal 1'*	*'Sal 2'*	*'Sal 3'*
Me	**(*N*)**	**40**	**57**	**25**	**20**	**15**	**33**	**30**	**10**	**27**	**13**	**34**
	154	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
	100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	1.000	1.000	1.000
Mdh	**(*N*)**	**36**	**50**	**23**	**12**	**10**	**23**	**14**	**10**	**10**	**12**	**22**
	167	0.000	0.020	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	126	0.000	0.000	0.022	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	100	1.000	0.970	0.978	1.000	1.000	1.000	1.000	0.000	1.000	1.000	1.000
	105	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
	46	0.000	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Idh	**(*N*)**	**40**	**60**	**25**	**20**	**15**	**35**	**30**	**30**	**27**	**14**	**35**
	131	0.000	0.000	0.000	0.000	0.000	0.014	0.000	0.000	0.019	0.000	0.000
	100	0.975	0.967	0.960	0.975	1.000	0.986	1.000	0.000	0.981	1.000	1.000
	85	0.025	0.017	0.040	0.025	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	55	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
	44	0.000	0.017	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Pgm	**(*N*)**	**38**	**54**	**25**	**20**	**14**	**34**	**29**	**29**	**27**	**14**	**35**
	100	1.000	1.000	1.000	1.000	0.000	0.000	0.034	0.000	0.000	0.000	0.000
	83	0.000	0.000	0.000	0.000	0.714	0.118	0.000	1.000	0.000	1.000	0.000
	66	0.000	0.000	0.000	0.000	0.071	0.118	0.103	0.000	0.000	0.000	0.229
	51	0.000	0.000	0.000	0.000	0.143	0.500	0.276	0.000	0.000	0.000	0.771
	43	0.000	0.000	0.000	0.000	0.071	0.265	0.586	0.000	1.000	0.000	0.000
Aat-1	**(*N*)**	**35**	**48**	**11**	**19**	**8**	**25**	**19**	**26**	**19**	**7**	**23**
	–100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	1.000	1.000	1.000
	–108	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
Aat-2	**(*N*)**	**39**	**59**	**22**	**20**	**14**	**34**	**29**	**30**	**26**	**14**	**35**
	156	0.000	0.000	0.000	0.000	0.179	0.441	0.241	0.000	0.000	0.000	0.686
	100	0.974	0.983	1.000	1.000	0.786	0.412	0.655	0.000	0.865	1.000	0.171
	44	0.026	0.017	0.000	0.000	0.036	0.147	0.103	1.000	0.135	0.000	0.143
Fumh-1	**(*N*)**	**30**	**49**	**21**	**18**	**15**	**33**	**29**	**30**	**25**	**14**	**35**
	333	0.000	0.000	0.048	0.000	0.000	0.030	0.000	0.000	0.000	0.000	0.029
	–100	1.000	1.000	0.952	1.000	1.000	0.955	1.000	0.000	1.000	1.000	0.957
	–767	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
	–1000	0.000	0.000	0.000	0.000	0.000	0.015	0.000	0.000	0.000	0.000	0.014
Fumh-2	**(*N*)**	**38**	**57**	**21**	**19**	**15**	**31**	**29**	**30**	**26**	**14**	**32**
	130	0.000	0.000	0.024	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	100	0.961	0.904	0.738	0.868	0.000	0.016	0.034	0.000	0.000	0.000	0.016
	42	0.039	0.096	0.238	0.132	1.000	0.984	0.966	0.000	1.000	1.000	0.984
	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
G3pdh	**(*N*)**	**40**	**59**	**22**	**20**	**14**	**35**	**27**	**30**	**26**	**14**	**34**
	283	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
	100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	1.000	1.000	1.000
Tpi	**(*N*)**	**40**	**59**	**25**	**20**	**15**	**35**	**30**	**30**	**27**	**14**	**35**
	151	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000	0.000	0.000
	126	0.013	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	100	0.988	1.000	1.000	1.000	1.000	1.000	1.000	0.000	1.000	1.000	1.000

Hosts are abbreviated by giving only the first letter of the generic name, and the three first letters of the species names (eg, *B. pen* = *Betula pendula*). See Results and Table 2 for explanation for the three 'Salicaceae' species ('Sal 1–3').

malate dehydrogenase (*Mdh*, EC 1.1.1.37), isocitrate dehydrogenase (*Idh*, EC 1.1.1.42), aspartate aminotransferase (*Aat*-1 and *Aat*-2, EC 2.6.1.1), triose-phosphate isomerase (*Tpi*, EC 5.3.1.1), fumarase hydratase (*Fumh*-1 and *Fumh*-2, EC 4.2.1.2), and glycerol-3-phosphate dehydrogenase (*G3pdh*, EC 1.1.1.8).

Data analysis

Allele frequencies and sample sizes were calculated by using *BIOSYS-1* version 1.7 (Swofford and Selander, 1989). Deviations from Hardy-Weinberg equilibrium (HWE) in individual loci were tested by Fisher's exact test in the Genetic Data Analysis program (*GDA*; Lewis and Zaykin, 2001). Ten thousand permutations were used in the tests, and the obtained significance levels were adjusted using a sequential Bonferroni correction (Rice, 1989). Linkage disequilibrium between pairs of polymorphic loci was tested similarly, but the genotypes in individual loci were preserved in the permutation procedure.

Clustering analyses were performed using the *TFPGA* version 1.3 program (Miller, 1997). UPGMA clustering analyses were based on Nei's (1978) unbiased genetic distances, and cluster support was inferred by bootstrapping 1000 times over loci. Because the samples collected from *S. aucuparia* and the *Alnus* species had no alleles in common (Table 1), genetic distances between these two groups could not be calculated. Consequently, two separate clustering analyses were performed: when the *S. aucuparia* sample was included, the *Alnus* samples were excluded, and vice versa.

Results

Allele frequencies and sample sizes in individual loci are presented in Table 1. Seven statistically significant deviations from HWE were found in the samples. With the exception of the *Idh* locus in the *B. pubescens* sample, all deviations occur in the *Pgm* and *Aat*-2 loci in all *Populus* and *Salix* samples (all Bonferroni corrected *P*:s < 0.05), caused by a deficiency of heterozygotes. In addition, these two loci are in statistically significant linkage disequilibria in all samples collected from these three host species (all uncorrected $P < 0.02$). In the other samples, no statistically significant linkage disequilibria were found.

Based on these observations, the individuals collected from *P. tremula*, *S. phylicifolia* and *S. caprea* were sorted according to their genotype at the *Pgm* locus, at which heterozygotes between some alleles were totally missing (only $Pgm^{51/66}$ heterozygotes were found); the resulting 'sorted' samples are shown in Table 1 as three putative species ('Salicaceae 1-3'), and their corresponding genotypes and hosts are shown in Table 2. Because of the linkage disequilibrium between the *Pgm* and *Aat*-2 loci, the sorting also leads to clear differences in allele frequencies at the *Aat*-2 locus (Tables 1 and 2). Three individuals having missing data in the *Pgm* locus were excluded from the sorted data set, as was one individual having a typical *Betula*/*Alnus* genotype. This last individual may represent an oviposition error (see Discussion). In the sorted data set, most deviations from HWE are no longer statistically significant in the 'Salicaceae 1-3' samples, but there is still a statistically significant deviation in 'Salicaceae 3' at the *Aat*-2 locus ($P < 0.001$). This may be caused by the fact that weakly stained $Aat\text{-}2^{44/156}$ heterozygotes may have been inadvertently scored as $Aat\text{-}2^{100/100}$ homozygotes, because the heterodimer is stained more strongly than the homodimers. However, excluding these individuals from the sorted data set has no meaningful effect on the results and conclusions in this study. The only statistically significant linkage disequilibrium in the sorted data set is between *Fumh*-1 and *Fumh*-2 in 'Salicaceae 3', but the disequilibrium is not significant if a Bonferroni correction is made (uncorrected $P = 0.027$).

The UPGMA clustering analysis using the original data set (Figure 1a) shows that the samples collected from the two *Betula* species and the two *Alnus* species form a distinct and well supported cluster. The samples collected from *Populus tremula* and the two *Salix* species are grouped together in a more heterogeneous cluster. The sample collected from *S. aucuparia* is extremely different from the other samples. When the clustering analyses are redone using the sorted data set, the results are otherwise similar to the original data, except that the heterogeneity within the Salicaceae-feeding cluster is increased (Figure 1b).

Thus, the results indicate that there are five *Phytobia* species that feed on the eight host species included in this study: the *Phytobia* larvae collected from the *Betula* and *Alnus* species represent a single species, as do the larvae from *S. aucuparia*. However, the larvae from *Populus tremula* and the two *Salix* species apparently represent three species with partially overlapping feeding ranges. The distribution of individuals on different host species is statistically significantly different between the Salicaceae-feeding species ($\chi^2_4 = 40.23$, $P < 0.001$; see Table 2). The assumptions of the χ^2-test are violated because some individuals were collected from the same tree, and thus they may be siblings (however, larvae representing different species could be found on single tree individuals). Thus, the test was repeated so that if multiple larvae representing a given species had been collected from a given tree individual, they were treated as only one larva, ie, it

Table 2 The three putative Salicaceae-feeding species that were obtained by sorting the individuals collected from *Populus tremula*, *Salix phylicifolia*, and *S. caprea* according to their genotype at the *Pgm* locus (see Results), and the host plants on which these individuals were found. Because of the linkage disequilibrium between *Pgm* and *Aat*-2, the sorting also results in clear differences in allele frequencies at the *Aat*-2 locus (see Table 1)

Putative species	*Allele(s) found in locus*		*Hosts*
	Pgm	*Aat-2*	
Salicaceae 1	43	100/44	*P. tremula* (1), *S. phylicifolia* (9), *S. caprea* (17)
Salicaceae 2	83	100	*P. tremula* (10), *S. phylicifolia* (4)
Salicaceae 3	51/66	44/100/156	*P. tremula* (3), *S. phylicifolia* (21), *S. caprea* (11)

Numbers in parentheses after each host species refer to the number of larvae collected from that host.

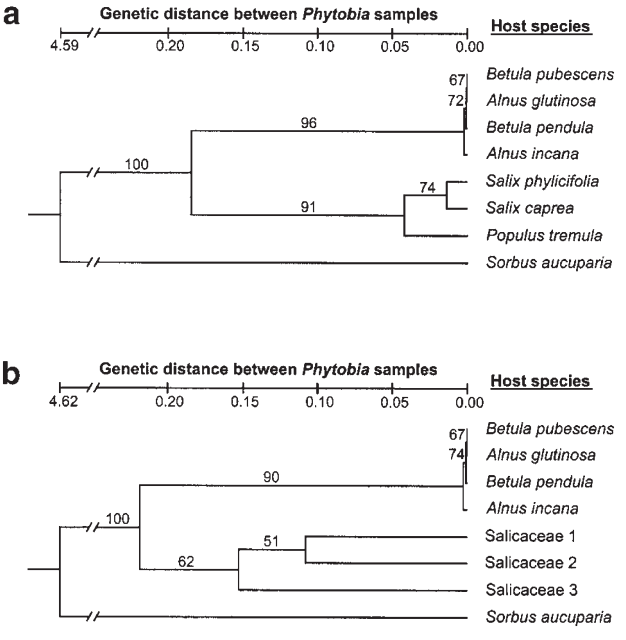

Figure 1 UPGMA clustering dendrograms of *Phytobia* samples collected from eight different host tree species, based on pairwise Nei's (1978) unbiased genetic distances calculated from 10 polymorphic allozyme loci. Host species are shown to the right of the dendrograms, numbers above branches are bootstrap proportions (%) from 1,000 replicates. Results are shown (a) for the original data set, and (b) for a data set in which the Salicaceae-feeding individuals have been sorted into three putative species (Salicaceae 1-3) according to their genotype at the *Pgm* locus (see Results and Table 2 for explanation, and list of hosts for these species). In both cases, analyses including the *S. aucuparia* sample were made while excluding the *Alnus* samples, and vice versa (see Materials and methods). Note the cut in the scale.

was assumed that they represent the offspring of a single female. The result of the analysis remains essentially the same ($\chi^2_4 = 25.76$, $P < 0.001$).

The aforementioned results are supported by the results obtained in a pilot study, conducted in 1997 using larvae collected from five of the host species in 1996 (n = 9–41 larvae per host species; results not shown). In the pilot survey, allele frequencies were close to the ones observed in this study and, consequently, the larvae collected from the two *Betula* species and *A. incana* were very different from the samples collected from the two *Salix* species. Likewise, all 'Salicaceae 1–3' species were observed in the *Salix* samples in the preliminary study.

Discussion

Although *Phytobia* has been extensively studied in Europe, the number of species that actually exist in the region is still unclear. As stated above, Kangas (1935, 1949) considered that there are five *Phytobia* species that feed on the tree species represented in this study: *P. betulae* on *Betula* spp., *P. aucupariae* on *S. aucuparia*, and three species (*P. cambii*, *P. barnesi*, and *P. tremulae*) on hosts belonging to the Salicaceae. Later, Spencer (1976) considered that the presence/absence of the mid-tibial bristle, which Hendel (in Barnes, 1933) and Kangas (1949) had mainly used to separate the Salicaceae-feeding species from each other, was not a clear enough character for species separation. Consequently, Spencer (1976, see also Spencer, 1990) considered that there exists only one European Salicaceae-feeding species (ie, *P. cambii*), and this view has later been followed by others (Martinez *et al*, 1985; von Tschirnhaus, 2000). In the end, von Tschirnhaus (2000) even synonymized *P. betulae* with *P. cambii*, which would mean that there would be only two European *Phytobia* species on the hosts in this study: *P. cambii* on Betulaceae and Salicaceae hosts, and *P. aucupariae* on *S. aucuparia*.

The results from our study clearly support Kangas' (1935, 1949) interpretation of the species situation in *Phytobia*. The five species that can be distinguished in our study are characterized by fixed differences in various loci. The clearest one is the *S. aucuparia* species (= *P. aucupariae*), and the apparently single species on *Betula* and *Alnus* (= *P. betulae*) is distinguished by the Pgm^{100} allele (Table 1). One individual having this allele was found on *S. caprea*, but because the individual was homozygous for the allele, it seems more likely that it represents an oviposition error rather than evidence for hybridization between species. Oviposition errors are probably a common phenomenon in phytophagous insects, although quantitative studies are rare (Larsson and Ekbom, 1995). The three Salicaceae species can be separated on the basis of their genotype at the *Pgm* locus (Table 2). We note that these Salicaceae-feeding types were also found in the preliminary survey that was conducted with samples collected from the study area in the previous summer (see Results). The three species apparently represent *P. cambii*, *P. barnesi*, and *P. tremulae*, but further study is needed in order to establish which 'allozyme species' corresponds to which morphospecies.

Studies on insect host races have shown that host-dependent disruptive selection can cause and maintain large differences in allele frequencies at allozyme loci, even when there is considerable gene flow between the host races (Feder *et al*, 1997; Filchak *et al*, 2000; Drès and Mallet, 2002). In the case of *Phytobia*, host-associated disruptive selection seems an unlikely explanation for the results above, because there are fixed differences at single or multiple loci and, more importantly, the 'Salicaceae 1–3' species overlap in their host use. Thus, our results indicate that all five species that can be identified on the basis of allozyme data are true biological species (*sensu* Mayr, 1963) that do not interbreed.

At the species level, the host associations can clearly be explained by the taxonomic affinities of the hosts. *Betula* and *Alnus*, which support a single *Phytobia* species (= *P. betulae*), belong to the plant family Betulaceae. In Central Europe, *P. betulae* has been reported to feed on *Corylus* and *Carpinus*, which also belong to the same family (Spencer, 1990). *Sorbus aucuparia*, which supports a very distinct *Phytobia* species (= *P. aucupariae*), belongs to the family Rosaceae, and the three species ('Salicaceae 1–3') with overlapping host associations all feed on trees belonging to the Salicaceae. Thus, the results conform to the common pattern that when an herbivorous insect is oligo- or polyphagous, the hosts tend to be closely related (Bernays and Graham, 1988).

However, on a larger scale it is evident that the evolution of host use in *Phytobia* cannot be explained by simple parallel cladogenesis (= cospeciation) between the flies and their host plants. Although UPGMA clustering is sensitive to rate heterogeneity in the lineages that are being studied (Swofford *et al*, 1996), the species occurring on *S. aucuparia* is so different from the four others that it is highly likely that it was the first one to diverge. In fact, the genetic distance between this species and the other samples is more typical of insects in separate genera (see Emelianov *et al*, 1995). This is also supported by the fact that *P. aucupariae* has a rather distinct morphology and life cycle: whereas the larvae of species that feed on Betulaceae and Salicaceae exit the trees in the autumn to pupate and overwinter on the ground, the larvae of *P. aucupariae* overwinter in the trees and pupate on the ground in the next spring (Kangas, 1949; Spencer, 1976). Species having a rather similar morphology and life cycle are also known from other tree species belonging to Rosaceae (Grossenbacher, 1915; Kangas, 1955; Spencer and Steyskal, 1986). However, the Rosaceae has probably been colonized independently at least twice, because some of the Rosaceae-feeding species overwinter as pupae on the ground (Spencer, 1976, 1990; Spencer and Steyskal, 1986). In any case, the branching pattern observed in our study, ie, (Rosaceae(Betulaceae, Salicaceae)), clearly contradicts the current hypothesis of the phylogenetic relationships between the host species represented in this study, ie, (Salicaceae(Betulaceae, Rosaceae)) (Angiosperm Phylogeny Group, 1998). A somewhat similar result was found by Scheffer and Wiegmann (2000) in the case of agromyzid *Phytomyza* leaf miners that mine in the leaves of hollies (*Ilex* spp.), because they also could rule out an explanation based on cospeciation. Together, these results from two distantly related agromyzid taxa provide further support for the view that strict parallel cladogenesis between insects and plants is a rare phenomenon (Mitter *et al*, 1991).

It has been suggested that mining larval habits may predispose insects to cospeciation with their hosts, but clearly the situation in *Phytobia* is very different from the situation in the *Phyllobrotica* and *Tetraopes* beetles dis-

cussed by Farrell and Mitter (1990, 1998). In contrast to the tissues of the hosts of these beetles (the Lamiales and Asclepiadaceae, respectively), the differentiating xylem in deciduous trees is probably not strongly defended by secondary compounds and, thus, host chemistry may present only weak barriers for colonization by *Phytobia*. This implies that larval endophagy *per se* does not affect the likelihood of parallel cladogenesis, but may do so in combination with heavy chemical defence of the hosts (see also Farrell and Mitter, 1990, 1998).

In conclusion, it is evident that in Finland the genus *Phytobia* is represented by at least five reproductively isolated species, and this is probably also true elsewhere in Europe. Allozyme electrophoresis appears to be a powerful tool for species identification in this genus. More thorough sampling of different tree species is needed, because it is likely that new *Phytobia* species can be found on unstudied, economically unimportant tree species. The host use of the individual species in this study seems to be explained by the taxonomic affinities of the host plants, but on a larger scale the evolution of host-plant associations cannot be explained by strict cospeciation between *Phytobia* and their hosts. A phylogenetic study of the whole *Phytobia* would add to the knowledge of evolution in the Agromyzidae, as well as in phytophagous insects in general.

Acknowledgements

We thank the Punkaharju Research Area of the Finnish Forest Research Institute for kindly providing sampling sites. We are especially grateful to Anne Matilainen, Henriikka Hyvönen, Tomi Nuora, and Jari Salminen for assistance in the collection of larvae, and to Jorma Tahvanainen, Pauline Roberts, and two anonymous reviewers for their comments on the manuscript. Funding for this study was provided by the Ministry of Agriculture and Forestry of Finland, the Emil Aaltonen Foundation, and the Academy of Finland.

References

Aguin-Pombo D (2002). Genetic differentiation among host-associated *Alebra* leafhoppers (Hemiptera: Cicadellidae). *Heredity* **88**: 415–422.

Angiosperm Phylogeny Group (1998). An ordinal classification for the families of flowering plants. *Ann MO Bot Gard* **85**: 531–553.

Barnes HF (1933). A cambium miner of basket willows (Agromyzidae) and its inquiline gall midge (Cecidomyidae). *Ann Appl Biol* **20**: 498–519.

Bernays E, Graham M (1988). On the evolution of host specificity in phytophagous arthropods. *Ecology* **69**: 886–892.

Condon MA, Steck GJ (1997). Evolution of host use in fruit flies of the genus *Blepharoneura* (Diptera: Tephritidae): cryptic species on sexually dimorphic host plants. *Biol J Linn Soc* **60**: 443–466.

De Barro PJ, Sherratt TN, Brookes CP, David O, Maclean N (1995). Spatial and temporal genetic variation in British field populations of the grain aphid *Sitobion avenae* (F.) (Hemiptera: Aphididae) studied using RAPD-PCR. *Proc R Soc Lond B* **262**: 321–327.

Diehl SR, Bush GL (1984). An evolutionary and applied perspective of insect biotypes. *Annu Rev Entomol* **29**: 471–504.

Drès M, Mallet J (2002). Host races in plant-feeding insects and their importance in sympatric speciation. *Phil Trans R Soc Lond B* **357**: 471–492.

Emelianov I, Mallet J, Baltensweiler W (1995). Genetic differentiation in *Zeiraphera diniana* (Lepidoptera: Tortricidae, the larch budmoth): polymorphism, host races or sibling species? *Heredity* **75**: 416–424.

Farrell B, Mitter C (1990). Phylogenesis of insect/plant interactions: have the *Phyllobrotica* leaf beetles (Chrysomelidae) and the Lamiales diversified in parallel? *Evolution* **44**: 1389–1403.

Farrell BD, Mitter C (1998). The timing of insect/plant diversification: might *Tetraopes* (Coleoptera: Cerambycidae) and *Asclepias* (Asclepiadaceae) have co-evolved? *Biol J Linn Soc* **63**: 553–577.

Feder JL, Chilcote CA, Bush GL (1988). Genetic differentiation between sympatric host races of the apple maggot fly *Rhagoletis pomonella*. *Nature* **336**: 61–64.

Feder JL, Roethele JB, Wlazlo B, Berlocher SH (1997). Selective maintenance of allozyme differences among sympatric host races of the apple maggot fly. *Proc Natl Acad Sci USA* **94**: 11417–11421.

Ferguson A (1980). *Biochemical Systematics and Evolution*. Blackie: Glasgow.

Filchak KE, Roethele JB, Feder, JL (2000). Natural selection and sympatric divergence in the apple maggot *Rhagoletis pomonella*. *Nature* **407**: 739–742.

Greene CT (1914). The cambium miner in river birch. *J Agric Res* **1**: 471–474 (+ illustrations).

Gregory R, Wallner W (1979). Histological relationship of *Phytobia setosa* to *Acer saccharum*. *Can J Bot* **57**: 403–407.

Grossenbacher JG (1915). Medullary spots and their cause. *Bull Torrey Bot Club* **42**: 227–239 (+ illustrations).

Hanson JB, Benjamin DM (1967). Biology of *Phytobia setosa*, a cambium miner of sugar maple. *J Econ Entomol* **60**: 1351–1355.

Jermy T (1984). Evolution of insect/host plant relationships. *Am Nat* **124**: 609–630.

Kangas E (1935). Die Braunfleckigkeit des Birkenholzes und ihr Uhrheber *Dendromyza* (*Dizygomyza*) *betulae* n. sp. Vorläufige Mitteilung. *Comm Inst For Fenn* **22**: 1–31.

Kangas E (1949). Die *Dendromyza*-Arten von Eberesche, Espe und Salweide und ihre Biologie. *Ann Ent Soc Fenn* **14** (Suppl): 106–117.

Kangas E (1955). *Dendromyza cerasiferae* n.sp. (Dipt., Agromyzidae) from England. *Ann Ent Soc Fenn* **21**: 165–170.

Larsson S, Ekbom B (1995). Oviposition mistakes in herbivorous insects: confusion or a step towards a new host plant? *Oikos* **72**: 155–160.

Lewis PO, Zaykin D (2001). *Genetic Data Analysis (GDA): Computer program for the analysis of allelic data*. Version 1.0 (d16c). Free program distributed by the authors over the internet from http://lewis.eeb.uconn.edu/lewishome/software.html.

Martinez M, Gumez JL, Munnier P (1985). Un ravageur mal connu: la mouche mineuse du cambium des peupliers. *Phytoma* **374**: 51–53.

Mayr E (1963). *Animal Species and Evolution*. Harvard University Press: Cambridge.

