Biological Journal of the Linnean Society, 2008, **94**, 105–114. With 3 figures

Morphological and mitochondrial DNA analyses indicate the presence of a hybrid zone between two species of leaf beetle (Coleoptera; Chrysomelidae) in Southern Spain

LAURENT GATTO[1], PATRICK MARDULYN[1]* and JACQUES M. PASTEELS[2]

[1]*Laboratory of Evolutionary Genetics, Université Libre de Bruxelles, 6041 Gosselies, Belgium*
[2]*Laboratoire de Biologie Animale et Cellulaire, