Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 43 (2007) 251–269
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Defining the limits of taxonomic conservatism in host–plant use for phytophagous insects: Molecular systematics and evolution of host–plant associations in the seed-beetle genus *Bruchus* Linnaeus (Coleoptera: Chrysomelidae: Bruchinae)

Gaël J. Kergoat [a,*], Jean-François Silvain [b], Alex Delobel [c], Midori Tuda [d], Klaus-Werner Anton [e]

[a] *INRA, Centre de Biologie et de Gestion des Populations (CBGP-UMR1062), Campus international de Baillarguet, 34988 Montferrier-sur-Lez, France*
[b] *IRD, UR 072 c/o CNRS, UPR 9034, Laboratoire Evolution, Génomes et Spéciation (LEGS), avenue de la Terrasse, 91198 Gif-sur-Yvette, France*
[c] *Laboratoire d'Entomologie (MNHN, DSE), 45 rue Buffon, 75005 Paris, France*
[d] *Institute of Biological Control, Faculty of Agriculture, Kyushu University, Hakozaki 6-10-1, Fukuoka 812-8581, Japan*
[e] *Grünewaldstr. 13, D-79312 Emmendingen, Germany*

Received 27 March 2006; revised 16 October 2006; accepted 28 November 2006
Available online 9 December 2006

Abstract

In this study, we have investigated the limits of taxonomic conservatism in host–plant use in the seed-beetle genus *Bruchus*. To reconstruct the insect phylogeny, parsimony and multiple partitioned Bayesian inference analyses were conducted on a combined data set of four genes. Permutation tests and both global and local maximum-likelihood optimizations of host preferences at distinct taxonomic levels revealed that host-fidelity is still discernible beyond the host–plant tribe level, suggesting the existence of more important than previously thought evolutionary constraints, which are further discussed in details. Our tree topologies are also mostly consistent with extant taxonomic groups. Through the analysis of this empirical data set we also provide meaningful insights on two methodological issues. First, Bayesian inference analyses suggest that partitioning by using codon positions greatly increase the accuracy of phylogenetical reconstructions. Regarding reconstruction of ancestral character states through maximum likelihood, the present study also highlights the usefulness of local optimizations. The issue of over-parameterization is also addressed, as the optimizations with the most parameter-rich models have returned the most counterintuitive results.

Keywords: Ancestral character state reconstruction; Bruchinae; *Bruchus*; Partitioned analyses; Phylogenetic inference; Phytophagous insects; Taxonomic conservatism

1. Introduction

Taxonomic conservatism in host–plant use, where phylogenetically related insects feed on phylogenetically related plants, is one of the most recognized patterns of insect host–plant interactions (e.g., see Ehrlich and Raven, 1964; Farrell, 2001; Farrell and Mitter, 1990; Janz and Nylin, 1998; Kergoat et al., 2004, 2005a; Silvain and Delobel, 1998). According to Ehrlich and Raven (1964), this conservatism in host association could be accounted for by the strong influence of plant secondary compounds since related host–plants generally share the same toxic compounds. Although several studies have established the influence of host–plant chemistry on the evolution of host use in several insect groups (Becerra, 1997; Futuyma and McCafferty, 1990; Kergoat et al., 2005b; Termonia et al., 2002), it has also become obvious that numerous other factors (e.g., behavorial factors, geographic distribution, genetic constraints or phenology of host–plants) may influence the evolution of

* Corresponding author.
E-mail address: kergoat@ensam.inra.fr (G.J. Kergoat).

1055-7903/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2006.11.026

insect host–plant associations (Becerra and Venable, 1999; Bernays, 2001; Dobler and Farrell, 1999; Futuyma et al., 1993; Kawecki and Mery, 2003; Morse and Farrell, 2005; Siemens et al., 1991; Thompson, 1993; Tuda et al., 2005, 2006) and thus explain the more or less pronounced extant patterns of taxonomic conservatism. Although various hierarchical levels of specialization onto particular plant lineages (e.g., family, tribe or genus) do exist (Johnson, 1980; Odegaard et al., 2005; Scheffer and Wiegmann, 2000), few studies have investigated their boundaries when testing for possible patterns of taxonomic conservatism (but see Wahlberg, 2001; Yotoko et al., 2005). In this study, we specifically address the question of the limits of taxonomic conservatism by investigating the phylogenetic relationships and host–plant use in a highly specialized genus of seed-beetles (Coleoptera, Bruchinae).

Among the large (40,000 species) family Chrysomelidae, the seed-beetles constitute a homogeneous group of 1700 species (Johnson et al., 2004). In this study, we follow the view of many authors who lowered seed-beetles to subfamily level (Lingafelter and Pakaluk, 1997; Reid, 1996; G.E. Morse, pers. comm.; but see also Kingsolver, 1995; Schmitt, 1998; Verma and Saxena, 1996). This change in taxonomy is supported by recent phylogenetic studies which have confirmed the inclusion of seed-beetles within the family Chrysomelidae and the position of chrysomelid subfamily Sagrinae as sister-group of Bruchinae (Duckett et al., 2003; Farrell, 1998; Farrell and Sequeira, 2004). As indicated by their common name, seed-beetles are characterized by a strong plant tissue specialization as their larvae only develop inside seeds even though a few species may complete their development in other plant parts (Hoffmann, 1945). They also show a high trend toward host-specialization: (i) they are generally monophagous or oligophagous; (ii) most bruchine tribes are affiliated to specific plant families: Bruchini on Fabaceae, Megacerini on Convolvulaceae, Pachymerini on Arecaceae, Spermophagini on Convolvulaceae and Malvaceae (Borowiec, 1987; Johnson, 1981). Interestingly, some bruchine genera exhibit stronger trends toward specialization as they are only associated with specific plant subfamilies, tribes or genera (Borowiec, 1987). For instance, species of *Sennius* (with the exception of a sole species) only develop on seeds of *Cassia* (Fabaceae, Caesalpinioideae) (Johnson, 1980). Another good example is given by the genus *Bruchus* which is almost exclusively associated with the tribe Vicieae of the Fabaceae (Delobel and Delobel, 2003, 2005).

As currently circumscribed the genus *Bruchus* Linnaeus, 1767 is composed of 36 valid species (summarized by Lukjanovitch and Ter-Minasian, 1957; revised by Anton, 2001; Borowiec, 1988; Wendt, 1993; new species added by Anton, 1999; Decelle, 1975, 1979; Ter-Minasian, 1968; Zampetti, 1993) divided into seven species groups by Borowiec (1988). *Bruchus* species are found predominantly in the Palearctic Region with only few species occurring in North Africa and Asia (Arora, 1977; Borowiec, 1987, 1988). Several species have been also accidentally introduced in North America, tropical Africa, Australia, and Japan (Borowiec, 1987; Lukjanovitch and Ter-Minasian, 1957; Morimoto, 1990). Species of *Bruchus* are well-defined by the following combinations of characters: (i) pronotum square or trapezoidal, emarginate on lateral margins near middle and with a denticle before the emargination in most species (see Lukjanovitch and Ter-Minasian, 1957 for details); (ii) middle tibia modified in male with apical spines or plates. Due to these distinctive external morphological characters, this genus has been erected in a specific subtribe, the subtribe Bruchina. *Bruchus* species are also characterized by unique male genitalia which distinguish themselves from other Bruchinae (Borowiec, 1987).

As in the case of most seed-beetles of temperate zones, the biology of *Bruchus* species is characterized by a univoltine life cycle (Huignard et al., 1990; Lukjanovitch and Ter-Minasian, 1957). Adults generally lay eggs on young pods from spring to summer (Huignard et al., 1990; N'Diaye and Labeyrie, 1990), and the subsequent larval development always occurs within a single seed (N'Diaye and Labeyrie, 1990; N'Diaye et al., 1992; Szentesi and Jermy, 1995). After their emergence, adults entered a period of reproductive diapause during autumn and winter (Huignard et al., 1990). This reproductive diapause lasts until spring and its termination is generally inducted by both photoperiod variations and pollen consumption (Tran et al., 1993; Tran and Huignard, 1992). Due to their strict univoltine life cycle, *Bruchus* are not granary pests of stored legume seeds (Southgate, 1979). However, the following species do cause major crop losses in the field: *B. lentis* on *Lens esculenta* (lentils); *B. pisorum* on *Pisum sativum* (field peas); and *B. rufimanus* on *Vicia faba* (broad beans) (Delobel and Tran, 1993; Lukjanovitch and Ter-Minasian, 1957; Smith, 1990).

Phylogenetic relationships of a representative sample of *Bruchus* species have been investigated by using the mitochondrial 12s rRNA (12S), cytochrome *b* (Cyt *b*), and cytochrome *c* oxidase subunit I (COI) genes, as well as the nuclear 28s rDNA (28S) gene. For the latter, we have sequenced a fragment which encompasses a small part of the extension segment D1 and most of the extension segment D2. The resulting phylogenetic hypotheses will allow us: (i) to test the monophyly of extant taxonomic groups using statistical tests; (ii) to study the evolution of host–plant use and the limits of taxonomic conservatism in *Bruchus* through multiple methods of optimization. In addition, this study will provide an opportunity to compare several recent methods (e.g., partitioning strategies in Bayesian inference; global and local optimizations under a maximum likelihood framework) using an empirical data set.

2. Materials and methods

2.1. Taxon sampling and species identification

Most of the specimens used in this study were reared from pods collected in the field from 2001 to 2004, and latter preserved in 95–100% ethanol. Dried specimens were

also used to encompass the largest taxon sampling (30 of the 36 known *Bruchus* species were thus sampled). Unfortunately, we were unable to recover suitable DNA templates for five species (*B. ervi*, *B. ibericus*, *B. lugubris*, *B. perezi* and *B. ulicis*) for which only dried specimens were available. Nonetheless, the remaining 25 species include members of the seven recognized taxonomic groups (see Table 1). In addition, species from several bruchine genera were assessed for use as outgroups: *Pachymerus cardo* a member of the tribe Pachymerini, and three members of the tribe Bruchini, subtribe Acanthoscelidina; *Acanthoscelides obtectus*, *Gibbobruchus* sp. and *Paleocanthoscelides gilvus*. The choice of the latter species as valid outgroups was based both on morphological data and on the results from previous studies (Borowiec, 1987; Kergoat et al., 2004, 2005b; Kergoat and Silvain, 2004; Poinar, 2005; Silvain and Delobel, 1998). Identification of species was conducted by K.-W. Anton and A. Delobel, who are recognized authorities in Old World bruchine taxonomy.

2.2. DNA extraction and polymerase chain reaction

Prior to DNA extractions, genitalia were removed from adults, mounted on microscope slide, and kept as vouchers in the Evolution, Génomes and Spéciation laboratory (LEGS) (CNRS UPR-9034, Gif/Yvette, France) (formerly Populations, Génétique et Evolution (PGE) laboratory). Whole individuals or just forelegs (for the large *P. cardo* and some rare dried specimens) were ground in phosphate buffered saline (PBS) buffer, and total DNA was extracted using the Quiagen DNAeasy tissue kit (Quiagen, Inc.). Polymerase chain reaction (PCR) amplifications were conducted as described previously (see Kergoat et al., 2004, 2005b for cycling conditions). All primer sequences are given in Table 2. PCR products were purified using Quiagen's PCR purification kit. Sequencing was carried out with an ABI 3100 automated sequencer (Applied Biosystems) with both strands sequenced for all taxa to minimize PCR artifacts and ambiguities. Further reading of the sequences was

Table 1
Taxon sampling

Genus species	Taxonomic		Genbank Accession No.			
	Groups	Locality	12S	Cyt *b*	COI	28S
Acanthoscelides						
obtectus (Say, 1831)		Giza (Eg.)	AY945982	AY947505	AY947519	DO307635
Bruchus						
affinis Frölich, 1799	*viciae*	Hte. Corse (Fr.)	AY390658	AY390721	AY390690	DO307636
altaicus Fahraeus, 1839	*atomarius*	Talysh (Az.)	DO307622	None	None	None
atomarius (Linnaeus, 1761)	*atomarius*	Htes. Alpes (Fr.)	DO307623	DO307664	DO307652	DO307637
brachialis Fahraeus, 1839	*brachialis*	Hte. Corse (Fr.)	AY390660	AY390723	AY390692	DQ307638
brisouti Kraatz, 1868	*brachialis*	(Mor.)	DO307624	DO307665	DO307653	None
canariensis Decelle (1975)	*brachialis*	Tenerife (Sp.)	None	None	DO307654	None
dentipes (Baudi, 1886)	*atomarius*	Vaucluse (Fr.)	AY390659	AY390722	AY390691	DQ307639
emarginatus Allard, 1868	*pisorum*	Vaucluse (Fr.)	DO307625	DO307666	DO307655	None
griseomaculatus Gyl., 1833	*rufipes*	Essone (Fr.)	DO307626	DO307667	DO307656	None
hamatus Miller, 1881	*brachialis*	(Tu.)	DO307627	DO307668	DO307657	None
laticollis Boheman, 1833	*brachialis*	Vaucluse (Fr.)	AY509807	AY509813	AY509810	DO307640
lends Frolich, 1799	*pisorum*	Hts-de-Seine (Fr.)	DO307628	DO307669	DO307658	None
libanensis Zampetti, 1993	*rufipes*	(Tu.)	DO307629	DO307670	DO307659	None
loti Paykull, 1800	*loti*	Oise (Fr.)	AY390661	AY390724	AY390693	DQ307641
luteicornis Illiger, 1794	*rufipes*	Vaucluse (Fr.)	AY390662	AY390725	AY390694	DO307642
occidentalis Luk.&T., 1957	*rufipes*	Hte. Corse (Fr.)	DO307630	DO307671	DO307660	DO307643
pisorum (Linnaeus, 1758)	*pisorum*	Basilicata (It.)	DQ307631	DQ307672	None	DQ307644
rufimanus Boheman, 1833	*atomarius*	Vaucluse (Fr.)	AY390663	AY390726	AY390695	DO307645
rufipes Herbst, 1783	*rufipes*	Hte. Corse (Fr.)	AY390664	AY390727	AY390696	DO307646
sibiricus Germar, 1824	*rufipes*	Aksu (Ka.)	DQ307632	None	DQ307661	None
signaticornis Gyll., 1833	*brachialis*	Herault (Fr.)	DQ307633	DQ307673	DQ307662	None
tristiculus Fahraeus, 1839	*tristis*	Vaucluse (Fr.)	AY390666	AY390729	AY390698	DO307647
tristis Boheman, 1833	*tristis*	Vaucluse (Fr.)	AY390667	AY390730	AY390699	DO307648
venustus Fahraeus, 1839	*brachialis*	Rhône (Fr.)	DQ307634	DQ307674	DQ307663	DQ307649
viciae Olivier, 1795	*viciae*	Basilicata (It.)	AY509808	AY509814	AY509811	DO307650
Gibbobruchus						
Sp.		Fr. Guyana (Fr.)	AY625331	AY625477	AY625428	None
Pachymerus						
cardo (Fahraeus, 1839)		Fr. Guyana (Fr.)	AY390636	AY390700	AY390668	AY625378
Paleacanthoscelides						
gilvus (Gyllenhal, 1839)		Vaucluse (Fr.)	AY390638	AY390702	AY390670	DO307651

Names of countries were abbreviated as follows: Azerbaijan (Az.); Egypt (Eg.); France (Fr.); Italy (It.); Kasakhstan (Ka.); Morocco (Mo.); Spain (Sp.); and Turkey (Tu.).

