Molecular Phylogenetics and Evolution 78 (2014) 14-24

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev





Multi-locus DNA sequence variation in a complex of four leaf beetle species with parapatric distributions: Mitochondrial and nuclear introgressions reveal recent hybridization



Maud C. Quinzin*, Patrick Mardulyn

Evolutionary Biology and Ecology, Université Libre de Bruxelles, CP 160/12, av. F.D. Roosevelt 50, 1050 Brussels, Belgium

ARTICLE INFO

Article history: Received 12 November 2013 Revised 5 April 2014 Accepted 2 May 2014 Available online 10 May 2014

Keywords: Cold-adapted Species tree Gene trees Speciation Species delimitation

ABSTRACT

While the importance for including multiple independent loci in phylogeographic studies is largely acknowledged, a majority of these still focus on a single species. We combine the study of multilocus DNA sequence variation (one mitochondrial and four unlinked nuclear fragments) at both the interand intraspecific levels to explore the evolutionary history of four closely related specialized herbivorous insects (Chrysomelidae, genus *Gonioctena*). Analyzing the sequences for samples collected across their entire European range allows us to (1) characterize the genetic boundaries among species, i.e. the degree of lineage sorting, (2) infer their phylogenetic relationships and (3) explore reproductive barriers among them in regions where their ranges overlap. For two sister species, we identify multiple independent cases of undirectional transfer of genetic material (introgression) at both mitochondrial and nuclear loci, demonstrating recent hybridization between both species in the overlapping regions of their range. The highlighted pattern of genetic variation suggests that *Gonioctena intermedia* expanded its range into that of *G. quinquepunctata*, and that both species may experience mutual exclusion. Overall, this study illustrates that interpreting intraspecific genetic variation for the purpose of evolutionary inference without the broader context of the other closely related species could lead to erroneous conclusions.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

DNA sequence variation among individuals has been extensively explored in the last two decades to infer phylogenetic relationships among species and to explore the evolution of different organismal features such as morphology, behavior, or physiology (e.g., Losos, 1998; Bossuyt and Milinkovitch, 2000; Sota and Ishikawa, 2004). DNA sequence data have also been used in separate studies to investigate patterns of variation at the intraspecific level, thought to reflect evolutionary histories of populations (e.g., Avise, 2000). The need to analyze several independent loci has arisen in both types of study. In the field of phylogenetic inference, it has long been known that a gene tree does not necessarily correspond to the species tree (Pamilo and Nei, 1988; Nichols, 2001). More recently, it has been shown that phylogeographic patterns may vary widely among loci, even in the absence of selection, due to the stochastic nature of gene transmission across generations (e.g., Hare, 2001; Rosenberg and Nordborg, 2002). Accordingly, to test any given hypothesis, corroborative evidence is required from independent loci; multilocus data sets are therefore becoming the norm, even in intraspecific studies, despite the difficulties associated with the sequencing of nuclear gene fragments in heterozygote individuals.

In addition, we argue further that it is important to explore intraspecific variation not only within a single species, but also, simultaneously, in the other most closely related species. Recent studies have highlighted several cases for which a fraction of the genetic variants found in one species are closer to those of - or even shared with-at least one other related species; this is due either to the persistence of ancestral polymorphism through time (i.e., incomplete lineage sorting between the two species; e.g., Nagl et al., 1998) or hybridization events, recent or past, that have occurred between species and lead to introgression phenomena (e.g., Gomez-Zurita and Vogler, 2003, 2006; Mardulyn et al., 2011; Nevado et al., 2011). In these cases, interpreting genetic variation in the context of a single species can be misleading, because its current pattern of variation stems from a broader evolutionary history involving other species. To illustrate this, we explore multilocus sequence variation within a complex of four sibling species, thereby combining the exploration of patterns of genetic

 ^{*} Corresponding author.
E-mail addresses: maudquinzin@gmail.com (M.C. Quinzin), pmarduly@ulb.ac.be
(P. Mardulyn).

relationships at both the inter- and intraspecies levels. By using both phylogenetic and phylogeographic methods, we attempt to maximize the use of the available historical signal and demonstrate the benefits of studying intraspecific genetic variation in a broader multi-species context. More specifically, we investigate genetic variation at five independent loci (1) to characterize the genetic boundaries among species, i.e. the degree of lineage sorting, (2) to infer their phylogenetic relationships and (3) to explore reproductive barriers among species whose geographic ranges partially overlap.

We focus on the cold-adapted leaf beetles of the subgenus Goniomena (genus Gonioctena) specialized on different host plants. The group appears ideal for our purpose, because its species display very little morphological and/or ecological differences among themselves, suggesting a recent history of speciation, and because their range partially overlap in Europe, creating the possibility of genetic exchange among species (Fig. 1). The range overlap is observed both at the scale of the species range and at the local scale, as the host plants of these species share similar habitats and are often found in close proximity. In addition, while the male genitalia seem to provide a valid diagnostic character to identify each species, the morphological distinction between G. guinguepunctata and G. intermedia is not straightforward in practice, as it is based on a small difference for which within-species variation occurs (Palmen, 1948; Franz and Palmen, 1950). For this reason, molecular markers appear ideal to delimit species objectively, and thus to clarify the geographic range of each species, and to investigate whether pre- or postzygotic barriers to reproduction exist to prevent hybridization. Finally, a pattern of strong host plant specialization that differs widely among the four species (see below) raises questions about the role such specialization had on the speciation process: did a host plant shift trigger a speciation event, or did the host plant shift follow a pre-existing differentiation?

We gather an extensive dataset on genetic variation for the four species by sequencing four nuclear and one mitochondrial loci for individuals from several populations in each species, spanning their entire geographic distribution. For two species, *G. intermedia* and *G. quinquepunctata*, the most difficult to differentiate based on

morphology and host plant use, we analyze larger samples and observe several cases of unidirectional introgression with the transfer of genetic material from *G. quinquepunctata* to *G. intermedia* in identified regions of secondary contact. We discuss how hybridization and demographic factors could have impacted the observed genetic patterns, and argue that collecting genetic variation data simultaneously from the four species allowed us to avoid erroneous interpretations.

2. Materials and methods

2.1. Study species

The four species investigated here form the subgenus *Goniome-na* within the genus *Gonioctena* (Kippenberg, 1994). They are univoltine (a single generation per year) and have a similar life cycle with four larval stages (Axelsson et al., 1974; Takizawa, 1976). Overall, the subgenus displays a highly fragmented boreo-montane distribution in Europe. It is largely present in the north, while restricted to mountainous (or in altitudes > 400–600 m) habitats in central and southern Europe. Yet, the geographic distributions vary among species.

An initial phylogenetic hypothesis is already available for the genus Gonioctena based on mitochondrial DNA sequence and allozyme variation (Mardulyn et al., 1997). Although its history of host plant association shows dramatic shifts among distantly-related plant families, the subgenus Goniomena is interesting because it displays, for only four sibling species of insects, such a high diversity of host plants including members of the families Rosaceae (Sorbus and Prunus), Salicaceae (Salix), and Betulaceae (Alnus and Corylus). Yet, each leaf beetle species appears to be specialized on a limited subset of these plants: G. pallida on Corylus avellana (Common hazel) and some Salix sp. (willows), G. interposita on Alnus viridis (Green alder), and both G. quinquepunctata and G. intermedia on Sorbus aucuparia (European mountain-ash) and Prunus padus (Bird cherry). In addition, we found that G. pallida accepted to feed on leaves of *P. padus* in the laboratory, although we did not find any wild populations on this host plant.



Fig. 1. Expected range for the four studied species, displayed by a dotted colored line and inferred from information gathered in the literature. The expected ranges for *G. intermedia* and *G. quinquepunctata* are not differentiated because the available information available *a priori* was ambiguous for these two species. Populations for which sampled individuals were sequenced for this study are indicated. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We found contradictory information from different sources in the literature regarding host plant association and geographic range for each species, especially for *G. intermedia* and *G. quinquepunctata* (Palmen, 1948; Franz and Palmen, 1950; Kippenberg, 1994; Mardulyn et al., 1997). These ambiguities could be attributed to the fact that their identification was relying almost exclusively on genital characteristics, which can be misleading because intraspecific variation at this level (e.g., Zaragoza-Caballero, 2007) sometimes induces subjective interpretations, even from specialists (Mutanen, 2005). Our preliminary species identification of collected individuals was based on the morphology of male genitalia, as described in the detailed works of Palmen (1948) and Franz and Palmen (1950), and on the host plant on which they were found. These species identification were then verified with the data on DNA sequence variation (see below).

Also, larvae of *G. pallida, G. intermedia* and *G. quinquepunctata* exhibit a cryptic green color, matching closely that of the leaves they feed on. This feature makes the larvae less conspicuous, and is probably an adaptation to avoid predators and/or parasites, and is not shared by other species in the genus.

2.2. Population sampling

Hereafter, we will refer to a population as a group of beetles associated to a patch of host trees (in the range of 100–200 m²). Populations were pooled into regions, defined by a mountain range or an area of relatively high elevation, isolated from other regions by large areas where the species is absent. The entire Scandinavian area (northern Europe) will thus be treated as a single region. This reflects the boreo-montane distribution. Information gathered from the literature – i.e., prior to analysis-(Palmen, 1948; Franz and Palmen, 1950; Kippenberg, 1994), over the distribution of each species in Europe, was used to direct our sampling efforts (Fig. 1). Samples from natural populations were collected between 2008 and 2010, in localities throughout the entire geographic range of

Table 1

Numbers of sampled localities and of collected individuals per region for each species.

each species (Table 1, Supporting information Table 1), including some localities in which two or more species live in sympatry (Fig. 1). We collected a maximum of one individual per tree to maximize genetic variability sampled at each locality. For G. pall*ida*, samples and DNA sequences were already available for many localities from a previous study (Mardulyn et al., 2009), but new sequences were obtained for some of the nuclear markers used here and/or from new localities. For G. intermedia and G. quinquepunctata, we sequenced a minimum of 10 individuals per region, distributed in separate areas in order to homogenize the sampling process as much as possible. Finally, individuals of G. interposita were collected from the only two regions where this species is known to occur: the Alps and the Carpathians. Host plants were recorded for all sampled populations. Because larvae have a limited ability to move and because they are required to feed almost at all time, their presence on a plant leaf was taken as reliable evidence that their respective species feed on that particular plant. Therefore, all host plant identifications for the four insect species were corroborated by multiple observations.

2.3. DNA extraction and sequencing

Genomic DNA was extracted using the Dneasy Tissue Kit from Qiagen (Hilden, Germany). Specimens preserved in alcohol were each dried, grinded in Qiagen ATL buffer and incubated overnight with proteinase K at 55 °C. The remaining DNA-extraction steps were conducted as described in the manufacturer's protocol. Using the polymerase chain reaction (PCR), we amplified one mitochondrial (cytochrome c oxydase subunit I, COI) and four nuclear (elongation factor 1 alpha, Elfac; ribosomal protein P0, RpP0; Actin protein, Actin; locus Wingless, Wg) gene fragments. Primer pairs were designed to specifically amplify targeted loci in all four species (Supporting information Table 2).

For the nuclear fragments, we used a combination of two approaches to reconstruct the allele sequence for multiple-site

Regions	n Localities	n Individuals for COI	n Individuals for Actin	n Individuals for Elfac	n Individuals for RpP0	n Individuals for Wg
G. intermedia						
Alps	10	46	30	42	38	46
Ardennes	3	9	10	14	11	15
Carpathians	4	25	25	27	21	28
Scandinavia (Finland)	6	15	7	14	10	15
Ural (Russia)	1	2	7	7	7	9
Total	24	97	79	104	87	113
G. auinauepunctata						
Alps	10	33	32	35	33	31
Carpathians	1	3	3	1	1	3
Massif Central	2	13	14	11	11	11
Pyrenees	2	24	18	20	20	25
Vosges	3	14	14	14	10	13
Total	18	87	81	81	75	83
G. interposita						
Alps	2	1	1	2	2	2
Carpathians	2	2	1	1	1	1
Total	4	3	2	3	3	3
G. pallida						
Alps	9	16	6	2	1	2
Carpathians	6	4	8	23	1	0
Germany	5	10	3	11	0	0
Massif central	4	9	6	7	0	1
Ore Monts	2	4	6	10	0	1
Scandinavia	4	7	9	9	1	1
Ural	2	5	0	1	1	0
Vosges	4	9	0	0	0	0
Altaï	2	3	7	7	1	1
Total	38	67	45	70	5	6

heterozygotes (i.e. with more than one polymorphic position). We first attempted to infer a maximum of allele sequences with the software PHASE v2.1.1 (Stephens et al., 2001; Stephens and Donnelly, 2003). Known allele sequences were used with PHASE to resolve the phase of heterozygous sequences, analyzing the (input) data set prepared with the web tool SeqPHASE (Flot, 2010; http://www.mnhn.fr/jfflot/seqphase). We used the default MR model, that allows explicitly for recombination, because the manual suggests it as more accurate. Preliminary analyses have shown that default settings of 100 burn-in iterations preceding 1000 main steps with a thinning interval of 1 were sufficient to observe consistency across 3 replicates. The seed number was modified for each of the three replicated runs, and convergence was verified manually. As suggested by Garrick and collaborators (2010), the default 90% threshold of phase certainty was chosen as the cut-off value below which haplotype reconstruction was considered unreliable. Moreover, these authors have shown that the common practice of omitting unresolved genotypes may induce systematic bias in the subsequent estimation of population genetics parameters because of possible exclusion of rare alleles and heterozygote genotypes. To avoid this risk, we cloned the heterozygous PCR products for which the phase could not be determined reliably. These amplified products were purified and ligated into a pGEM-3Z vector (Promega), then transferred to E. coli JM109 competent cells. At least five clones per PCR product were sequenced, and these sequences were compared to the one initially obtained by direct sequencing. When the sequences of all five clones were identical or when evidence of PCR recombination occurred, additional clones were sequenced to ensure an accurate allele sequence reconstruction.

2.4. Gene trees, species tree, and divergence times

Sequences were easily aligned manually in the sequence editor Se-Al (Rambaut, 2002). For two aligned data sets with indels, corresponding to loci Elfac and RpPO, gaps were considered as missing data and their presence/absence recoded as separate characters. These characters were added at the end of all sequences and used in NETWORK analyses, but not in maximum likelihood and Bayesian phylogenetic analyses (see below). Multi-nucleotide gaps were treated as follow: if of different lengths and starting at the same site (in either direction), they were treated as one multistate character, with each size class considered as a different character state. When gaps of equal or different sizes overlapped but started at a different site (in both directions), they were treated as separate characters. Sequences that were identified as imported from other species through hybridization (evidence for introgression from conflicting historical signal among loci, see below in Section 3.3) were excluded from the data set prior to all phylogenetic analyses, including for the purpose of estimating divergence time.

As the complexity of a phylogenetic estimation increases dramatically with the number of sequences to analyze (e.g., Felsenstein, 2004), we did not include all of our *G. intermedia* and *G. quinquepunctata* sequences in the data set for all phylogenetic analyses. Maximum Likelihood and Bayesian analyses were conducted on each locus separately with a reduced data set. That is, only a subset of different allelic variants from the entire available set were used per species for *G. intermedia* and *G. quinquepunctata*. Indeed, using all sequences available for these two species, many of which are multiple copies of the same alleles, would have included redundant information not useful for the purpose of inferring phylogenetic relationships among the four studied species. Instead, we used variants from different regions of the geographic distribution, while at the same time maximizing genetic variability within each species, using the allele networks (see results) as a guide. The same strategy was used to choose several sequences of *G. pallida* for the COI, Elfac and Actin loci from the data set already published by Mardulyn et al. (2009). The sequences of outgroup species belonging to four other subgenera (*Gonioctena Gonioctena, Gonioctena Spartophila, Gonioctena Spartomena* and *Gonioctena Spartoxena*; Supporting information Table 3) were also included in these analyses.

Maximum likelihood analyses were performed using the software PhyML 3.0 (Guindon et al., 2010) with a substitution model that varies from locus to locus (COI, GTR + G + [gamma]; Elfac, HKY + G; Wg, GTR + G; Actin, GTR + G + [gamma]; RpP0, HKY + G) and chosen using the Akaike Information criterion (AIC) with the program [ModelTest 0.1 (Posada, 2008), optimized equilibrium nucleotide frequencies, an estimated proportion of invariable sites and four categories of substitution rates distributed following a gamma distribution whose shape parameter was also estimated from the data. Ten separate analyses were run, each from a different random starting tree, except for bootstrap analyses (500 iterations each) in which each replicate was analyzed with a single run, started from a neighbor-joining tree (option BION]). Bayesian inferences were performed with MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). For each locus, two independent runs were launched, each with four MCMC chains (1 million generations per chain) among which three were heated. The cold chain was sampled every hundred steps and we verified that enough swapping occurred among chains, by checking that the mean standard deviation of frequencies was less than 0.01, as suggested in the program manual. The same nucleotide substitution models as in the maximum likelihood analyses were used. All other analysis parameters were set to their default values.

Instead of concatenating sequences from different loci into a single data set for a "total evidence" analysis, we used an algorithm that infer the species tree from multiple loci by explicitly modeling intra-species polymorphism and the possibility of incomplete lineage sorting among species, *BEAST (Heled and Drummond, 2010). Because this approach relies on a coalescence model to describe within-species sequence variation, which assumes that sequences were randomly sampled from each species, and because we used it not only to estimate phylogenetic relationships among species but also species divergence times, this analysis was performed on a much larger data set that included all available sequences for all loci, with BEAST version 1.6.2. We ran a MCMC chain of 1.5×10^9 steps. The nucleotide substitution models, clock models and tree models were all considered unlinked among loci. A HKY + G nucleotide substitution model and a strict molecular clock were assumed for each locus. Although no mutation rate specifically estimated for Gonioctena is available, the mutation rate of the mitochondrial fragment COI has been estimated for various insect taxa. We assumed that the COI mutation rate in Gonioctena falls somewhere within the range of the different insect mutation rates that have been estimated for this particular gene in previous studies, defining a prior range of 1.15% (estimated for a Lepidoptera species by Nazari and Sperling (2007)) to 4% (estimated for Orthoptera by Shapiro et al. (2006)) substitutions per nucleotide and per million years. A uniform prior ranging from 0.0115 to 0.0400 substitution per nucleotide and per million years was thus defined for the clock rate of the COI locus. Clock rates were estimated for all loci. A Yule process was chosen as species tree prior. The ploidy level of the nuclear loci was set to "autosomal nuclear". while that of the COI locus to "mitochondrial". For other parameters, including the priors for all model parameters, we chose the default settings. Sampling of the MCMC chain was set to generate a result file of 10,000 trees. Those trees were analyzed with TreeAnnotator version 1.6.2, previously discarding the first 1000 trees as burn-in, and the log file with Tracer version 1.5, both included in the BEAST package. The final estimated species tree was visualized with FigTree version 1.3.1 (http://tree.bio.ed.ac.uk/). For each run, we ensured that the effective sample size of the most important estimated parameters (posterior probability and likelihood of the species tree, as well as the clock rates for each locus), were above 100.

2.5. Genetic variation and allele networks

In the case of *G. intermedia* and *G. quinquepunctata*, for which we collected and sequenced a larger number of individuals across their entire geographic distribution, evolutionary relationships among haplotypes and geographic characterization of genetic variability was explored with the construction of a median-Joining allele network (Bandelt et al., 1999) inferred for each gene fragment separately with NETWORK v4.610 (http://www.fluxus-engineer-ing.com/sharenet.htm), using default settings. Sequence alignments that include recoded gap characters for Elfac (36 additional characters) and RpP0 (4 additional characters) were used.

In addition, to statistically test whether a significant genetic structure characterizes the distribution of diversity among regions within both species, we also conducted AMOVA analyses with ARLEQUIN v3.5.1. (Excoffier and Lischer, 2010) using both allele frequency and genetic distance information (Excoffier et al., 1992). The significance test was based on 10,000 permutations.

3. Results

3.1. Sequence variation, species identification and species range

The five matrices of aligned sequences, after the pruning of both 5'- and 3'-ends to avoid trailing gaps, display a total fragment length (excluding gap coding) of 987, 1061, 405, 479 and 619 nucleotides for COI, Elfac, Wg, RpP0, and Actin, respectively. These DNA sequence matrices are each characterized by 173, 170, 50, 41 and 37 polymorphic sites (again, excluding gap coding). The frequency and geographic distribution of each allele are given in Supporting information Table 1 for all five markers.

Species identification based on the morphology of the male genitalia matches the well-differentiated groups delimited by DNA sequence variation at multiple loci (see below), thus confirming the reliability of the male genitalia as a diagnostic character. The identification of the collected individuals was then used to refine our knowledge of the geographic ranges of *G. intermedia* and *G. quinquepunctata*, for which contradictory informations were found in the literature (as summarized in Fig. 1). Fig. 2 presents the identification of *G. intermedia* and *G. quinquepunctata* populations sampled in this study, based on the molecular markers. From there, it becomes obvious that the distributions of both species only partially overlap. Focusing on the central and southern part of their range, *G. quinquepunctata* is essentially present in the west (Alps, Vosges, Massif Central and Pyrenees), even though it is also



Fig. 2. Distribution map of all sampled populations of *G. intermedia* (white circles) and *G. quinquepunctata* (black circles) in Europe, along with regions defined for the analyses. Species identification is based on sequences from all loci. Populations of *G. intermedia* in which past and/or current introgression has been detected are highlighted (gray circles). All studied populations included individuals from a single species.

observed in the Carpathians (in a single locality, Fig. 2), while *G. intermedia* is found only in the east (central and eastern part of the Alps, Carpathians, Ardennes and in Asia, Ural mountains). In central Europe, their distributions overlap mainly in the center of the Alps, and somewhat in the Carpathians. In northern Europe (Scandinavia), although we only collected populations of *G. intermedia*, *G. quinquepunctata* is likely present as well, because it has been observed in the field by others (H. Silfverberg, *pers. com.*), and because we found evidences of recent introgression between both species in this region (see below). It is worth noting that although *G. intermedia* and *G. quinquepunctata* feed on the

same host plants (see below), they were never found simultaneously in the same locality. In comparison, the distribution of *G. pallida*, as described in Mardulyn et al. (2009), is more widespread (Fig. 1); the species is found in all regions of Europe where *G. intermedia* and *G. quinquepunctata* were collected, with the exception of the Pyrenees, and extends its range much more to the east, as far as the Altaï mountains. On the contrary, *G. interposita* displays a more restricted geographic distribution, in accordance with what has been described in the literature (Franz and Palmen, 1950), as it is only found in the Carpathians and the Eastern part of the Alps.



Fig. 3. Maximum likelihood gene trees inferred with phyML for each of the four nuclear (a) and mitochondrial (b) loci for the group Goniomena including Gonioctena intermedia (IM), G. quinquepunctata (Q), G. pallida (Pal) and G. interposita (IP).

3.2. Phylogenetic trees and divergence time estimates

Gene trees estimated from the ML analyses for each separate locus (Fig. 3) are very similar to estimates from the Bayesian analyses (data not shown). While the monophyly of each of the four species is confirmed in only two trees (Elfac and COI), the non-monophyly of some of the species observed in the other trees are never supported by high bootstrap values. The lack of support for species monophyly with some loci could be a consequence of incomplete lineage sorting or of insufficient phylogenetic signal. The same unrooted four-species tree ((*G. quinquepunctata, G. intermedia*), (*G. pallida, G. interposita*)) was observed with each locus, but the position of the root, as determined by outgroups, varies among trees.

The ^{*}BEAST analysis, seeking to infer a species tree by analyzing multiple independent loci, and which explicitly models the possibility of incomplete lineage sorting, resulted in the same species tree topology (Fig. 4). Relative divergence times estimated among species are also presented and we cannot reject the hypothesis that both splits occurred within the same time frame. In addition, the phylogenetic relationships displayed on Fig. 4 show that larval morphology within the subgenus experienced convergent evolution: green larvae are shared by *G. intermedia, G. quinquepunctata,* and *G. pallida,* but is not found elsewhere in genus *Gonioctena.* Therefore, either this feature has been acquired twice independently in the subgenus, or it has been acquired once in the ancestor of the four species, and was then lost in *G. interposita.*

The genetic identification of sampled individuals, together with our records of the host plant on which they were found, confirm that *G. pallida* and *G. interposita* feed on different host plants, but that *G. intermedia* and *G. quinquepunctata* feed on the same host trees, *Sorbus aucuparia* and *Prunus padus*. Because these last two leaf beetles are sister species (Fig. 4), it is likely that they share their diet with their common ancestor.

3.3. Allele networks and hybridization

Median-joining (MJ) networks summarizing genetic variation for *G. intermedia* and *G. quinquepunctata* are displayed for each locus in Fig. 5 (see Supporting Information Figs. 1–5 for networks with allele identification). For all loci, there is an unambiguous separation between sequences from the two species: each network is compatible with a hypothesis of monophyly of each species, even though a single mutation distinguishes them on the Actin network. Within species, all networks revealed a strong association between genetic variation and geographic distribution: colors identifying regions are not randomly distributed among alleles. This was confirmed by AMOVA analyses, that resulted in large and highly significant Φ_{st} values (evaluating genetic partitioning among regions; Table 2). Although complete lineage sorting among regions is not achieved yet, even for the mitochondrial locus, the strong phylogeographic structure observed suggests a long history of isolation in both species, at least between populations of the Alps and the rest of the distribution.

While clusters identified using the morphology of male genitalia correspond well to those inferred with the five molecular markers, allowing to identify each individual either as *G. intermedia* or *G. quinquepunctata*, we detected a few conflicts in the clusters



Fig. 4. Species tree constructed with ^{*}BEAST (150 million steps); mean estimated divergence times in millions of years (above the branches) and their 95% HPD interval (values under the branches and represented by purple lines on the branch) are given. Associated host plant trees and larvae morphology are displayed for all four species. Image of male aedeage morphology are also shown for *G. intermedia* and *G. quinquepunctata*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



Fig. 5. Allele networks displayed for each locus; each pie chart represents an observed allele whose frequency is proportional to the circle surface. Colors indicate regions where gene copies were sampled. When more than one mutation step separate two alleles on the network, mutations are depicted by dashes crossing branches, or, when more convenient, the number of mutations is displayed in a square box. Small plain black circles characterize unsampled alleles and red boxes identify alleles that have presumably been introgressed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2 AMOVA Φ_{st} for *G. intermedia* and *G. quinquepunctata.*

Locus	Fixation index (Φ_{st})	p-Value of significance test*			
G. intermedia					
COI	0.641	<0.0000			
Actin	0.237	<0.0000			
Elfac	0.080	0.002 ± 0.00014			
RpP0	0.105	<0.0000			
Wg	0.340	<0.0000			
G. quinquepunctata					
COI	0.314	<0.0000			
Actin	0.227	<0.0000			
Elfac	0.381	<0.0000			
RpP0	0.317	<0.0000			
Wg	0.295	<0.0000			

* With 10,000 permutations.

obtained with the mitochondrial locus and the four nuclear markers. Indeed, twenty-seven individuals unambiguously identified as *G. intermedia* by the nuclear markers harbor a typical *G. quinquepunctata* mitochondrial allele. Mitochondrial alleles corresponding to these individuals are highlighted and identified in Fig. 5. We interpret this pattern as the result of unidirectional introgression of the *G. quinquepunctata* mitochondrial genome into individuals of *G. intermedia*, probably mediated by hybridization.

An alternative explanation of incomplete lineage sorting of the mitochondrial genome between both species is highly unlikely given the large number (64) of mutations separating the two species, compared to the low number (0–4) of mutations separating the putatively introgressed alleles in *G. intermedia* individuals from alleles found in *G. quinquepunctata*.

Interestingly, these G. intermedia individuals harboring a typical G. quinquepunctata mitochondrial allele are quite geographically dispersed; they are found in two populations from Scandinavia (Finland), one from the Carpathians (Romania) and four from the Alps (Switzerland) (Figs. 2 and 5). They thus probably indicate multiple independent hybridization events. This is further supported by the fact that many genetically distantly-related alleles of G. quinquepunctata origin were found in G. intermedia individuals. If a single introgression event had occurred, we should have expected to find only one allele (or at least several closely-related alleles) of foreign origin in G. intermedia. Furthermore, with the exception of two alpine populations, all individuals sampled in these localities were characterized by an introgressed mitochondrial genome (i.e., no G. intermedia individual there harbors a mitochondrial genome typical of that species). The overall observation that introgressed specimens were found only in regions were the two species co-occur offers additional evidence against incomplete lineage sorting, as such a geographic structuring would not be expected under this hypothesis.

In a couple cases, a similar conflict in species affiliation was observed, but this time among nuclear loci. One individual from an alpine G. intermedia population displaying an introgressed mitochondrial genome also harbors one of its two wingless alleles (h17) that is one mutation apart from the most frequent G. quinquepunctata allele, while seven mutations away from the closest G. intermedia allele. The exact same pattern, involving the same introgressed allele, is observed for an individual collected this time in Finland (Scandinavia). Again, because these two G. intermedia individuals with an introgressed nuclear variant were sampled more than 2000 km apart, they are likely the result of two independent hybridization events. In the second case, a G. intermedia individual collected in the Alps and already harboring a mitochondrial haplotype characteristic of *G. quinquepunctata*, also carries the most frequent G. quinquepunctata allele, this time for the Actin locus (allele h14). These nuclear alleles found in G. intermedia individuals are most likely of *G. auinauepunctata* origin, and offer thus additional evidence for past hybridization events between the two species. Interestingly, no G. quinquepunctata individual displays typical G. intermedia alleles, such that introgression appears unidirectional.

4. Discussion

4.1. Species characterization

The DNA sequence data showed first that the complex was unambiguously divided into four genetically well-differentiated groups that could each be associated with a different previously defined species based on male genitalia morphology. Using DNA sequence variation at multiple loci to identify the species to which all sampled individuals belong, along with host plant records from collecting trips, we were able to better characterize the current geographic distribution and feeding behavior of each species. In the wild, G. pallida feeds on several Salix species and on Corylus avellana with the exception of one population from Scandinavia that feeds on Prunus padus. G. pallida displays the largest geographic distribution, spanning most of Europe and also extending into Northwest Asia. Although G. quinquepunctata and G. intermedia both feed on Prunus padus and Sorbus aucuparia, they are characterized by mostly distinct geographic distributions in Europe. G. quin*quepunctata* is mainly present in the west of central and southern Europe, *G. intermedia* is mainly found in the north and the east; however, the two distributions overlap in a large part of the Alps, to a lesser extent in the Carpathians, and probably in parts of Scandinavia. Finally, G. interposita has its distribution restricted to two mountain ranges: the eastern Alps and the Carpathians and feeds on a single species in the wild, Alnus viridis.

4.2. Phylogeny and speciation

The phylogenetic hypothesis obtained previously for the genus *Gonioctena* and based on mitochondrial DNA sequences of only a few individuals (Mardulyn et al., 1997), was here confirmed using five independent loci and a much larger set of sequences per species. In the light of this confirmed phylogenetic hypothesis, we can discuss the evolution of two specific traits within the group, larval morphology and host plant specialization.

The species tree of Fig. 4 suggests parallel evolution of larval morphology, as *G. intermedia*, *G. quinquepunctata* and *G. pallida* are the only species in the genus *Gonioctena* for which the larvae are not characterized by an external black pigmentation. Instead, the three species larval stages exhibit a cryptic green color, matching closely that of the leaves they feed on, yet they are not united in a clade. An alternative evolutionary explanation, the loss of the

cryptic green color (evolutionary reversal) in *G. interposita*, sister species of *G. pallida*, is equally parsimonious.

Focusing on host plant specialization, the same group of three species is characterized by its ability to feed on *Prunus padus*, again, a feature not shared by *G. interposita*. If we consider this as evidence that feeding on *P. padus* is an ancestral behavior, then it was lost in *G. interposita*. Because other examples of lineage differentiation in specialist phytophagous insects have been shown to be induced by a shift towards a new host plant (e.g., Stireman et al., 2005; Borer et al., 2011; Matsubayashi et al., 2011), a similar mechanism could be at the origin of the split between *G. interposita* and *G. pallida*. Indeed, the two species feed on very different plants, while sharing many life cycle and ecological traits. Because their current geographic distributions completely overlap, a hypothesis of sympatric speciation involving specialization on different plants seems likely.

A similar hypothesis was also initially considered for G. interme*dia* and *G. auinguepunctata* as it was first thought they specialized on two different host plants (Mardulyn et al., 1997). However, it is safely rejected because our host plant records and species identifications showed that both species feed on the same two plant species. Since, in addition, their current geographic distributions only partially overlap, a hypothesis of allopatric speciation is more plausible in this case. Because our data show that reproductive isolation between the two species is not reached yet, and because the genetic distance between the two species is large, it can only be concluded that geographic isolation has separated them during the largest part of their evolutionary history. Overall, our data show that the different ranges displayed by these two leaf beetles do not result from specialization on different host plants that are characterized by different geographic distributions, but other factors might be involved (e.g., differentiation of their climatic niche).

We also inferred divergence times among species, although the absolute values of these estimates should be considered cautiously, because based on a range of mutation rates of a single locus (COI) inferred for other insects (Nazari and Sperling, 2007; Shapiro et al., 2006). These estimates do, nonetheless, suggest that the divergences are more than several hundred thousands of years old, which places the speciation events long before the last glacial episode. Regarding the relative timing of the two last speciation events, one leading to the divergence between *G. pallida* and *G. interposita*, the other leading to *G. quinquepunctata* and *G. intermedia*, the two estimated time intervals overlap. These events may thus have occurred during a relatively similar time period.

4.3. Hybridization and introgression

Analysis of the large dataset obtained for *G. quinquepunctata* and *G. intermedia* showed that hybridization events occurred between the two sister species in the areas where their distributions overlap. We detected transferred genetic material from *G. quinquepunctata* to *G. intermedia*, both in the mitochondrial and nuclear genomes. In general, signs for the introgression of nuclear DNA from one species to another is expected to progressively fade out, and ultimately disappear, as recombination mingles the transferred nuclear fragment with the genetic material of the host species across generations. The fact that nuclear introgression is still detectable in our samples offers strong evidence for the recent nature of these hybridization events. This is corroborated further by the locations of the introgressed individuals, only found in those areas were the two species currently co-exist.

Given the strong level of genetic divergence observed between *G. intermedia* and *G. quinquepunctata*, for two loci in particular (ElFac and COI), it is clear that they have been diverging from each other for a long period of time, likely more than 400,000 years according to our estimates (Fig. 4). Adding the fact that the two species are characterized by different geographic distributions,

and yet have recently been able to hybridize in the two (or three) parts of their range in which they are found in sympatry, it suggests they have first differentiated in an allopatric geographic setting. The current overlapping parts of their range thus corresponds most likely to zones of secondary contact.

All cases of introgression detected in this study, both with nuclear and mitochondrial markers, involved the transfer of genetic material in the same direction, from the G. quinquepunctata genome into that of G. intermedia. Moreover, in four out of six populations of G. intermedia in which mitochondrial introgression is observed, all individuals harbor only G. quinquepunctata mitochondrial alleles. While natural selection is sometimes invoked as the best explanation for the complete replacement of the mitochondrial genome of one species by that of another, following introgression (e.g., Bachtrog et al., 2006; Gompert et al., 2008), it appears that a purely demographic explanation is also possible (e.g. Jordal et al., 2006; Nevado et al., 2009, 2011: Sequeira et al., 2011). Indeed, when a species colonizes a new area already occupied by a closely related species with which interbreeding is possible, massive introgression from the local to the invading species is expected, irrespective of the relative densities of the two species and without the need for selection to occur. This was demonstrated in a theoretical study based on computer simulations by Currat et al. (2008). The pattern of asymmetrical introgression uncovered by our analyses could therefore be a consequence of the colonization by G. intermedia of areas previously occupied by G. quinquepunctata alone; the range shift could possibly have followed the climate warming that occurred at the end of the last glaciation. This hypothesized colonization would have been favored by similarities in the life history traits of the two species, in particular in their feeding and mating behavior. Finally, we cannot reject the possibility that the observed asymmetrical gene introgression is simply a consequence of asymmetric barriers to reproduction (e.g., Scribner and Avise, 1994; Takami et al., 2007; Bolnick et al., 2008). Either pre- or post-zygotic barriers could have prevented hybridization between G. intermedia females and G. auinguepunctata males.

Furthermore, an interesting feature of the secondary contact zones highlighted here is the absence of a sampled population (defined here as a group of beetles associated with a patch of host plant) in which the two species coexist. In other words, even in those areas where the two species occur in sympatry, each population included exclusively individuals of G. quinquepunctata or G. intermedia (including the twelve populations sampled in the portion of the Alps in which the ranges of the two species overlap). Given that this pattern has most probably emerged independently in each identified secondary contact zone between the two species, it is unlikely to have appeared by chance. It could rather be interpreted as evidence for mutual exclusion (e.g., Ribeiro and Spielman, 1986; Kuno, 1992; Yoshimura and Clark, 1994). Because both species share the same food resources and habitats, it could be hypothesized that competition has favored one species over the other in the local populations were they meet. However, plant resources in sampled insect populations were always largely in excess and the possibility of it becoming a limiting factor to population expansion seems highly unlikely. While competition for food resource can probably be excluded, selection may still favor one species over the other, leading to large differences in population sizes between the two species in localities where they meet. If this occurs, individuals from the minority species could have no other opportunity but to mate with individuals of the other species. If hybrids are less viable, the entire process could lead to the disappearance of the minority species at the scale of the local population, while patterns of introgression remain as traces of hybridization. Note that unidirectional introgression could also have favored G. intermedia thanks to its ability to create new combinations of genetic material through the introduction of genes from the *G. quinquepunctata* mitochondrial genome. The important role of introgression in biological invasion has indeed already been stressed elsewhere (e.g. Hänfling et al., 2002; Rieseberg et al., 2007).

Detection of introgression is increasingly reported in the scientific literature thanks to a rising number of studies comparing mitochondrial and nuclear markers (e.g., Gomez-Zurita and Vogler, 2003, 2006; Leaché and Cole, 2007; Gompert et al., 2008; Wahlberg et al., 2009; Sequeira et al., 2011). Only a few studies so far have reported introgression of nuclear loci (e.g., Melo-Ferreira et al., 2009; Bossu and Near, 2009; Nevado et al., 2011) and in each case, like in our study, the associated signal for introgression was lower than that found for the mitochondrial genome. Here, the observation of a few nuclear introgressions was only possible thanks to an extended sampling effort and the sequencing of five independent DNA fragments. With fewer individuals, introgression at the nuclear level would not have been detected.

5. Conclusion

Overall, the sequence data analyzed here significantly increased our knowledge on the specific population history of the *Goniomena* species complex, a group of herbivorous insects adapted to cold climate. Combining analyses of sequence variation for multiple loci at both inter- and intra-species levels has provided information that could not have been retrieved in a single locus and/or a single species study. Most notably, mitochondrial and nuclear introgression would have gone unnoticed. We would not only have been unable to detect the occurrence of hybridization between two species, but in addition, many sampled individuals could have been assigned to the wrong species, thereby biasing several of our inferences, including the estimated species ranges or divergence times. This strengthens the importance to combine DNA sequence variation from multiple species and multiple loci in a single evolutionary study (e.g., Ballard and Whitlock, 2004; Dupuis et al., 2012).

Data archival

DNA sequences are available from Genbank under accession numbers: KJ785947 to KJ786215.

Acknowledgments

We are grateful to Y. Mikhailov, A. and H. Mardulyn for their help in collecting samples. We thank J. Cillis (IRSNB) for his help in electronic microscopy, H. Silfverberg for providing us with his personal information on the geographic distribution of Goniomena species in Scandinavia and A. Hougron for her contribution to the sequencing of the RpPO locus. M.Q. was supported by a FRIA scholarship from the Belgian Fund for Scientific Research (FRS-FNRS) and Van buuren Funds, and P.M. is Research Associate at the FRS-FNRS. This research project was funded by the FRS-FNRS (FRFC – convention n° 2.4554.09).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2014.05. 003.

References

Avise, J.C., 2000. Phylogeography: The History and Formation of Species. Harvard Univ. Press, Cambridge, MA.

Axelsson, B., Bossata, E., Lohm, U., Persson, T., Tenow, O., 1974. Energy flow through a larval population of *Phytodecta pallidus* L. (Col., Chrysomelidae) on *Corylus* avellana L. Zoon 2, 49–55.

- Bachtrog, D., Thornton, K., Clark, A., Andolfatto, P., 2006. Extensive introgression of mitochondrial DNA relative to nuclear genes in the *Drosophila yakuba* species group. Evolution 60, 292–302.
- Ballard, J., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13, 729–744.
- Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.
- Bolnick, D.I., Turelli, M., López-Fernández, H., Wainwright, P.C., Near, T.J., 2008. Accelerated mitochondrial evolution and "Darwin's corollary": asymmetric viability of reciprocal F1 hybrids in Centrarchid fishes. Genetics 178, 1037– 1048.
- Borer, M., van Nort, T., Arrigo, N., Buerki, S., Alvarez, N., 2011. Does a shift in host plants trigger speciation in the Alpine leaf beetle Oreina speciosissima (Coleoptera, Chrysomelidae)? BMC Evol. Biol. 11, 310.
- Bossu, C.M., Near, T.J., 2009. Gene trees reveal repeated instances of mitochondrial DNA introgression in orangethroat darters (percidae: etheostoma). Syst. Biol. 58, 114–129.
- Bossuyt, F., Milinkovitch, M.C., 2000. Convergent adaptive radiations in Madagascan and Asian ranid frogs reveal covariation between larval and adult traits. Proc. Natl. Acad. Sci. USA 97, 6585–6590.
- Currat, M., Ruedi, M., Petit, R.J., Excoffier, L., 2008. The hidden side of invasions: massive introgression by local genes. Evolution 62, 1908–1920.
- Dupuis, J.R., Roe, A.D., Sperling, F.A.H., 2012. Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. Mol. Ecol. 21, 4422–4436.
- Excoffier, L., Lischer, 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Res. 10, 564–567.
- Excoffier, L., Smouse, P., Quattro, J., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479–491.

Felsenstein, J., 2004. Inferring Phylogenies. Sinauer Associates, Sunderland, MA.

- Flot, J.-F., 2010. Seqphase: a web tool for interconverting phase input/output files and fasta sequence alignments. Mol. Ecol. Res. 10, 162–166.
- Franz, H., Palmen, E., 1950. Beitrag zur kenntnis der untergattung Goniomena Kirby der Gattung Phytodecta (Col., Chrysomelidae). Ann. Entomol. Fennici 16, 14–18.
- Garrick, R.C., Sunnucks, P., Dyer, R., 2010. Nuclear gene phylogeography using PHASE: dealing with unresolved genotypes, lost alleles, and systematic bias in parameter estimation. BMC Evol. Biol. 10, 118.
- Gómez-Zurita, J., Vogler, A.P., 2003. Incongruent nuclear and mitochondrial phylogeographic patterns in the *Timarcha goettingensis* species complex (Coleoptera, Chrysomelidae). J. Evol. Biol. 16, 833–843.
- Gómez-Zurita, J., Vogler, A.P., 2006. Testing introgressive hybridization hypotheses using statistical network analysis of nuclear and cytoplasmic haplotypes in the leaf beetle *Timarcha goettingensis* species complex. J. Mol. Evol. 62, 421–433.
 Gompert, Z., Forister, M.L., Fordyce, J.A., Nice, C.C., 2008. Widespread mito-nuclear
- Gompert, Z., Forister, M.L., Fordyce, J.A., Nice, C.C., 2008. Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in Lycaeides. Mol. Ecol. 17, 5231–5244.
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59, 307–321.
- Hänfling, B., Hellemans, B., Volckaert, F.A.M., Carvalho, G.R., 2002. Late glacial history of the cold-adapted freshwater fish *Cottus gobio* revealed by microsatellites. Mol. Ecol. 11, 1717–1729.
- Hare, M.P., 2001. Prospects for nuclear gene phylogeography. Trends Ecol. Evol. 16, 700–706.
- Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27, 570–580.
- Jordal, B.H., Emerson, B.C., Hewitt, G.H., 2006. Apparent "sympatric" speciation in ecologically similar herbivorous beetles facilitated by multiple colonizations of an island. Mol. Ecol. 15, 2935–2947.
- Kippenberg, H., 1994. Chrysomelinae. In: Lohse, G.A., Lucht, W.H. (Eds.). Die k\u00e4fer mitteleuropas, Band 14, 3.Supplementband. Goecke and Evers, Krefeld, Germany. pp. 9–92.
- Kuno, E., 1992. Competitive exclusion through reproductive interference. Res. Popul. Ecol. 34, 275–284.
- Leaché, A.D., Cole, C.J., 2007. Hybridization between multiple fence lizard lineages in an ecotone: locally discordant variation in mitochondrial DNA, chromosomes, and morphology. Mol. Ecol. 16, 1035–1054.
- Losos, J.B., 1998. Contingency and determinism in replicated adaptive radiations of island lizards. Science 279, 2115–2118.
- Mardulyn, P., Milinkovitch, M., Pasteels, J.M., 1997. Phylogenetic analyses of DNA and allozyme data suggest that *Gonioctena* leaf beetles (Coleoptera; Chrysomelidae) experienced convergent evolution in their history. Syst. Biol. 46, 722–747.
- Mardulyn, P., Mikhailov, Y.E., Pasteels, J.M., 2009. Testing phylogeographic hypotheses in a Euro-Siberian cold-adapted leaf beetle with coalescent simulations. Evolution 63, 2717–2729.
- Mardulyn, P., Othmezouri, N., Mikhailov, Y.E., Pasteels, J.M., 2011. Conflicting mitochondrial and nuclear phylogeographic signals and evolution of host-plant

shifts in the boreo-montane leaf beetle *Chrysomela lapponica*. Mol. Phylogenet. Evol. 61, 686–696.

- Matsubayashi, K.W., Kahono, S., Katakura, H., 2011. Divergent host plant specialization as the critical driving force in speciation between populations of a phytophagous ladybird beetle. J. Evol. Biol. 24, 1421–1432.
- Melo-Ferreira, J., Alves, P.C., Freitas, H., Ferrand, N., Boursot, P., 2009. The genomic legacy from the extinct *Lepus timidus* to the three hare species of Iberia: contrast between mtDNA, sex chromosomes and autosomes. Mol. Ecol. 18, 2643–2658.
- Mutanen, M., 2005. Delimitation difficulties in species splits: a morphometric case study on the *Euxoa tritici* complex (Lepidoptera, Noctuidae). Syst. Entomol. 30, 632–643.
- Nagl, S., Tichy, H., Mayer, W.E., Takahata, N., Klein, J., 1998. Persistence of neutral polymorphisms in Lake Victoria cichlid fish. Proc. Natl. Acad. Sci. USA 95, 14238–14243.
- Nazari, V., Sperling, F.A.H., 2007. Mitochondrial DNA divergence and phylogeography in western Palaearctic Parnassiinae (Lepidoptera: Papilionidae): how many species are there ? Insect Syst. Evol. 38, 121–138.
- Nevado, B., Koblmüller, S., Sturmbauer, C., Snoeks, J., Usano-Alemany, J., Verheyen, E., 2009. Complete mitochondrial DNA replacement in a Lake Tanganyika cichlid fish. Mol. Ecol. 18, 4240–4255.
- Nevado, B., Fazalova, V., Backeljau, T., Hanssens, M., Verheyen, E., 2011. Repeated unidirectional introgression of nuclear and mitochondrial DNA between four congeneric Tanganyikan cichlids. Mol. Biol. Evol. 28, 2253–2267.
- Nichols, R., 2001. Gene trees and species trees are not the same. Trends Ecol. Evol. 16, 358–364.
- Palmen, E., 1948. Zur Systematik finnischer Chrysomeliden. 4. Phytodecta (Goniomena) quinquepunctata F. Ph. intermedius Hellies. Ann. Entomol. Fennici 14, 110.
- Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. Mol. Biol. Evol. 5, 568–583.
- Posada, D., 2008. JModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25, 1253–1256.
- Rambaut, A., 2002. Se-Al: sequence alignment editor. https://evolve.zoo.ox.ac.uk/. Ribeiro, J.M.C., Spielman, A., 1986. The satyr effect: a model predicting parapatry and species extinction. Am. Nat. 128, 513–528.
- Rieseberg, L.H., Kim, S.-C., Randell, R.A., Whitney, K.D., Gross, B.L., Lexer, C., Clay, K., 2007. Hybridization and the colonization of novel habitats by annual sunflowers. Genetica 129, 149–165.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
- Rosenberg, N.A., Nordborg, M., 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Genetics 3, 380–390.
- Scribner, K.T., Avise, J.C., 1994. Population cage experiments with a vertebrate: the temporal demography and cytonuclear genetics of hybridization in *Gambusia* fishes. Evolution 48, 155–171.
- Sequeira, F., Sodré, D., Ferrand, N., Bernardi, J.A.R., Sampaio, I., Schneider, H., Vallinoto, M., 2011. Hybridization and massive mtDNA unidirectional introgression between the closely related Neotropical toads *Rhinella marina* and *R. schneideri* inferred from mtDNA and nuclear markers. BMC Evol. Biol. 11, 264.
- Shapiro, L.H., Strazanac, J.S., Roderick, G.K., 2006. Molecular phylogeny of Banza (Orthoptera: Tettigoniidae), the endemic katydids of the Hawaiian Archipelago. Mol. Phylogenet. Evol. 41, 53–63.
- Sota, T., Ishikawa, R., 2004. Phylogeny and life-history evolution in Carabus (subtribe Carabina: Coleoptera, Carabidae) based on sequences of two nuclear genes. Biol. J. Linn. Soc. 81, 135–149.
- Stephens, M., Donnelly, P., 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 73, 1162– 1169.
- Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68, 978–989.
- Stireman, J.O., Nason, J.D., Heard, S.B., 2005. Host-associated genetic differentiation in phytophagous insects: general phenomenon or isolated exceptions? Evidence from a goldenrod-insect community. Evolution 59, 2573–2587.
- Takami, Y., Nagata, N., Sasabe, M., Sota, T., 2007. Asymmetry in reproductive isolation and its effect on directional mitochondrial introgression in the parapatric ground beetles *Carabus yamato* and *C. albrechti*. Popul. Ecol. 49, 337–346.
- Takizawa, H., 1976. Larvae of the genus *Gonioctena* Chevrolat (Coleoptera, Chrysomelidae): descriptions of japanese species and the implications of larval characters for the phylogeny. Kontyû 44, 444–468.
- Wahlberg, N., Weingartner, E., Warren, A.D., Nylin, S., 2009. Timing major conflict between mitochondrial and nuclear genes in species relationships of *Polygonia* butterflies (Nymphalidae: Nymphalini). BMC Evol. Biol. 9, 92.
- Yoshimura, J., Clark, C.W., 1994. Population dynamics of sexual and resource competition. Theor. Popul. Biol. 45, 121–131.
- Zaragoza-Caballero, S., 2007. A new species of Photinus (Coleoptera: Lampyridae: Photinini) from Jalisco Mexico with comments on intraspecific aedeagal variability and a key to the species of the subgenus *Paraphotinus*. Zootaxa 1437, 61.