McPheron BA, Smith DC, Berlocher SH (1988). Genetic differences between host races of *Rhagoletis pomonella*. *Nature* **336**: 64–66.

Miller MP (1997). *Tools for Population Genetics Analyses (TFPGA) 1.3: A Windows program for the analysis of allozyme and molecular population genetic data*. Computer software distributed by author.

Mitter C, Farrell B, Futuyma DJ (1991). Phylogenetic studies of insect-plant interactions: insights into the genesis of diversity. *Trends Ecol Evol* **6**: 290–293.

Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583–590.

Nowakowski JT (1962). Introduction to a systematic revision of the family Agromyzidae (Diptera) with some remarks on host plant selection by these flies. *Ann Zool Warsz* **20**: 67–183.

Nyman T (2002). The willow bud galler *Euura mucronata* Hartig

(Hymenoptera: Tenthredinidae): one polyphage or many monophages? *Heredity* **88**: 288–295.
Nyman T, Roininen H, Vuorinen JA (1998). Evolution of different gall types in willow-feeding sawflies (Hymenoptera: Tenthredinidae). *Evolution* **52**: 465–474.
Pornkulwat S, Skoda SR, Thomas GD, Foster JE (1998). Random amplified polymorphic DNA used to identify genetic variation in ecotypes of the European corn borer (Lepidoptera: Pyralidae). *Ann Entomol Soc Am* **91**: 719–725.
Rice, WR (1989). Analyzing tables of statistical tests. *Evolution* **43**: 223–225.
Roininen H, Vuorinen J, Tahvanainen J, Julkunen-Tiitto R (1993). Host preference and allozyme differentiation in shoot galling sawfly, *Euura atra*. *Evolution* **47**: 300–308.
Scheffer S (2000). Molecular evidence of cryptic species within the *Liriomyza huidobrensis* (Diptera: Agromyzidae). *J Econ Entomol* **93**: 1146–1151.
Scheffer S, Wiegmann BM (2000). Molecular phylogenetics of the holly leafminers (Diptera: Agromyzidae: *Phytomyza*): species limits, speciation, and dietary specialization. *Mol Phylogenet Evol* **17**: 244–255.
Spencer KA (1976). The Agromyzidae (Diptera) of Fennoscandia and Denmark. *Fauna Ent Scand* **5**: 1–606.
Spencer KA (1990). *Host Specialization in the World Agromyzidae (Diptera)*. Kluwer Academic Publishers: London.
Spencer KA, Steyskal GC (1986). *Manual of the Agromyzidae (Diptera) of the United States*. United States Department of Agriculture: Washington.
Süss H, Müller-Stoll WR (1980). Das fossile Holz *Pruninium gummosum* Platen emend. Süss u. Müller-Stoll aus dem Yellowstone Nationalpark und sein Parasit *Palaeophytobia prunorum* sp. nov. nebst Bemerkungen über Markflecke. In: Vent W (ed) *100 Jahre Arboretum Berlin*, Humboldt University: Berlin. pp 343–364.
Swofford DL, Selander RB (1989). *BIOSYS-1. A computer program for the analysis of allelic variation in population genetics and biochemical systematics*. Release 1.7. University of Illinois: Urbana.
Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996). Phylogenetic inference. In: Hillis DM, Moritz C, Mable BK (eds) *Molecular Systematics*, 2nd edn. Sinauer: Sunderland. pp 407–514.
von Tschirnhaus M (1971). Unbekannte Stridulationsorgane bei Dipteren und ihre Bedeutung für Taxonomie und Phylogenetik der Agromyziden. *Beitr Ent* **21**: 551–579.
von Tschirnhaus M (2000). Agromyzidae. *Nova Suppl Ent* **14**: 133–149.
Vuorinen J (1984). Electrophoretic expression of genetic variation and duplicate gene activity in vendace, *Coregonus albula* (Salmonidae). *Hereditas* **101**: 85–96.
Ylioja T (2000) Relationship between *Phytobia betulae* and its host tree *Betula* sp.. *Faculty of Forestry, University of Joensuu, Research Notes* **112**: 1–26.
Ylioja T, Saranpää P, Roininen H, Rousi M (1998). Larval tunnels of *Phytobia betulae* (Diptera: Agromyzidae) in birch wood. *J Econ Entomol* **91**: 175–181.