Table 2
Names, sequences, and references of primers used

Gene	Name of primer	Sequence of primer (5′ → 3′)	Reference
12S	SR-J-14233	AAG AGC GAC GGG CGA TGT GT	Simon et al. (1994)
	SR-N-14588	AAA CTA GGA TTA GAT ACC CTA TTA T	Simon et al. (1994)
Cyt *b*	CP1	GAT GAT GAA ATT TTG GAT C	Harry et al. (1998)[a]
	CB-J-10933	TAT GTA CTA CCA TGA GGA CAA ATA TC	Simon et al. (1994)
	CB-N-11367	ATT ACA CCT CCT AAT TTA TTA GGA AT	Simon et al. (1994)
COI	Cl-J-1751	GGA TCA CCT GAT ATA GCA TTC CC	Simon et al. (1994)
	Cl-N-2191	CCC GGT AAA ATT AAA ATA TAA ACT TC	Simon et al. (1994)
	TONYA	GAA GTT TAT ATT TTA ATT TTA CCG GG	Monteiro and Pierce (2001)
	HOBBES	AAA TGT TGN GGR AAA AAT GTT A	Monteiro and Pierce (2001)
28S	28S-01	GAC TAC CCC CTG AAT TTA AGC AT	Choong-Gon et al. (2000)
	28SR-01	GAC TCC TTG GTC CGT GTT TCA AG	Choong-Gon et al. (2000)

[a] With modifications.

conducted through Sequencing Analysis (ABI) software and the new sequences generated in this study were deposited in GenBank (see Table 1 for accession numbers and voucher information). Unlike the sequences of coding genes (i.e., Cyt *b* and COI), the sequences of ribosomal genes (i.e., 12S and 28S) presented some variations in length. Their alignment was performed using ClustalX (Thompson et al., 1997) with default option settings. The alignment produced by ClustalX was then reviewed by eye in Seaview (Galtier et al., 1996). The resulting combined data set (2945 bp in length) was deposited to Treebase under accession number SN2588-10051. No significant base composition heterogeneity was detected between taxa for the four genes (12S: $\chi^2 = 43.11$, df = 84, $P = 0.99$; Cyt *b*: $\chi^2 = 50.12$, df = 84, $P = 0.99$; COI: $\chi^2 = 44.81$, df = 84, $P = 0.99$; 28S: $\chi^2 = 4.82$, df = 84, $P = 1.00$). With gaps treated as fifth position, 569 positions were informative under parsimony (see Table 3 for more detailed information on the molecular data set).

2.3. Phylogenetic analyses and hypothesis testing

Parsimony (MP) and Bayesian inference (BI) methods were used to reconstruct phylogenetic relationships among taxa. Among the four possible outgroups, we used *P. cardo* as an outgroup for all analyses. This choice was based on results from previous studies (Kergoat and Silvain, 2004), as well as on both morphological and paleontological data that indicate a basal position of tribe Pachymerini within the subfamily Bruchinae (Borowiec, 1987; Kingsolver, 1965; Poinar, 1999, 2005). In addition, an analysis (not figured) of an extended data set (with multiple specimens of the same species) was performed. No species-level paraphyly was detected for the five species for which additional specimens (from distinct localities) were included.

2.3.1. Maximum parsimony

All MP analyses were performed using PAUP* version 4.0b10 (Swofford, 2003). Heuristic searches were conducted using tree-bisection-reconnection (TBR) branch swapping, 1000 random-addition replicates, and a MaxTrees's value of 500. To test the heterogeneity between the four genes, we used the incongruence length difference test (ILD; Farris et al., 1994), as implemented in PAUP*, with all invariant characters excluded (Cunningham, 1997). Since the result of the partition-homogeneity test was not significant ($P > 0.05$), we chose to perform an analysis of the combined data set. The latter approach was preferred over separate analyses for two reasons: (i) all gene sequences were not obtained for all species, thus limiting the scope of separate analyses; (ii) in absence of data heterogeneity, adding in more data from distinct sources generally increase phylogenetic accuracy estimates (Bull et al., 1993; Huelsenbeck et al., 1996; Soltis et al., 1998; Wheeler et al., 1993), even if several sequences are missing (Wiens, 1998, 2003 and

Table 3
Distribution of invariant (INV) and parsimony informative (PI) characters among the four genes (with gaps treated as fifth position)

	Length	INV sites	% INV	PI sites	% PI	A–T bias
12S	402	260	64.67	60	14.92	76.84
Cyt *b*	782	490	62.65	206	26.34	69.48
COI	1018	670	65.81	276	27.11	67.01
28S	743	656	88.29	27	3.63	41.58
First positions (Cyt *b* + COI)	600	477	79.50	79	13.16	56.91
Second positions (Cyt *b* + COI)	600	563	93.83	16	2.66	60.98
Third positions (Cyt *b* + COI)	600	119	19.83	387	64.50	85.94
Stem regions (12S + 28S)	627	517	82.45	43	6.85	55.67
Loop regions (12S + 28S)	518	401	77.41	44	8.49	60.93
All sites	2945	2076	70.49	569	19.32	64.22

especially Wiens, 2006 for a review of this issue). Relative support of nodes for MP analyses was assessed by non-parametric bootstrap (Felsenstein, 1985a) procedures (1000 pseudoreplicates of 100 random-addition replicates were used), as implemented in PAUP*. Here, we have considered the nodes supported by bootstrap values $\geqslant 70\%$ as strongly supported following Hillis and Bull (1993). In addition, Bremer support (BS; Bremer, 1988, 1994) and partitioned Bremer support (PBS) values (Baker and DeSalle, 1997) were estimated, using TreeRot version 2.0 (Sorenson, 1999). Given the lack of clear statistical interpretations for the BS (Debry, 2001), a somewhat arbitrarily threshold ($BS \geqslant 4$; Felsenstein, 1985b) was used to identify well-supported nodes using BS values.

2.3.2. Bayesian inference

BI analyses were carried out using MrBayes version 3.11 (Huelsenbeck and Ronquist, 2001). For data sets consisting of multiple genes, the use of partition-specific models of evolution is advocated (Nylander et al., 2004; Yang, 1996), as it increases the fit of the evolutionary models with the data. By allowing subsets of the data (e.g., codon positions) to evolve under distinct models and parameters, an increase in both phylogenetic accuracy and posterior probability estimates is expected. However, the choice of partitions can be problematic, as countless partitioning strategies are envisageable. In addition, the use of smaller partitions increases the risk of random error associated with the estimation of model parameters (Brandley et al., 2005). Here, we follow the view of several authors (Brandley et al., 2005; Nylander et al., 2004) who propose the use of the Bayes factor (B_F) as an objective criterion to choose among several partitioning strategies in partitioned BI analyses. The Bayes factor is given by the ratio of the harmonic means of the likelihoods (sampled from the posterior) of the two analyses (respectively H_0 and H_1) in competition (Brandley et al., 2005; Nylander et al., 2004). Harmonic means of the likelihoods can be estimated by using the *sump* option in MrBayes (with the burnin period specified). In the study of Brandley et al. (2005), a fixed threshold was used to determine whether a given strategy was better than another (i.e., $2 \ln (B_F \geqslant 10)$; see also Kass and Raftery, 1995 for more details). At the end of their study, the former authors have nonetheless indicated that this criterion of $2 \ln (B_F \geqslant 10)$ was likely not stringent enough because all the observed positive values were generally well above this value. Here, we propose the use of a more stringent threshold which takes into account the difference in number of parameters between each competing strategy, in a similar way to the likelihood ratio test (LRT) statistic. In our study, degrees of freedom are equal to the numbers of additional parameters which are required by the most complex strategies. When comparing two strategies (H_0 and H_1), this variable (i.e., the number of additional parameters of the more complex strategy) is used to determine the critical value of the χ^2 distribution test statistic from standard statistical tables (with $\alpha = 0.05$). In addition, for comparisons involving strategies with the same number of parameters, we chose to use in a more conservative way the lowest critical value found in the statistical tables (i.e., 3.84). The alternative partitioning strategy (H_1) is rejected if the value of $2 \ln (B_F)$ is above the critical value corresponding to the estimated degree of freedom. Ultimately, the optimal strategy will be the strategy not rejected in any comparison and with the fewest number of partitions (to limit the risk of random error). For this study, we have compared eight partitioning strategies (summarized in Table 4), for which partitions were defined with reference to gene identity (12S, Cyt *b*, COI and 28S), codon positions (for the coding genes) and secondary structures (for the ribosomal genes). To identify stem and loop regions, secondary structure models (Clark et al., 1984; Gillespie et al., 2004; Page, 2000) were used. Results of the above strategies will be further compared and discussed to see if an increase in resolution and branch support is obtained through the use of specific partitions. Best-fit models of evolution for each gene of the combined data set were determined by using the Akaike information criterion (AIC), as implemented in Modeltest version 3.06 (Posada and Crandall, 1998). Since the results from the AICs indicated that the GTR+I+G model (Gu et al., 1995; Yang, 1994) was the best-fit model for all genes, this model was applied to each data partition. Two independent BI runs were carried out to identify whether convergence of clade posterior probabilities has been reached (Huelsen-

Table 4
Partitioning strategies used in this study

Partitioning strategy	Definition
P1	Non-partitioned data set
$P4_a$	12S + Cyt *b* + COI + 28S
$P4_b$	Cyt *b* + COI + stems (12S + 28S) + loops (12S + 28S)
$P5_a$	12S + 28S + lst pos. (Cyt *b* + *COI)* +2nd pos. (Cyt *b* + *COI)* + 3rd pos. (Cyt *b* + *COI)*
$P5_b$	Stems (12S + 28S) + loops (12S + 28S) +lst pos. (Cyt *b* + *COI)* + 2nd pos. (Cyt *b* + *COI)* +3rd pos. (Cyt6 + COI)
P6	Cyt *b* + COI + stems (12S) + loops (12S) + stems (28S) + loops (28S)
P8	12S + 28S + lst pos. (Cyt *b*) +2nd pos. (Cyt *b*) + 3rd pos. (Cyt *b*) + lst pos. (COI) +2nd pos.(COI) +3rd pos.(COI)
P10	Stems (12S) + loops (12S) + stems (28S) + loops (28S) +lst pos. (Cyt *b*) +2nd pos. (Cyt *b*) +3rdpos. (Cyt *b*) +lst pos. (COI) +2nd pos. (COI) +3rd pos. (COI)

beck et al., 2002; Miller et al., 2002). Each run consisted of four Markov chains (with incremental heating) of 2×10^6 generations, with random starting trees, default priors and trees sampled every 100 generations (branch lengths were also saved). A burn-in period of 1×10^5 generations was defined for all BI runs (stationarity was assessed graphically, by plotting likelihood scores against generations of the chains). For each partitioning strategy, BI results were generated using the pooled tree samples from the stationary phases of the two independent runs. Support of nodes for BI analyses was given by clade posterior probability (CPP) estimates. Since recent studies have suggested that Bayesian posterior probabilities are less conservative than non parametric bootstrap values, especially for short internodes (Alfaro et al., 2003; Erixon et al., 2003), only clades with posterior probabilities $\geqslant 90\%$ were considered as well supported in BI analyses.

2.3.3. Hypothesis testing

A priori hypotheses (i.e., the monophyly of each *Bruchus* taxonomic group) were compared statistically with a posteriori phylogenetic hypotheses (i.e., the trees obtained through MP and BI analyses). Here, we have chosen to use the likelihood-based nonparametric Shimodaira–Hasegawa test (SH; Shimodaira and Hasegawa, 1999), because it can be properly applied to compare a priori and a posteriori hypotheses (Buckley et al., 2001; Goldman et al., 2000). The constrained trees (in which *Bruchus* taxonomic groups were monophyletic) were built using Treeview version 1.66 (Page, 2001). For both a priori and a posteriori hypotheses, branch lengths were further reestimated in PAUP* using a GTR+I+G model. The reestimated log likelihoods (RELL) method (Kishino et al., 1990), as implemented in PAUP*, was used to resample the log likelihoods (1000 replicates) in the SH tests.

2.4. Host–plant information

As emphasized by numerous authors (e.g., Delobel and Delobel, 2003; Jermy and Szentesi, 2003; Johnson et al., 2004), available literature on bruchines often includes doubtful host–plant records. Misidentifications and non rigorous observations in the field (e.g., when catching adults on various flowering plants with the assumption that these plants are their host–plants) are generally responsible for these erroneous records. When possible, it is thus preferable to use data from studies in which host–plant associations are determined by extensive sampling of potential host–plant seeds in the field and further monitoring of adult emergences. Here, we have based our study on unequivocal data from three studies in which adults were reared from pods (Delobel and Delobel, 2003, 2005; Jermy and Szentesi, 2003). In addition, we have performed a critical examination of the major review of Lukjanovitch and Ter-Minasian (1957) to complete the host–plant records of most species (Table 5). Accurate additional host–plant records were also found in Anton (1998) and Morimoto (1990), for *B. altaicus* and *B. loti*, respectively. When necessary, host–plant names from the literature were updated by using the International Legume Database and Information Services database (ILDIS; http://www.ildis.org). Detailed information on host–plant taxonomy were also included using Kupicha (1983) and the Germplasm Resources Information Network database (GRIN; http://www.ars-grin.gov). Regarding taxonomy, *Cracca* PETERM. 1847 was used as a correct name of the subgenus following Jaaska (2005). Despite an exhaustive review of the existing literature, we were unable to recover host–plant records for three species (*B. brisouti*, *B. canariensis* and *B. sibiricus*) for which host–plants are still unknown. The latter finding underlines the fact that the majority of studies on the evolution of host–plant associations in phytophagous insects have to rely on potentially incomplete information, and caution must be therefore be taken to avoid hasty conclusions in such studies.

2.5. Data categorization and ancestral state estimation

To examine the evolution of host–plant associations in *Bruchus*, host–plant ancestral character states were mapped onto the seed-beetle phylogeny using distinct hierarchical taxonomical levels.

2.5.1. Data categorization

Since all species of *Bruchus* are known to feed on plants belonging to the tribe Vicieae (some species have also been reported to feed on the phylogenetically related tribe Cicereae, but reliable records are still lacking), various subtribal levels were used to investigate possible conservatisms of host–plant use on lower taxonomic levels. From a review of recent systematic studies (Jaaska, 2005; Kenicer et al., 2005; Steele and Wojciechowski, 2003), the following assumptions were made on tribe Vicieae phylogenetic relationships: (i) the monophyly of the tribe Vicieae is strongly supported by both molecular and morphological data; (ii) genera *Lathyrus*, *Lens* and *Pisum* are monophyletic, whereas preliminary analyses (Steele and Wojciechowski, 2003) suggest that the genus *Vicia* is paraphyletic with respect to the other genera in tribe Vicieae; (iii) within *Lathyrus*, current subgeneric (subgen. *Lathyrus* and *Orobus*) and sectional classifications are mostly retrieved with some notable exceptions (e.g., the placement of the section *Lathyrostylis*); (iv) within *Vicia*, while most of the species do cluster in two clades which correspond to the two subgenera (i.e., *Cracca* and *Vicia*), there is no evidence of a strong support of current sectional classifications (with the exception of some sections). With reference to the above information, first we have categorized host–plant information at the genus level. The genus *Vicia* was coded as a single character, in spite of its possible paraphyletic status. Five character states were thus used: (i) genus *Lathyrus*; (ii) genus *Lens*; (iii) genus *Pisum*; (iv) genus *Vicia*; (v) non-Vicieae genera. Second, we have categorized host–plant information at the subgenus level (for the genera

Table 5
Bruchus host–plant records

Plant species	Subgenus	Section	*Bruchus* species[a]
Lat. annuus L.	*Lathyrus*	*Lathyrus*	*tristiculus*[3]
Lat. aphaca L.	*Orobus*[*]	*Aphaca*	*laticollis*[3,5], *tristiculus*[2,5]
Lat. cicera L.	*Lathyrus*	*Lathyrus*	*luteicornis*[2], *rufimanus*[2], *tristis*[2] *tristiculus*[3,5]
Lat. clymenum L.	*Lathyrus*	*Clymenum*	*brachialis*[2], *tristiculus*[2]
Lat. digitatus (M.Bieb.)Fiori	*Lathyrus*[*]	*Lathyrostylis*	*viciae*[3]
Lat. grandiflorus Sibth.&Smith	*Lathyrus*	*Lathyrus*	*affinis*[3]
Lat. hirsutus L.	*Lathyrus*	*Lathyrus*	*tristiculus*[2,3,4]
Lat. japonicus Willd.	*Orobus*	*Orobus*	*loti*[6]
Lat. latifolius L.	*Lathyrus*	*Lathyrus*	*affinis* [2,3,4], *atomarius*[4], *tristiculus*[5]
Lat. linifolius (Reich.)Bässler	*Orobus*	*Orobus*	*dentipes*[3]
Lat. niger (L.) Bernh.	*Orobus*	*Orobus*	*atomarius*[4], *viciae*[4,5]
Lat. nissolia L.	*Lathyrus*	*Nissolia*	*loti*[4], *tristiculus*[2]
Lat. odoratus L.	*Lathyrus*	*Lathyrus*	*tristiculus*[3,4]
Lat. palustris L.	*Orobus*	*Orobus*	*atomarius*[3]
Lat. Pannonicus (Jacq.)Garcke	*Lathyrus*[*]	*Lathyrostylis*	*atomarius*[4], *viciae*[4]
Lat. pratensis L.	*Orobus*[*]	*Pratensis*	*affinis* [4], *atomarius*[5], *loti*[2,4], *viciae*[2,4]
Lat. sphaericus Retz.	*Lathyrus*	*Linearicarpus*	*tristiculus* [2,3], *viciae*[2,3]
Lat. sylvestris L.	*Lathyrus*	*Lathyrus*	*affinis*[3,4,5], *atomarius*[4]
Lat. tuberosus L.	*Lathyrus*	*Lathyrus*	*affinis* [3,4], *altaicus*[1], *atomarius*[5], *loti*[5]
Lat. venetus (Miller)Wohlf.	*Orobus*	*Orobus*	*rufimanus*[3]
Lat. vernus (L.) Bernh.	*Orobus*	*Orobus*	*atomarius*[4,5], *loti*[5]
Lens culinaris Medik.			*lentis*[2], *signaticornis*[5]
Pisum sativum L.			*pisorum*[3,4,5]
Vic. bithynica (L.)L.	*Vicia*	*Bithynicae*	*rufimanus*[3]
Vic. cassubica L.	*Cracca*	*Cassubicae*	*atomarius*[4]
Vic. cracca L.	*Cracca*	*Cracca*	*atomarius*[5], *hamatus*[5], *libanensis*[4], *luteicornis*[2], *occidentalis*[3,4], *rufipes*[2], *signaticornis*[3,5], tristiculus[2], venustus[2,4,5]
Vic. dumetorum L.	*Cracca*	*Vicilla*	*atomarius*[5]
Vic. faba L.	*Vicia*	*Faba*	*atomarius*[5], *dentipes*[1,5], *rufimanus*[3,5]
Vic. grandiflora Scop.	*Vicia*	*Vicia*	*luteicornis*[4]
Vic. hybrida L.	*Vicia*	*Hypechusa*	*ruflmanus*[3]
Vic. hyrcanica Fisch.&C.Mey.	*Vicia*	*Hypechusa*	*dentipes*[1,5]
Vic. lutea L.	*Vicia*	*Hypechusa*	*dentipes*[2,3], *ruflmanus*[2,3,4]
Vic. monantha Retz.	*Cracca*	*Cracca*	*rufipes*[3], *signaticornis*[5]
Vic. narbonensis L.	*Vicia*	*Narbonensis*	*rufimanus*[3]
Vic. onobrychioides L.	*Cracca*	*Pedunculatae*	*ruflmanus*[2]
Vic. pannonica Crantz	*Vicia*	*Hypechusa*	*brachialis*[3,5], *ruflmanus*[2,4]
Vic. parviflora Cav.	*Cracca*	*Ervum*	*griseomaculatus*[5]
Vic. peregrina L.	*Vicia*	*Peregrinae*	*emarginatus*[2,3], *ruflmanus*[2]
Vic. pisiformis L.	*Cracca*	*Vicilla*	*atomarius*[4,5]
Vic. pubescens (DC.)Link	*Cracca*	*Ervum*	*brachialis*[2]
Vic. sativa L.	*Vicia*	*Vicia*	*atomarius*[5], *brachialis*[3], *dentipes*[5], *luteicornis*[2,3,4,5] *rufipes*[2,3,5]
Vic. sepium L.	*Vicia*	*Atossa*	*atomarius*[2,3,4,5] *rufipes*[5], *luteicornis*[5]
Vic. sparsi flora Ten.	*Cracca*	*Cassubicae*	*atomarius*[4]
Vic. tenuifolia Roth	*Cracca*	*Cracca*	*atomarius*[2], *brachialis*[2,4], *libanensis*[4], *occidentalis*[4], *venustus*[4]
Vic. tetrasperma (L.)Schreber	*Cracca*	*Ervum*	*griseomaculatus*[2,3,5], *rufipes*[2]
Vic. villosa Roth	*Cracca*	*Cracca*	*brachialis*[2,4,5], *rufimanus*[3], *rufipes*[3,5]

Detailed host–plant systematic is given for genera *Lathyrus* and *Vicia* (the changes in taxonomic nomenclature made in this study are indicated by asterisks). [a] Numbers in this column refer to the following articles: (1) Anton (1998); (2) Delobel and Delobel (2003); (3) Delobel and Delobel (2005); (4) Jermy and Szentesi (2003); (5) Lukjanovitch and Ter-Minasian (1957); (6) Morimoto (1990).

Lathyrus and *Vicia* only). Given the fact that extant classifications of *Lathyrus* exhibit major discrepancies at the subgenus level (Asmussen and Liston, 1998), we have used the results of Kenicer et al. (2005) to redefine the two *Lathyrus* subgenera. Here, we have considered a subgenus *Lathyrus* which includes sections *Clymenum*, *Lathyrus*, *Lathyrostylis*, *Linearicarpus*, *Neurolobus* and *Nissolia*, and a subgenus *Orobus* which includes the sections *Aphaca*, *Notolathyrus*, *Orobus* and *Pratensis* (clade A in Kenicer et al., 2005). Six character states were used: (i) subgen. *Lathyrus*; (ii) subgen. *Orobus*; (iii) subgen. *Cracca*; (iv) subgen. *Vicia*; (v) other Vicieae; (vi) non-Vicieae. It was also critical to

deal with the issue of optimization of multiple associations, as several species of *Bruchus* were able to develop on plants belonging to distinct genera and/or subgenera. This finding underlines the fact that extant patterns of host–plant association may result from a progressive expansion of host range (Kergoat et al., 2005a). Interestingly, the corresponding *Bruchus* species were generally strongly associated with a specific genera and/or subgenera. Consequently, we chose to only consider the majority host–plant genera and/or subgenera in the coding of the corresponding character states. Although being not entirely satisfactory, this treatment of data was preferred over alternative methods (e.g., Janz and Nylin, 1998; Wahlberg, 2001) which generally involve parsimonious optimizations and have their own bias (Lopez-Vaamonde et al., 2003). Optimizations at the section level were not performed because *Bruchus* species with more than one host–plant were generally able to feed on plants belonging to distinct sections without exhibiting strong preferences for specific sections. In addition, the data categorization of *Bruchus* host–plant information at the section level is a problematic issue since several sections are likely paraphyletic (Asmussen and Liston, 1998; Steele and Wojciechowski, 2003).

2.5.2. Ancestral state estimation

Maximum-likelihood (ML) models were used to infer ancestral character states because, in sheer contrast with MP optimizations, likelihood-based optimizations can take into account branch lengths (more changes are expected on long branches if branch lengths are time proportional), and they allow the assessment of uncertainty in ancestral trait reconstruction (Belshaw and Quicke, 2002; Pagel, 1999; Schluter et al., 1997). For comparison purposes, we conducted global and local ML optimizations (Pagel, 1999), using Mesquite version 1.06 (Maddison and Maddison, 2005) and Multistate version 0.8 (Pagel, 2003), respectively. In global optimizations, ancestral character states at all nodes are reconstructed using a sole estimate for the parameters of the model of evolution whereas local optimizations estimate parameters separately for each possible ancestral character states at a node (Mooers and Schluter, 1999; Pagel, 1999). As a consequence, local optimizations are supposed to outperform global optimizations as they maximize likelihoods and result in the best fit of the model (Pagel, 1999). Given that both programs require trees with no missing data and given the fact that host–plant records were unavailable for three *Bruchus* species, a subset of our original data set (with *B. brisouti*, *B. canariensis* and *B. sibiricus* excluded) was analyzed. The analysis of this data set was carried out through BI (to include branch length estimates), by using the partitioning scheme which has been considered as optimal in previous analyses of the complete data set. For all analyses, we considered that the support of one state over another (at a given node) was significant if the difference between their log-likelihoods was superior or equal to 2.0 (Schluter et al., 1997; Pagel, 1999).

The first global optimizations were performed using the one-parameter Markov k-state model (Mk1; Lewis, 2001), as implemented in Mesquite. In this generalization of the Jukes–Cantor model, the rates of change parameter are constrained to be equal (Maddison and Maddison, 2005). Subsequently, we have used Multistate to carry out additional global optimizations using more complex models which are detailed below. The choice of a model of trait evolution (and the estimation of the associated transition rates) is indeed a critical issue because complex models generally require a lot of data (i.e., large trees) to provide accurate ancestral character state estimates (Mooers and Schluter, 1999; Schluter et al., 1997). For both global and local optimizations under Multistate, we thus followed the view of several authors (Mooers and Schluter, 1999; Pagel, 1999) who advocated the use of simpler models when possible (i.e., if the latter do not lead to a significant reduction in the likelihood). To study the evolution of character traits with n states, up to $n\ (n-1)$ parameters can be estimated through Multistate (Pagel, 2003). Consequently, we have used default models with 20 parameters (for the first character with five states) and 30 parameters (for the second character with six states) in some of our optimizations (i.e., local optimizations which are abbreviated as L20 and L30). Since no significant reduction in the likelihood was found when constraining forward and backward rates to be equal, simpler models (with twice less parameters; 10 or 15) were used in addition to the default models with a view to compare their respective results. For global optimizations these models were abbreviated as G10 and G15 whereas L10 and L15 were used to name the models which were used in local optimizations.

Finally, permutation tail probability tests (PTP; Faith and Cranston, 1991), as implemented in PAUP*, were performed as an alternative way to investigate whether host–plant association is correlated with phylogeny of *Bruchus*. host–plant association character states were randomized across the tips of the phylogeny 10,000 times. Within-character randomization was only applied to the ingroup taxa (i.e., *Bruchus* species) to avoid misleading PTP scores (Trueman, 1996). The resulting frequency distribution of tree lengths was then used to estimate whether the observed tree length was significantly shorter than expected under a random model (Maddison and Slatkin, 1991). Following Kelley and Farrell (1998), we have also performed additional PTP tests by adding to the phylogeny two non sequenced species for which accurate host–plant records are known (*Lens culinaris* for *B. ervi* and *Vicia cirrhosa* for *B. hierroensis*). Their respective placements on the existing phylogeny were supported by numerous morphological characters which indicate close relationships of *B. ervi* with *B. lentis* (Borowiec, 1988; Hoffmann, 1945), and of *B. hierroensis* with *B. canariensis* (Decelle, 1979).

3. Results

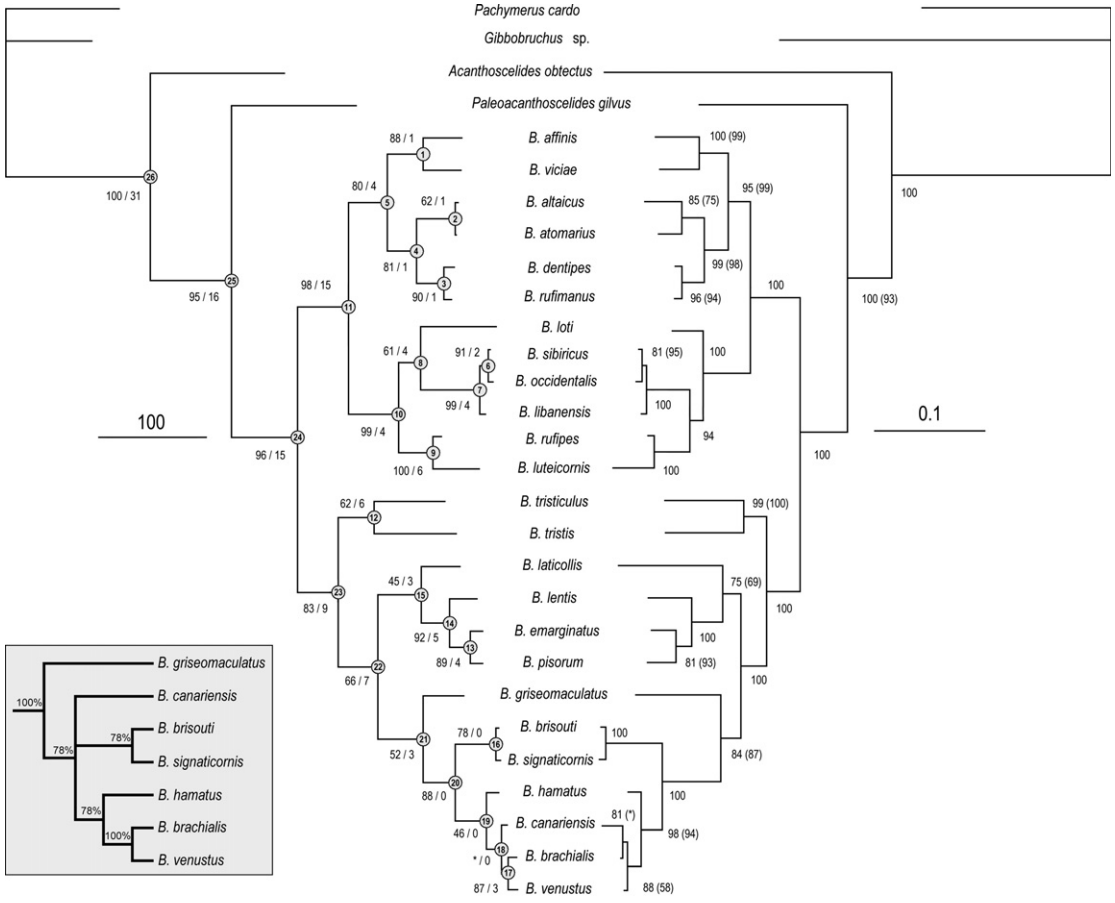
3.1. Phylogenetic analyses and hypothesis testing

3.1.1. Maximum parsimony

The analysis of the combined data set yielded nine most-parsimonious trees (2215 steps; CI = 0.534; RI = 0.593) that differed among themselves only in the position of *B. canariensis* (one of the nine most-parsimonious trees is shown on Fig. 1). On average, MP trees are well supported by bootstrap values (bootstrap $\geqslant$70% for 18 of the 26 nodes), whereas BS values provided a lesser support (BS $\geqslant$4 for 14 of the 26 nodes). All basal nodes are strongly supported, and the genus *Bruchus* is recovered monophyletic with a high support (bootstrap of 96%, BS of 15). Within *Bruchus* species, groups *affinis*, *atomarius*, *pisorum*, and *tristis* are recovered monophyletic, whereas groups *brachialis* and *rufipes* are found paraphyletic. The examination of PBS values indicated a low level of conflicting data (only 3 of the 104 values were negatives), in accordance with the result of the ILD test. Almost all positive values came from the three mitochondrial genes, whereas a negligible contribution of the 28S gene is suggested by the PBS values. The latter result could be likely accounted for by missing sequences and the relative low number of PI characters of this gene (as indicated in Table 3).

3.1.2. Bayesian inference

For each partitioning strategy, the two independent runs converged on similar likelihood scores and reached stability around 4 to 5×10^5 generations. According to the Bayes factor criterion, the most complex strategy (i.e., involving the greatest number of partitions) was optimal (Table 6). Interestingly, partition-rich strategies were not always the best ones, since in some cases less complex strategies have performed better (i.e., $P5_a$ vs P6, $P5_b$ vs P6 and $P5_a$ vs P8).



Nodes	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
COI	0	0	0	0	0	0	0	0.5	0	0	6	0.56	2	4.5	1	0	0	0	0	0	1	6.5	4	0	-3	1
Cytochrome *b*	0	0	0	0	0	2	0	0	0	0	3	2.75	0	0	2	0	3	0	0	0	2	0	3	0	12	28
12s rRNA	1	1	1	1	4	0	4	3.5	6	4	7	2.69	2	0.5	0	0	0	0	0	0	0	1	2	15	6	2
28s rDNA	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	-0.5	0	0	1	0

Fig. 1. Phylogenetic relationships of *Bruchus* species. The tree on the left corresponds to one of the nine most-parsimonious trees (2215 steps; CI = 0.534; RI = 0.593) from the parsimony analysis of the combined data set. The tree of the right corresponds to the result of the partitioned BI analyses conducted using the most optimal strategy (P10). For the MP tree, numbers at nodes indicate both bootstrap values (left) and BS values (right). In addition PBS values are given for each node, on the bottom left of the figure (all nodes are labelled accordingly). On the left, a pruned consensus tree of the nine most-parsimonious trees is also figured. For the BI tree, numbers at nodes indicate the CPP values of the P10 strategy. Additional values (under bracket) correspond to the CPP values of an alternative strategy ($P4_a$). Identical CPP were recovered for both strategies when no additional values are given.

Table 6
Comparisons of all partitioning strategies using Bayes factors

Harmonic Mean	Mean CPP	H_1 H_0	PI	$P4_a$	$P4_b$	$P5_a$	$P5_b$	P6	P8	P10
14111.94	92.61	P1*	—	50.99	50.99	65.17	65.17	79.08	106.39	133.25
13872.10	88.62	$P4_a$*	479.68	—	3.84	21.02	21.02	36.41	65.17	92.80
13991.13	89.19	$P4_b$*	241.62	−238.06	—	21.02	21.02	36.41	65.17	92.80
13246.22	89.58	$P5_a$	1731.44	1251.76	1489.82	—	3.84	21.02	50.99	79.08
13347.11	89.12	$P5_b$	1529.66	1049.98	1288.04	−201.78	—	21.02	50.99	79.08
13837.57	88.31	P6*	548.74	69.06	307.12	−1182.70	−980.92	—	36.41	65.17
13228.75	90.04	P8	1766.38	1286.70	1524.76	34.94	236.72	1217.64	—	36.41
13193.53	90.62	P10	**1836.82**	**1357.14**	**1595.20**	**105.38**	**307.16**	**1288.08**	**70.44**	—

2 ln (Bp) values are figured on the left side of the data matrix: bold values indicate the 2 ln (B_F) comparisons used in determining the optimal partitioning strategy whereas italic values indicate comparisons in which partition-rich strategies are rejected in favor of less complex ones. Critical values of the χ^2 distribution are figured on the right side of the data matrix. In addition, for each partitioning strategy, both mean clade posterior probabilities (CPP) and harmonic mean -ln L values are given. Four partitioning strategies indicated by asterisks yield a similar alternative topology.

The four analyses with the highest mean likelihood scores (i.e., $P5_a$, $P5_b$, P8 and P10) yielded the same topology (Fig. 1). A very similar topology (not shown), with an alternative placement of *B. canariensis*, was recovered by the four other partitioning strategies. BI topologies were mostly congruent with MP trees, only differing in the positions of *B. loti* and *B. canariensis* (not in all MP trees for the latter species). Furthermore, the similarity between the results of both inference methods was also statistically supported by non significant SH tests (e.g., $P = 0.463$ for the two topologies shown in Fig. 1). Overall, the BI topology corresponding to the optimal partitioning strategy (P10) is well supported (CPP $\geqslant$90% for 19 of the 26 nodes). Likewise to MP trees, basal nodes are strongly supported (CPP of 100% for most basal nodes), and a monophyletic genus *Bruchus* is recovered with a strong support (100%). Regarding the monophyly of *Bruchus* taxonomic groups, the same groups (groups *brachialis* and *rufipes*) are also found paraphyletic under BI. Unexpectedly, the non-partitioned analysis (P1) yielded the highest arithmetic mean CPP value (92.61), followed by the three best strategies (mean CPP of 90.62, 90.04 and 89.58 for P10, P8 and $P5_b$, respectively; see Table 5 for details).

3.1.3. Hypothesis testing

Constrained trees were built to specifically address the monophyly of groups *brachialis* and *rufipes*. SH tests failed ($P = 0.274$ and $P = 0.304$ when using unconstrained MP and BI trees, respectively) to reject the alternative hypothesis of a monophyletic group *brachialis*. On the contrary, SH tests significantly rejected the alternative hypothesis of a monophyletic group *rufipes* ($P < 0.001$ for both tests).

3.2. Ancestral state estimation

3.2.1. Global vs local ML optimizations

For more clarity, focus has been given to the results of the global optimizations performed using a simple model (i.e., the Mk1 model), and the results of the local optimizations performed by constraining forward and backward rates to be equal (i.e., L10 and L15 optimizations). The results of the other optimizations (i.e., G10 and G15: global optimizations with 10- and 15-parameters; L20 and L30: local optimizations with 20- and 30-parameters) will be further discussed in the text. Overall, a similar evolution of host–plant associations is suggested by the two methods of optimization, both at the host–plant genus (Figs. 2 and 3) and subgenus level (Figs. 4 and 5), as the global estimates are generally consistent with the local estimates (see Supplementary material for details). This finding is supported by a highly significant Pearson's correlation which is found between the estimates provided by the two methods ($R = 0.91$, $P < 0.001$ at the genus level; $R = 0.85$, $P < 0.001$ at the subgenus level). Nonetheless, some discrepancies are noticeable when comparing the results of the two optimization methods. First, in contrast with local optimizations, global optimizations with a Mk1 model yield ambiguous and puzzling ancestral character state values at the base of the tree (nodes 24, 23 and 22). As opposed to the character states exhibited by the outgroups, a very low probability is found for the preference for non-Vicieae host–plants in the three deepest nodes. Second, the two methods yield contradictory results for some nodes in which the most likely ancestral character states are different (nodes 5, 10 and 21 for the first character, and node 17 for the second character), thus suggesting distinct patterns of evolution. Finally, a greater number of significantly supported ancestral characters are recovered by global ML optimizations using a Mk1 model (nine vs five), when only considering the ingroup taxa (i.e., the *Bruchus* species).

The use of more complex models (with 10- and 15-parameters) in global optimizations result in reconstructions which are between those obtained previously: (i) intermediate values are recovered for the probabilities associated with each ancestral character states; (ii) at the base of the tree, the ancestral character state values for two of the three nodes (nodes 24 and 23) are consistent with the result of the previous local optimizations (with a high probability associated to non-Vicieae feeding); (iii) a greater number of significantly supported ancestral characters (nine) is recovered by the these global optimizations in comparison with previous local optimizations, when only considering the

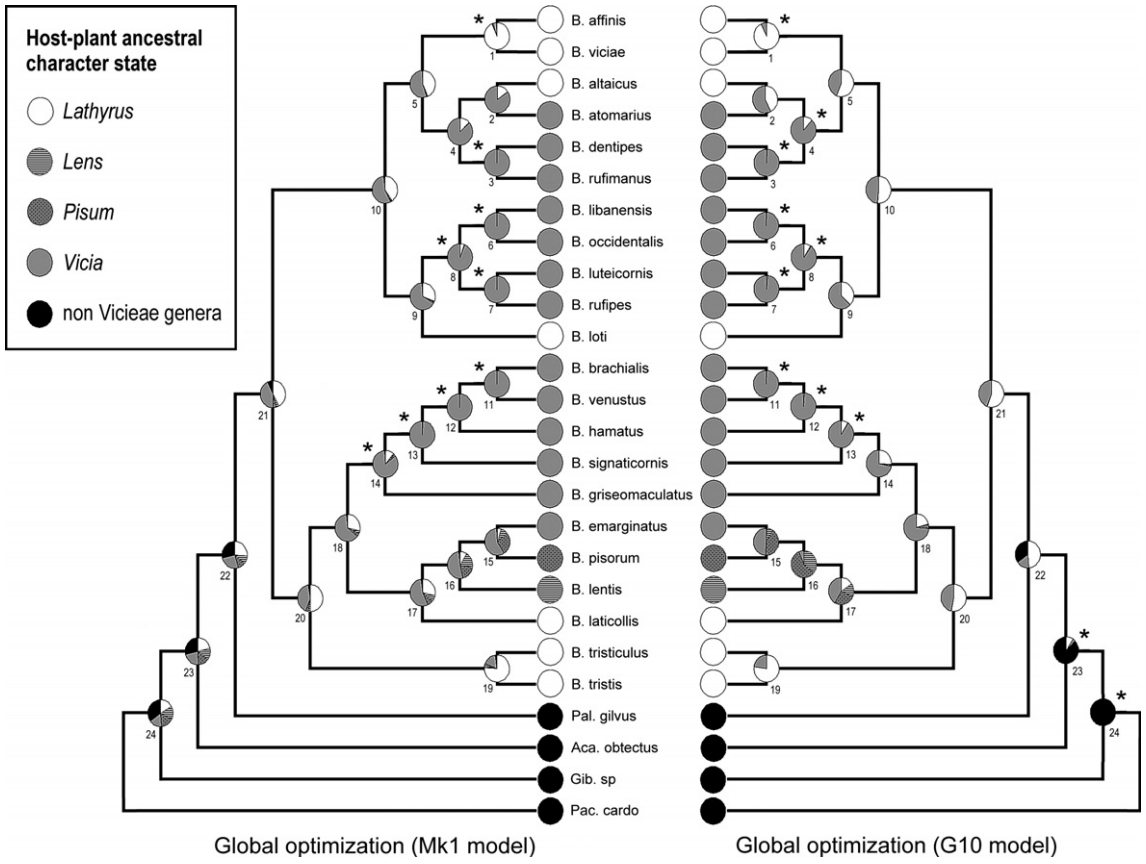


Fig. 2. Mirror image of the ML global optimizations at the genus level (using both the topology and branch lengths obtained under BI). On the left cladogram, ancestral character states are reconstructed under Mesquite using global optimizations and a Mk1 model. On the right cladogram, ancestral character states are reconstructed under Multistate using global optimizations and a 10-parameter model (G10 model). Probabilities of character states are figured at the nodes with pie diagrams (see Supplementary material for detailed values). Asterisks indicate nodes with significantly supported character states.

ingroup taxa. As expected a more important correlation is also found between the previous local estimates and the newer global estimates ($R = 0.95$ at the genus level and $R = 0.93$ at the subgenus level). In general, the corresponding evolutionary pattern is in agreement with those obtained using the methods detailed beforehand. In contrast, it is not the case for the local optimizations performed using 20- and 30-parameter models. For the latter analyses, we found numerous counterintuitive values which obviously overestimate the probabilities of having a preference toward genera *Lens* and *Pisum*. For example, at the host–plant genus level, a summed probability of 36.57% is found for the node 14 (for which the corresponding clade does not include species which feed on *Lens* spp. or *Pisum* spp.) whereas unexpectedly high probabilities are also found for some nodes at the subgenus level (e.g., see nodes 2 and 9 in Fig. 5). In addition, puzzling values are recovered at the base of the tree (node 24) for both optimizations (host–plant genus and subgenus level; see Figs. 3 and 5).

3.2.2. Host–plant genus level

The mapping of the evolution of host–plant associations does not recover a clear pattern at the basal and intermediate levels of the *Bruchus* phylogeny (Figs. 2 and 3), as most values are not statistically supported. On the other hand, in most terminal levels of the tree there is some evidence for a trend toward taxonomic conservatism, as related species generally share the same ancestral character states (with significant statistical support). Species belonging to groups *affinis* and *tristis* are thus clearly associated with *Lathyrus* spp. whereas species belonging to groups *atomarius*, *brachialis* and *rufipes* are generally associated with *Vicia* spp. Having said that, various loss and gain events are also suggested in several cases (e.g., in the clade which groups *B. laticollis*, *B. lentis*, *B. emarginatus* and *B. pisorum*), thus indicating a more dynamic pattern. For the default data set, no significant phylogenetic signal was recovered by the PTP test at the host–plant genus level ($P = 0.268$). Interestingly, a nearly significant value ($P = 0.052$) was found for the data set with a larger taxon sampling.

3.2.3. Host–plant subgenus level

As expected given the results of the previous optimizations, the assignment of ancestral character states was unclear at the basal and intermediate level of the tree (Figs. 4 and 5). Regarding the terminal level of the tree, the same

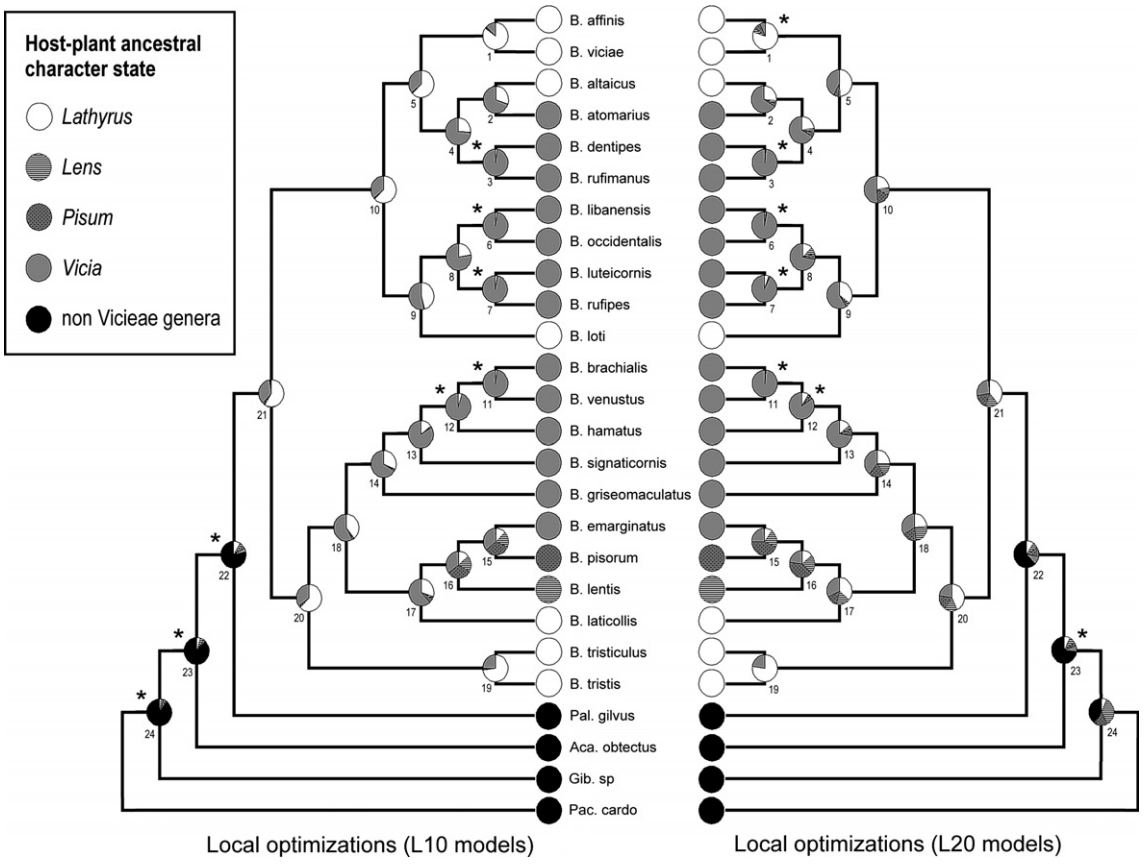


Fig. 3. Mirror image of the ML local optimizations at the genus level (using both the topology and branch lengths obtained under BI). On the left cladogram, ancestral character states are reconstructed under Multistate using local optimizations (the FOSSIL command was used) and 10-parameter models (L10). On the right cladogram, ancestral character states are reconstructed under Multistate using local optimizations (the FOSSIL command was used) and 20-parameter models (L20). Probabilities of character states are figured at the nodes with pie diagrams (see Supplementary material for detailed values). Asterisks indicate nodes with significantly supported character states.

nodes are found significantly supported by the comparison of likelihood scores. It is interesting to note that a trend toward taxonomic conservatism is still suggested at the subgenus level, because most of the closely-related *Bruchus* species are found to feed on plants belonging to the same subgenus (with the obvious exceptions of the species exclusively associated with *Lens* and *Pisum*). At this hierarchical level of optimization, the PTP tests yield significant values for both data sets ($P = 0.016$ for the default data set and $P = 0.002$ for the extended data set).

4. Discussion

4.1. Methodological issues

4.1.1. Partitioned analyses

Several useful findings can be drawn from the results of the partitioned analyses of the combined data set. For our data set, only the use of codon positions in partitioning strategies has systematically led to a significant increase of the mean likelihood scores (and presumably of phylogenetic accuracy as well). In contrast, using only the secondary structure of ribosomal genes (strategies $P4_b$ and P6) did not lead to such an increase and yielded a presumably suboptimal topology. Interestingly, these observations are consistent with those obtained in the study of Brandley et al. (2005). Collectively, these empirical results therefore suggest that partitioned analyses which use codon positions may likely outperform analyses which use standard ''one partition per gene'' or ''secondary structure-based'' strategies. Finally, our partitioned analyses yield a somewhat counterintuitive result, as the highest mean CPP value was recovered by the non-partitioned analysis (P1). A likely explanation can be found in Nylander et al. (2004) who have suggested that there is general tendency for oversimplified models to be associated with excessive credibilities in topologies that may not be correct.

4.1.2. Global vs local ML optimizations

With the exception of the most parameter-rich optimizations, namely the local optimizations with 20- and 30-parameters, global and local optimizations recover similar patterns of evolution of host–plant preferences, at both the genus and subgenus level. Nonetheless, due to their incorrect assessments of character state estimates for basal nodes, global optimizations appear as less reliable estima-

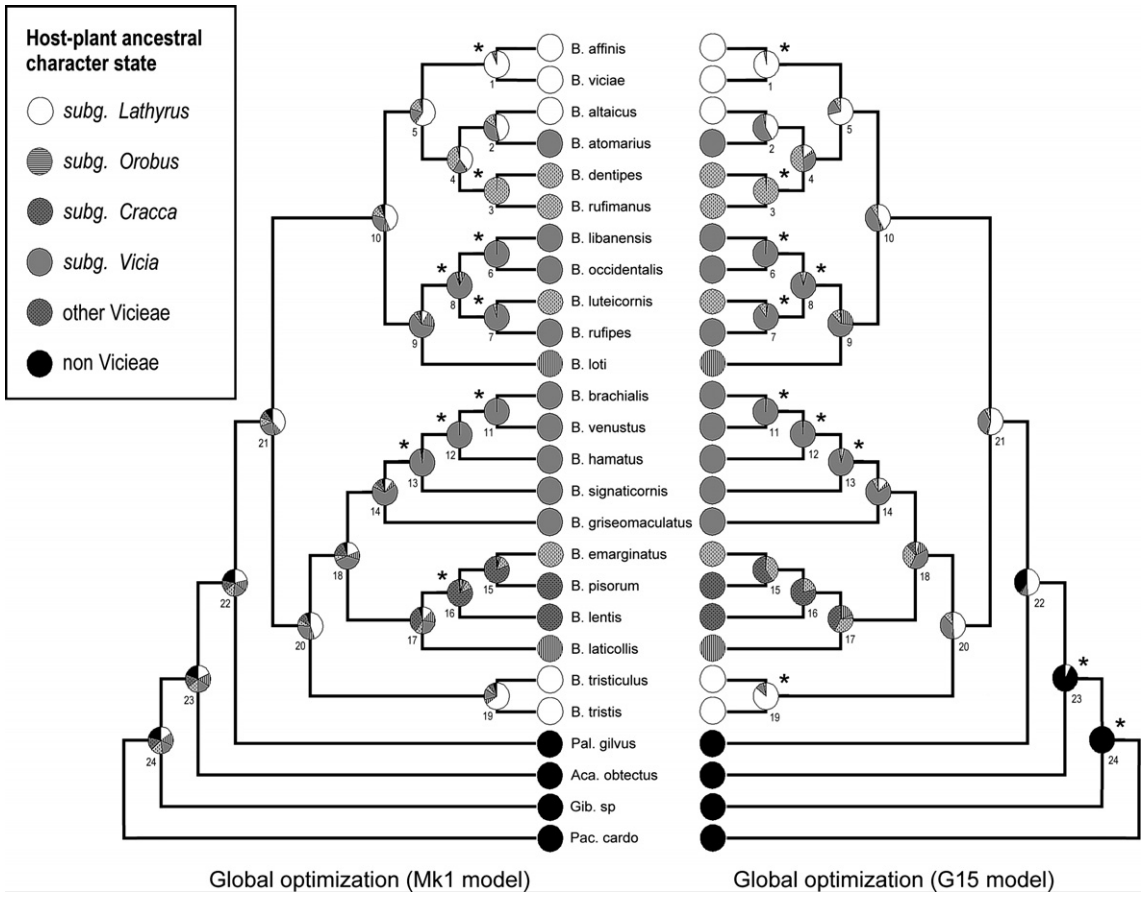


Fig. 4. Mirror image of the ML global optimizations at the subgenus level (using both the topology and branch lengths obtained under BI). On the left cladogram, ancestral character states are reconstructed under Mesquite using global optimizations and a Mk1 model. On the right cladogram, ancestral character states are reconstructed under Multistate using global optimizations and a 15-parameter model (G15 model). Probabilities of character states are figured at the nodes with pie diagrams (see Supplementary material for detailed values). Asterisks indicate nodes with significantly supported character states.

tors in comparison with local optimizations, in agreement with Pagel (1999). The counterintuitive results which were obtained using optimizations with 20- and 30-parameters can be likely accounted for by an over-parameterization. Since parameter-rich models obviously require more data per parameter, the use of models with 20- and 30-parameters for our data set was certainly limited by the number of terminal taxa which were used (Mooers and Schluter, 1999).

4.2. *Bruchus* systematics

4.2.1. Phylogenetic relationships within *Bruchus*

As detailed before, both MP and BI analyses yield very similar and well-supported phylogenetic hypotheses which are mostly consistent with the taxonomic groupings of Borowiec (1988). Although tree topologies from each analysis differ in some details (i.e., the positions of *B. loti* and *B. canariensis*), the differences are in no case statistically significant. Under MP, the position of *B. loti* within group *rufipes* is rather weakly supported (bootstrap of 61%, BS of 4). In contrast, BI analyses recover a more basal position for this species (i.e., a sister group relationship with a clade composed of four members of group *rufipes*) with high support (CPP of 100%). Since there is no strong support (under MP) for the inclusion of *B. loti* within members of group *rufipes*, and since this alternative position is not found to be statistically significant, we are more inclined to favor the results of BI analyses regarding the phylogenetical placement of *B. loti*. The position of *B. canariensis* is not clearly resolved in our analyses (although an apical position within group *brachialis* is suggested), not only in MP (the nine most-parsimonious trees are not in agreement for the placement of this species), but also in BI (results of distinct partitioned analyses only differ in the position of *B. canariensis*). This lack of resolution can likely be accounted for by the missing sequences for this species. Under BI, the four most optimal strategies indicate a close and relatively well-supported relationship of *B. canariensis* with *B. brachialis* (CPP of 81%, 78, 80 and 76% for P10, P8, $P5_b$ and $P5_a$, respectively) whereas the four less optimal strategies suggest a sister-group relationship of *B. canariensis* with the clade which groups *B. brachialis* and *B. venustus* (CPP of 63, 63, 62 and 35% for P6, $P4_b$, $P4_a$ and P1, respec-

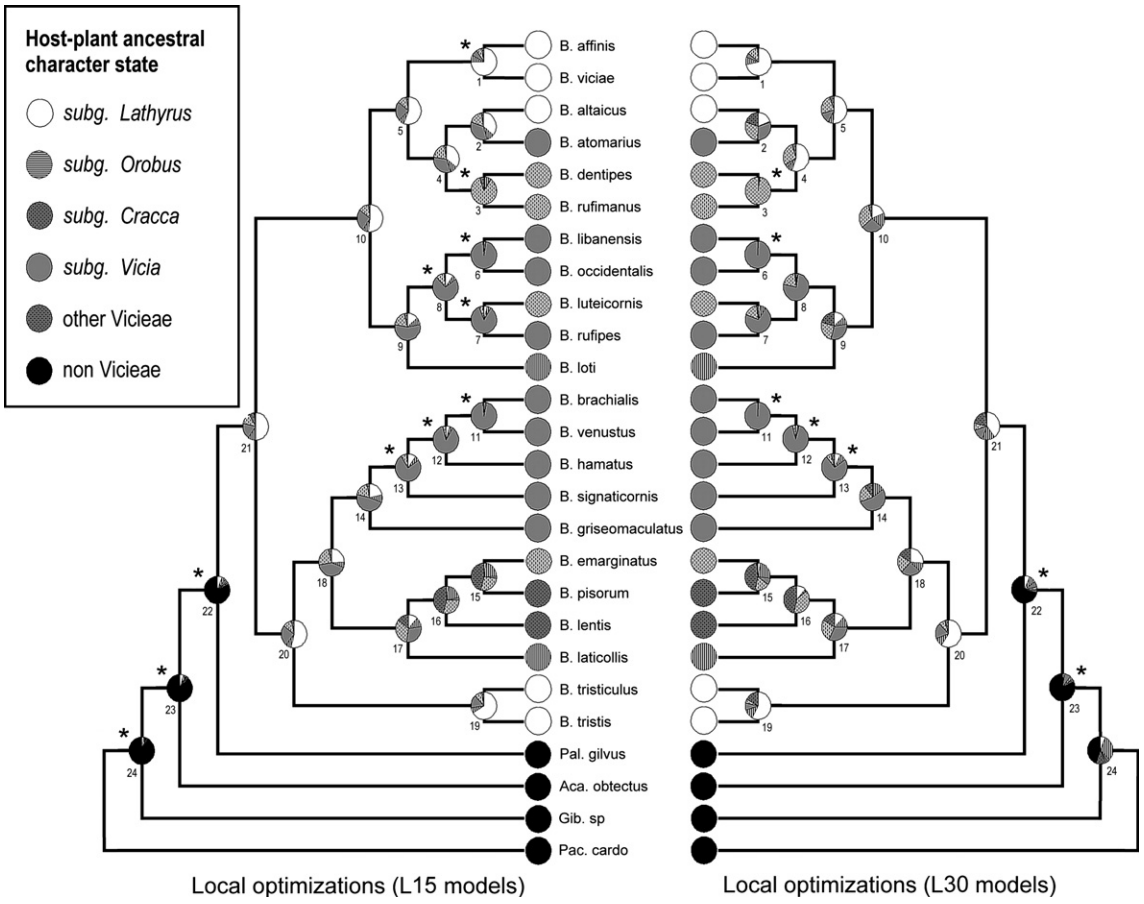


Fig. 5. Mirror image of the ML local optimizations at the subgenus level (using both the topology and branch lengths obtained under BI). On the left cladogram, ancestral character states are reconstructed under Multistate using local optimizations (the FOSSIL command was used) and 15-parameter models (L15). On the right cladogram, ancestral character states are reconstructed under Multistate using local optimizations (the FOSSIL command was used) and 30-parameter models (L30). Probabilities of character states are figured at the nodes with pie diagrams (see Supplementary material for detailed values). Asterisks indicate nodes with significantly supported character states.

tively). The same two alternative positions were also recovered in several of the most-parsimonious trees, also supporting in a convincing way the supposed apical position of *B. canariensis* within group *brachialis*.

4.2.2. Taxonomic groups

Our phylogenetical analyses (see Fig. 1) strongly support the monophyly of groups *affinis*, *atomarius*, *pisorum* and *tristis* as currently defined (Borowiec, 1988). Unfortunately, no conclusions can be drawn on the status of group *loti*, because we were not able to obtain sequences for *B. lugubris* (therefore *B. loti* was the sole representative of the group *loti* in our molecular data set). In our phylogenetical analyses *B. laticollis* appears as the sister-species of members of group *pisorum*, thus suggesting a paraphyletic group *brachialis*. Nonetheless this placement is weakly supported (bootstrap of 45%, BS of 3 and CPP of 68%) and not statistically significant according to the SH tests. As a consequence, we still favor the null hypothesis of a monophyletic group *brachialis*. Another group whose monophyly is questioned by our results is the group *rufipes*. This group is rendered paraphyletic by the position of *B. griseomaculatus* (under MP and BI) and the position of *B. loti* (under MP only). As underlined in the precedent paragraph, the inclusion of *B. loti* within group *rufipes* was not supported in a convincing way (bootstrap of 61%, SH test not significant). On the contrary, constraining a monophyletic group *rufipes* by moving *B. griseomaculatus* results in a highly significant SH test ($P < 0.001$). The latter result was not so surprising, because some morphological evidences (e.g., differences in the shape of parameres, absence of a characteristic sclerite in the distal part of the internal sac) have already suggested that *B. griseomaculatus* is somewhat unrelated to other members of group *rufipes*. Based on both molecular and morphological evidences, we therefore reject the monophyly of group *rufipes* as currently defined. *B. griseomaculatus* is also not closely related to members of group *brachialis* because it does not possess the specific features of this group, like the enlarged fore tibiae in males. Since this species is also morphologically quite distinct from members of other extant taxonomic groups, we propose to assign this species to a group of its own (with the obvious consequence of recovering a monophyletic group *rufipes*).

4.3. Limits of taxonomic conservatism

In this study, we have investigated the limit of taxonomic conservatism in a specialized genus of phytophagous beetle, finding evidence for a trend toward taxonomic conservatism at both host–plant genus and subgenus level. However, this trend is not so obvious when examining the reconstruction of ancestral character states under ML. In order to discuss that discrepancy, we have to take into account two important factors. First, as underlined in previous studies (Pagel, 1999; Morse and Farrell, 2005), the difference in 2.0 log units (which was used to determine whether a character state was significantly supported or not) appears as a very conservative criterion in our analyses. As a consequence several nodes with a high support (e.g., as high as 85.83%) in a given character state were not significantly supported in the ML optimizations (see the Supplementary material for details). Second, as illustrated by the results of the PTP tests, the inclusion of non sequenced species for which host–plants are accurately known (i.e., *B. ervi* and *B. hierroensis*) will likely increase the discernible trend toward taxonomic conservatism in the various ML optimizations. It is also important to note that though the results of ML optimizations were ambiguous for deeper nodes, they nonetheless suggested an ancestral association with either genera *Lathyrus* or *Vicia*. Having said that, our analyses have been likely influenced (in both directions) by the way we have treated multiple host–plant associations. Indeed, in our treatment of these associations, the information associated with the wider host-range exhibited by some species was lost (e.g., for *B. atomarius*). The latter observation underlines the fact that taxonomic conservatism in host–plant use is not the sole important feature in the evolution of host–plant associations in *Bruchus*. Host shifts are likely under multiple evolutionary constraints and must be rather viewed as progressive processes in which some species are able to expand and/or reduce their host-range over-time (Bernays, 1998). While we did not test specifically for a trend toward so-called generalist or specialist species in *Bruchus*, preliminary analyses (not shown) have suggested that there was no clear apical distribution for neither specialists (here defined as species which were associated with a sole genus) nor generalist species.

4.4. Factors influencing Bruchus host–plant associations

The results of our study are interesting in the light of determining which factors better explain the observed patterns of host–plant associations in the *Bruchus*—Vicieae model. Indeed, our finding of a trend toward taxonomic conservatism below the host–plant tribe level suggests the influence of strong constraints on the evolution of *Bruchus* host–plant associations. Although the genus as a whole is specialized on plants from the tribe Vicieae, each *Bruchus* species is thus restricted to a given set of host–plants. Several factors may be advocated to explain the far from random pattern of host–plant associations in the *Bruchus*—Vicieae model.

4.4.1. Host-selection behavior

In bruchines, host-selection behaviors are likely decisive to understand why potential hosts are not fed upon when present (Siemens et al., 1991). In *Bruchus*, the host-selection behavior tends to be determined by the females' oviposition behavior rather than visual cues (e.g., N'Diaye and Labeyrie, 1990). It has been also shown that females are sensitive to deterrent chemical stimuli when selecting an oviposition site (Annis and O'Keeffe, 1984; Jermy and Szentesi, 1978). As phylogenetically related plants likely share more similar chemical compounds, we can thus suppose that both host range and potential host shift of *Bruchus* species are likely influenced by the evolution of the females' chemoreception system. In a recent study, Jermy and Szentesi (2003) suggested that the evolution of host specialization (and host-switches) in seed-beetles may result primarily from the evolution of the nervous system (with reference to chemoreception). While we agree with them in recognizing the importance of host-selection behavior (see also the recent review of Bernays in 2001), we also think that the lack of related experimental studies (on both the evolution of the females' chemoreception system and the nature of plant chemicals of young pods) do not allow the assessment of the relative importance of this factor in the evolution of host–plant use in *Bruchus*.

4.4.2. Host-suitability

First of all, and as underlined by Szentesi and Jermy (1995), host-suitability for *Bruchus* is limited by seed morphology. Since all *Bruchus* species develop to adults within a single seed, they are not able to develop in flat or very small seeds in contrast with other species of seed-beetles which are able to feed on several seeds. Second, multiple plant defense mechanisms are also involved which undoubtedly influence host-suitability. For instance, in response to oviposition or egg hatch, several species of *Lathyrus* are known to stimulate cell divisions and callus development on pods to impede the larval development (Annis and O'Keeffe, 1984). Interestingly, this unique form of induced resistance is specifically mediated by a novel class of natural products, the bruchins, which have been found up to now only in seed-beetles (Doss et al., 1997). In this spectacular example, the plants have apparently developed a specific defense versus their seed-beetle predators. Finally, numerous chemical defenses (e.g., non-protein and pyrimidine amino acids, protease inhibitors) with well-known or supposed toxic effects on the development of seed-beetle larvae (e.g., Bleiler and Rosenthal, 1988; Huignard et al., 1996; Janzen et al., 1977) are found in the seed-coats and in the cotyledons of the seeds. The non-protein amino acid L-canavanine is found in several *Vicia* species within the subgenus *Cracca* (Bell et al., 1978), whereas Bowman–Birk inhibitors are found in all Vicieae genera (e.g., Weder and Kahleyss, 1998). Regarding pyrimidine amino acids, three distinct compounds are found: (i) lathyrine in species of *Lathyrus* and in several *Lens* species excluding *Lens culinaris* (Bell, 1962; Rozan

et al., 2001); (ii) vicine and convicine in several *Vicia* species within the subgenus *Vicia* (Ramsay and Griffiths, 1996); (iii) willardiine and isowillardiine in *P. sativum* (Brown and Turan, 1995). Collectively, these defenses have been shown to be effective against the majority of seed-beetles, as only very few species outside the genus *Bruchus* are able to develop on Vicieae seeds (Johnson, 1981; Kergoat et al., 2005a). Since all these defenses are not uniformly distributed throughout Vicieae (e.g., several toxic compounds are only found is specific subgenera), we can suppose that their absence/presence play a decisive role in determining *Bruchus* host–plant specificity, restricting host–plant use and limiting potential host shifts (thus accounting for the observed pattern of taxonomic conservatism). The resulting specialization in host–plant use will likely occur because of evolutionary trade-offs (Cornell and Hawkins, 2003): a species that excels in bypassing a given defense (e.g., using detoxifying pathways) will conversely lose the ability to bypass other defenses as well.

4.4.3. Perspectives

While experimental data on the determinism of host–plant selection behavior in seed-beetles are still lacking, a few studies have already provided meaningful insights with reference to genetic determinism in host-suitability. For instance in the seed-beetle *Callosobruchus maculatus*, Huignard et al. (1996) have demonstrated that the ability to develop in seeds with high level of vicine and convicine was under the control of a major dominant gene which controls the activity of a β-glucosidase. In the absence or non-activity of this β-glucosidase (in recessive homozygous individuals), vicine and convicine are not hydrolyzed in a toxic aglycone, thus permitting the adaptation of some individuals to *V. faba*. Nevertheless, more physiological, genetical and behavioral researches are definitely required to draw any general conclusion about the mechanisms which could be accounted for the patterns of host–plant associations we observed for *Bruchus* and more generally for many seed-beetles we have studied.

Acknowledgments

We are very grateful to all who helped use obtain some of the specimens which were used in this study. We also wish to thank L. Borowiec for the information on *Bruchus* taxonomic groups. G.J.K. thanks M.-L. Cariou and M. Takagi for laboratory support. This research was partially supported by a postdoctoral grant (Grant No. PE05017) from the Japan Society for the Promotion of Science (JSPS) for G.J.K.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2006.11.026.

References

Alfaro, M.E., Zoller, S., Lutzoni, F., 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol. Biol. Evol. 20, 255–266.

Annis, B., O'Keeffe, L.E., 1984. Response of two *Lathyrus* species to infestation by the pea weevil *Bruchus pisorum* L. (Coleoptera: Bruchidae). Entomol. Exp. Appl. 35, 83–87.

Anton, K.-W., 1998. Results of the Czechoslovak-Iranian entomological expeditions to Iran 1970, 1973 and 1977 Coleoptera Bruchidae. Cas. Nar. Muzea Rada Prirod. 167, 73–90.

Anton, K.-W., 1999. Two new species of the *Bruchus brachialis* group from the Mediterranean region (Coleoptera: Bruchidae: Bruchinae). Linzer Biol. Betr. 31, 655–660.

Anton, K.-W., 2001. Bemerkungen zur faunistik und taxonomie mitteleuropäischer samenkäfer (Coleoptera: Bruchidae). Folia Entomol. Hungarica 62, 43–49.

Arora, G.L., 1977. Taxonomy of the Bruchidae (Coleoptera) of Northwest India. Orient. Insects Suppl. 7, 1–132.

Asmussen, C., Liston, A., 1998. Chloroplast DNA characters, phylogeny and classification of *Lathyrus* (Fabaceae). Am. J. Bot. 85, 962–969.

Baker, R.H., DeSalle, R., 1997. Multiple sources of character information and the phylogeny of the Hawaiian drosophilids. Syst. Biol. 46, 645–673.

Becerra, J.X., 1997. Insects on plants: macroevolutionary chemical trends in host use. Science 276, 253–256.

Becerra, J.X., Venable, D.L., 1999. Macroevolution of insect–plant associations: the relevance of host biogeography to host affiliation. Proc. Natl. Acad. Sci. USA 22, 12626–12631.

Bell, E.A., 1962. Association of ninhydrin-reacting compounds in the seeds of 49 species of *Lathyrus*. Biochem. J. 83, 225–229.

Bell, E.A., Lackey, J.A., Polhill, R.M., 1978. Systematic significance of canavanine in the Papilioinoideae (Faboideae). Biochem. Syst. Ecol. 6, 201–212.

Belshaw, R., Quicke, D.L.J., 2002. Robustness of ancestral character state estimates: evolution of life history strategy in ichneumonoid parasitoids. Syst. Biol. 51, 450–477.

Bernays, E.A., 1998. Evolution of feeding behaviour in insect herbivores. Bioscience 48, 35–44.

Bernays, E.A., 2001. Neural limitations in phytophagous insects: implications for diet breadth and evolution of host affiliation. Annu. Rev. Entomol. 46, 703–727.

Bleiler, J.A., Rosenthal, G.A., 1988. Biochemical ecology of canavanine-eating seed predators. Ecology 69, 427–433.

Borowiec, L., 1987. The genera of seed-beetles (Coleoptera, Bruchidae). Polsk. Pismo Entomol. 57, 3–207.

Borowiec, L., 1988. Bruchidae-Strakowce (Insecta: Coleoptera). Fauna Polski, tom 11. PWN, Warszawa.

Brandley, M.C., Schmitz, A., Reeder, T.W., 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Syst. Biol. 54, 373–390.

Bremer, K., 1988. The limits of amino-acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42, 795–803.

Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295–304.

Brown, E.G., Turan, Y., 1995. Pyrimidine metabolism and secondary product formation; biogenesis of albizziinen 4-hydroxyhomoarginine and 2,3-diaminopropanoic acid. Phytochemistry 40, 763–771.

Buckley, T.R., Simon, C., Shimodaira, H., Chambers, G.K., 2001. Evaluating hypotheses on the origin and evolution of the New Zealand alpine cicadas (Maoricicada) using multiple-comparison tests of tree topology. Mol. Biol. Evol. 18, 223–234.

Bull, J.J., Huelsenbeck, J.P., Cunningham, C.W., Swofford, D.L., Waddell, P.J., 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. 42, 384–397.

Clark, C.G., Tague, B.W., Ware, V.C., Gerbi, S.A., 1984. *Xenopus laevis* 28S ribosomal RNA: a secondary structure model and its evolutionary and functional implications. Nucleic Acids Res. 12, 6197–6220.

Choong-Gon, K., Hong-Zhang, Z., Imura, Y., Tominaga, O., Su, Z.-H., Osawa, S., 2000. Pattern of morphological diversification in the *Leptocarabus* ground beetles (Coleoptera: Carabidae) as deduced from mitochondrial ND5 gene and nuclear 28S rDNA sequences. Mol. Biol. Evol 17, 137–145.

Cornell, H.V., Hawkins, B.A., 2003. Herbivore responses to plant secondary compounds: a test of phytochemical coevolution theory. Am. Nat. 161, 507–522.

Cunningham, C.W., 1997. Can three incongruence tests predict when data should be combined? Mol. Biol. Evol. 14, 733–740.

Debry, R.W., 2001. Improving interpretation of the decay index for DNA sequence data. Syst. Biol. 50, 742–752.

Decelle, J., 1975. Les Bruchidae des îles Canaries. Bull. Ann. Soc. R. Entomol. Belg. 111, 109–142.

Decelle, J., 1979. Une nouvelle espèce de *Bruchus* (Coleoptera: Bruchidae) des îles Canaries. Vieraea 8, 143–146.

Delobel, A., Delobel, B., 2003. Les plantes-hôtes des bruches (Coleoptera, Bruchidae) de la faune de France, une analyse critique. Bull. mens. Soc. linn. Lyon 72, 199–221.

Delobel, A., Tran, M., 1993. Les Coléoptères des denrées alimentaires entreposées dans les régions chaudes. Faune tropicale XXXII. Orstom/CTA, Paris.

Delobel, B., Delobel, A., 2005. Les plantes-hôtes des bruches, errata et données nouvelles. Bull. mens. Soc. linn. Lyon 74, 277–291.

Dobler, S., Farrell, B.D., 1999. Host use evolution in *Chrysochus* milkweed beetles: evidence from behaviour, population genetics and phylogeny. Mol. Ecol. 8, 1297–1307.

Doss, R.P., Oliver, J.E., Proebsting, W.M., Potter, S.W., Kuy, S., Clement, S.L., Williamson, R.T., Carney, J.R., DeVilbiss, E.D., 1997. Bruchins: insect-derived plant regulators that stimulate neoplasm formation. Proc. Natl. Acad. Sci. USA 97, 6218–6223.

Duckett, C.N., Gillespie, J.J., Kjer, K.M., 2003. Relationships among the subfamilies of Chrysomelidae inferred from small subunit ribosomal DNA and morphology, with special emphasis on the relationship among the flea beetles and the Galerucinae. In: Jolivet, P., Schmitt, M., Santiago-Blay, J. (Eds.), New contributions in Chrysomelidae biology. SPB Academic Publishing, The Netherlands, pp. 3–18.

Ehrlich, P.R., Raven, P.H., 1964. Butterflies and plants: a study in coevolution. Evolution 18, 586–608.

Erixon, P., Svennblad, B., Britton, T., Oxelman, B., 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. Syst. Biol. 52, 665–673.

Faith, D.P., Cranston, P.S., 1991. Could a cladogram this short have arisen by chance alone? On permutation tests for cladistic structure. Cladistics 7, 1–28.

Farrell, B.D., 1998. "Inordinate fondness" explained: Why are there so many beetles? Science 281, 555–559.

Farrell, B.D., 2001. Evolutionary assembly of the milkweed fauna: cytochrome oxidase I and the age of *Tetraopes* beetles. Mol. Phylogenet. Evol. 18, 467–478.

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

Farrell, B.D., Sequeira, A.S., 2004. Evolutionary rates in the adaptive radiation of beetles on plants. Evolution 58, 1984–2001.

Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics 10, 315–319.

Felsenstein, J., 1985a. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Felsenstein, J., 1985b. Confidence limits on phylogenies with a molecular clock. Syst. Zool. 34, 152–161.

Futuyma, D.J., Keese, M.C., Scheffer, S., 1993. Genetic constraints and the phylogeny of insect-plant associations: responses of *Ophraella communa* (Coleoptera: Chrysomelidae) to host–plants of its congeners. Evolution 47, 888–905.

Futuyma, D.J., McCafferty, S.S., 1990. Phylogeny and the evolution of host–plant associations in the leaf beetle genus *Ophraella* (Coleoptera, Chrysomelidae). Evolution 44, 1885–1913.

Galtier, N., Gouy, M., Gautier, C., 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. 12, 543–548.

Gillespie, J., Cannone, J., Gutell, R., Cognato, A., 2004. A secondary structure model of the 28S rRNA expansion segments D2 and D3 from rootworms and related leaf beetles (Coleoptera: Chrysomelidae; Galerucinae). Insect Mol. Biol. 13, 495–518.

Goldman, N., Anderson, J.P., Rodrigo, A.G., 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. 49, 652–670.

Gu, X., Fu, Y.-X., Li, W.-H., 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. Mol. Biol. Evol. 12, 546–557.

Harry, M., Solignac, M., Lachaise, D., 1998. Molecular evidence for parallel evolution of adaptive syndromes in fig-breeding *Lissocephala* (Drosophilidae). Mol. Phylogenet. Evol. 9, 542–551.

Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a methods for assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182–192.

Huelsenbeck, J.P., Bull, J.J., Cunningham, C.W., 1996. Combining data in phylogenetic analysis. Trends Ecol. Evol. 11, 152–157.

Huelsenbeck, J.P., Larget, B., Miller, R.E., Ronquist, F., 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. Syst. Biol. 51, 673–688.

Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Huignard, J., Baehr, J.C., Desroches, P., Mandon, N., 1996. Adaptation of a *Callosobruchus maculatus* strain to *Vicia faba*, as its new host–plant. Entomol. Exp. Appl. 80, 156–159.

Huignard, J., Dupont, P., Tran, B., 1990. Coevolutionary relations between bruchids and their host–plants. The influence of the physiology of the insects. In: Fuji, K, Gatehouse, A.M.R., Johnson, C.D., Mitchel, R., Yoshida, T. (Eds.), Bruchids and Legumes: Economics, Ecology and Coevolution: Proceedings, Kluwer Academic Publishers, pp. 171–179.

Hoffmann, A., 1945. Coléoptères Bruchidae et Anthribidae. In: Lechevalier, P. (Ed.), Faune de France— 44, Paris, pp. 1–184.

Jaaska, V., 2005. Isozyme variation and phylogenetic relationships in *Vicia* subgenus *Cracca* (Fabaceae). Ann. Bot. 96, 1085–1096.

Janz, N., Nylin, S., 1998. Butterflies and plants: a phylogenetic study. Evolution 52, 486–502.

Janzen, D.H., Juster, H.B., Bell, E.A., 1977. Toxicity of secondary compounds to the seed-eating larvae of the bruchid beetle *Callosobruchus maculatus*. Phytochemistry 16, 223–227.

Jermy, T., Szentesi, A., 1978. The role of inhibitory stimuli in the choice of oviposition site by phytophagous insects. Entomol. Exp. Appl. 24, 258–271.

Jermy, T., Szentesi, A., 2003. Evolutionary aspects of host–plant specialization—a study on bruchids (Coleoptera: Bruchidae). OIKOS 101, 196–204.

Johnson, C.D., 1980. The use of host preferences as taxonomic characters of bruchid beetles (Coleoptera: Bruchidae) feeding in the seeds of *Cassia* (Leguminosae). J. Kansas Entomol. Soc. 53, 27–34.

Johnson, C.D., 1981. Seed beetle host specificity and the systematics of the Leguminosae. In: Polhill, R.M., Raven, P.H. (Eds.), Advances in Legume Systematics, The Royal Botanic Gardens, Kew, pp. 995–1027.

Johnson, C.D., Southgate, B.J., Delobel, A., 2004. A revision of the Caryedontini (Coleoptera: Bruchidae: Pachymerinae) of Africa and the Middle East. Mem. Am. Entomol. Soc. 44, 1–120.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–775.

Kawecki, T.J., Mery, F., 2003. Evolutionary conservatism of geographic variation in host preference in *Callosobruchus maculatus*. Ecol. Entomol. 28, 449–456.

Kelley, S.T., Farrell, B.D., 1998. Is specialization a dead end? The phylogeny of host-use in *Dendroctonus* bark beetles (Scolytidae). Evolution 52, 1731–1743.

Kenicer, G.J., Kajita, T., Pennington, R.T., Murata, J., 2005. Systematics and biogeography of *Lathyrus* (Leguminosae) based on internal transcribed spacer and cpDNA sequence data. Am. J. Bot. 99, 1199–1209.

Kergoat, G.J., Alvarez, N., Hossaert-McKey, M., Faure, N., Silvain, J.-F., 2005a. Parallels in the evolution of the two largest New and Old World seed-beetle genera (Coleoptera, Bruchidae). Mol. Ecol. 14, 4003–4021.

Kergoat, G.J., Delobel, A., Fédière, G., Le Rü, B., Silvain, J.-F., 2005b. Both host–plant phylogeny and chemistry have shaped the African seed-beetle radiation. Mol. Phylogenet. Evol. 35, 602–611.

Kergoat, G.J., Delobel, A., Silvain, J.-F., 2004. Phylogeny and host-specificity of European seed beetles (Coleoptera, Bruchidae), new insights from molecular and ecological data. Mol. Phylogenet. Evol. 32, 855–865.

Kergoat, G.J., Silvain, J.-F., 2004. Le genre *Bruchidius* (Coleoptera: Bruchidae) est-il monophylétique? Apports des méthodes de parcimonie, maximum de vraisemblance et inférence bayésienne. Biosystema 22, 113–125.

Kingsolver, J.M., 1965. A new fossil bruchid genus and its relationships to modern genera (Coleoptera: Bruchidae: Pachymerinae). Coleopt. Bull. 19, 25–30.

Kingsolver, J.M., 1995. On the family Bruchidae. Chrysomela Newsl. 30, 3.

Kishino, H., Miyata, T., Hasegawa, M., 1990. Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. J. Mol. Evol. 30, 151–160.

Kupicha, F.K., 1983. The infrageneric structure of *Lathyrus*. Notes Roy. Bot. Gard. Edinb. 41, 209–244.

Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913–925.

Lingafelter, A., Pakaluk, J., 1997. Comments on the Bruchinae and Chrysomelidae. Chrysomela Newsl. 33, 3–4.

Lopez-Vaamonde, C., Godfray, H.C.J., Cook, J.M., 2003. Evolutionary dynamics of host–plant use in a genus of leaf-mining moths. Evolution 57, 1804–1821.

Lukjanovitch, F.K., Ter-Minasian, M.E., 1957. Zhuki-zernovski (Bruchidae). Fauna SSSR, Zhestkokrylye, 24, 209 pp. Moscow.

Maddison, W.P., Maddison, D.R., 2005. Mesquite: a modular system for evolutionary analysis. Version 1.06 <http://mesquiteproject.org>.

Maddison, W.P., Slatkin, M., 1991. Null models for the number of evolutionary steps in a character on a phylogenetic tree. Evolution 45, 1184–1197.

Miller, R.E., Buckley, T.R., Manos, P.S., 2002. An examination of the monophyly of morning glory taxa using Bayesian phylogenetic inference. Syst. Biol. 51, 740–753.

Monteiro, A., Pierce, N.E., 2001. Phylogeny of *Bicyclus* (Lepidoptera: Nymphalidae) inferred from COI, COII, and Ef-1α gene sequences. Mol. Phylogenet. Evol. 18, 264–281.

Mooers, A.O., Schluter, D., 1999. Reconstructing ancestor states with maximum-likelihood: support for one- and two-rate models. Syst. Biol. 48, 623–633.

Morimoto, K., 1990. A synopsis of the bruchid fauna of Japan. In: Fuji, K, Gatehouse, A.M.R., Johnson, C.D., Mitchel, R., Yoshida, T. (Eds.), Bruchids and Legumes: Economics, Ecology and Coevolution: Proceedings, Kluwer Academic Publishers, pp. 131–140.

Morse, G.E., Farrell, B.D., 2005. Ecological and evolutionary diversification of the seed beetle genus *Stator* (Coleoptera: Chrysomelidae: Bruchinae). Evolution 59, 1315–1333.

N'Diaye, S., Fabres, G., Labeyrie, V., 1992. Modalités de la compétition larvaire intraspécifique chez *Bruchus affinis* (Coleoptera, Bruchidae) dans les graines de *Lathyrus sylvestris* (Leguminosae, Fabaceae). Bull. Soc. Entomol. Fr. 97, 135–144.

N'Diaye, S., Labeyrie, V., 1990. Etude de l'adaptation de *Bruchus affinis à Lathyrus sylvestris*: analyse de la mortalité avant l'installation des larves dans la graine. Entomol. Exp. Appl. 55, 195–204.

Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53, 47–67.

Odegaard, F., Diserud, O.H., Ostbye, K., 2005. The importance of plant relatedness for host utilization among phytophagous insects. Ecol. Lett. 8, 612–617.

Page, R.D.M., 2000. Comparative analysis of secondary structure of insect mitochondrial small subunit ribosomal RNA using maximum weighted matching. Nucleic Acids Res. 28, 3839–3845.

Page, R.D.M., 2001. TreeView. Version 1.66 <http://taxonomy.zoology.gla.uc.uk/rod/rod.html>.

Pagel, M., 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. Syst. Biol. 48, 612–622.

Pagel, M., 2003. Multistate. available via <www.ams.reading.ac.uk/zoology.pagel/>. School of animal and microbial sciences, University of Reading, Reading, UK.

Poinar Jr., G., 1999. A fossil palm bruchid, *Caryobruchus dominicanus* sp. n. (Pachymerini: Bruchidae) in Dominican amber. Entomol. Scand. 30, 219–224.

Poinar Jr., G., 2005. A cretaceous palm bruchid, *Mesopachymerus antiqua*, n. gen., n. sp. (Coleoptera: Bruchidae: Pachymerini) and biogeographical implications. Proc. Entomol. Soc. Wash. 107, 392–397.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Ramsay, G., Griffiths, D.W., 1996. Accumulation of vicine and convicine in *Vicia faba* and *V. narbonensis*. Phytochemistry 42, 63–67.

Reid, C., 1996. More on the family Bruchidae. Chrysomela Newsl. 31, 3.

Rozan, P., Kuo, Y.-H., Lambein, F., 2001. Amino acids in seeds and seedlings of the genus *Lens*. Phytochemistry 58, 281–289.

Scheffer, S.J., Wiegmann, B.M., 2000. Molecular phylogenetics of the holly leafminers (Diptera: Agromyzidae: *Phytomyza*): species limits, speciation, and dietary specialization. Mol. Phylogenet. Evol. 17, 244–255.

Schluter, D., Price, T., Mooers, A.O., Ludwig, D., 1997. Likelihood of ancestor states in adaptive radiation. Evolution 51, 1699–1711.

Schmitt, M., 1998. Again, bruchid classification. Chrysomela Newsl. 36, 3–4.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Siemens, D.H., Johnson, C.D., Woodman, R.L., 1991. Determinants of host range in bruchid beetles. Ecology 72, 1560–1566.

Silvain, J.-F., Delobel, A., 1998. Phylogeny of west African *Caryedon* (Coleoptera: Bruchidae): congruence between molecular and morphological data. Mol. Phylogenet. Evol. 9, 533–541.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann. Entomol. Soc. Am. 87, 651–702.

Smith, A.M., 1990. Pea weevil (*Bruchus pisorum* L.) and crop-loss implications for management. In: Fuji, K, Gatehouse, A.M.R., Johnson, C.D., Mitchel, R., Yoshida, T. (Eds.), Bruchids and Legumes: Economics, Ecology and Coevolution: Proceedings, Kluwer Academic Publishers, pp. 105–114.

Soltis, D.E., Soltis, P.S., Mort, M.E., Chase, M.W., Savolainen, V., Hoot, S.B., Morton, C.M., 1998. Inferring complex phylogenies using parsimony: an empirical approach using three large DNA data sets for Angiosperms. Syst. Biol. 47, 32–42.

Sorenson, M.D., 1999. TreeRot.v2.. Boston University, Boston.

Southgate, B.J., 1979. Biology of the Bruchidae. Annu. Rev. Entomol. 24, 449–473.

Steele, K.P., Wojciechowski, M.F., 2003. Phylogenetic analyses of tribes Trifolieae and Vicieae, based on sequences of the plastid gene *mat*K (Papilionoideae: Leguminosae). In: Klitgaard, K., Bruneau, A. (Eds.),

Advances in Legume Systematics, part 10, Higher Level Systematics. The Royal Botanic Gardens, Kew, pp. 355–370.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Szentesi, A., Jermy, T., 1995. Predispersal seed predation in leguminous species: seed morphology and bruchid distribution. OIKOS 73, 23–32.

Ter-Minasian, M.E., 1968. A new species of the genus *Bruchus* L. (Coleoptera: Bruchidae) from the fauna of USSR. Rev. Entomol. URSS 47, 181–183.

Termonia, A., Pasteels, J., Windosr, D.M., Milinkovitch, M.C., 2002. Dual chemical sequestration: a key mechanism in transitions among ecological specialization. Proc. R. Soc. Lond. B 269, 1–6.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Thompson, J.N., 1993. Preference hierarchies and the origin of geographical specialization in host use in swallowtail butterflies. Evolution 47, 1585–1594.

Tran, B., Darquenne, J., Huignard, J., 1993. Changes in responsiveness to factors inducing diapause termination in *Bruchus rufimanus* (Boh.) (Coleoptera: Bruchidae). J. Insect. Physiol. 39, 769–774.

Tran, B., Huignard, J., 1992. Interactions between photoperiod and food affect the termination of reproductive diapause in *Bruchus rufimanus* (Boh.), (Coleoptera, Bruchidae). J. Insect. Physiol. 38, 633–637.

Trueman, J.W.H., 1996. Permutation tests and outgroups. Cladistics 12, 253–261.

Tuda, M., Chou, L.-Y., Niyomdham, C., Buranapanichpan, S., Tateishi, Y., 2005. Ecological factors associated with pest status in *Callosobruchus* (Coleoptera: Bruchidae): high host specificity of non-pests to Cajaninae (Fabaceae). J. Stored Prod. Res. 41, 31–45.

Tuda, M., Ronn, J., Buranapanichpan, S., Wasano, N., Arnqvist, G., 2006. Evolutionary diversification of the bean beetle genus *Callosobruchus* (Coleoptera: Bruchidae): traits associated with stored-product pest status. Mol. Ecol. 15, 3541–3551.

Verma, K.K., Saxena, R., 1996. The status of Bruchidae as a family. Chrysomela Newsl. 32, 3.

Wahlberg, N., 2001. The phylogenetics and biochemistry of host–plant specialization in melitaeine butterflies (Lepidoptera: Nymphalidae). Evolution 55, 522–537.

Weder, J.K.P., Kahleyss, R., 1998. Isolation and characterisation of four trypsin-chymotrypsin inhibitors from lentil seeds. J. Sci. Food. Agric. 78, 429–434.

Wendt, H., 1993. *Bruchus ecalcaratus* K. Daniel, 1906—ein Synonym zu *Bruchus rufimanus* Boheman, 1833 (Coleoptera, Chrysomeloidea, Bruchidae). Deutsche Entomol. Zeit. 40, 161–165.

Wheeler, W.C., Cartwright, P., Hayashi, C.Y., 1993. Arthropod phylogeny: a combined approach. Cladistics 9, 1–39.

Wiens, J.J., 1998. Combining data sets with different phylogenetic histories. Syst. Biol. 47, 568–581.

Wiens, J.J., 2003. Missing data, incomplete taxa, and phylogenetic accuracy. Syst. Biol. 52, 528–538.

Wiens, J.J., 2006. Missing data and the design of phylogenetic analyses. J. Biomed. Inf. 39, 34–42.

Yang, Z., 1994. Estimating the pattern of nucleotide substitution. J. Mol. Evol. 39, 105–111.

Yang, Z., 1996. Maximum likelihood models for combined analyses of multiple sequence data. J. Mol. Evol. 42, 587–596.

Yotoko, K.S.C., Prado, P.I., Russo, C.A.M., Solferini, V.N., 2005. Testing the trends towards specialization in herbivore-host–plant associations using a molecular phylogeny of *Tomoplagia* (Diptera: Tephritidae). Mol. Phylogenet. Evol. 35, 701–711.

Zampetti, M.F., 1993. Una nuova specie di *Bruchus* del Libano (Coleoptera, Bruchidae). Fragm. Entomol. Roma 24, 215–218.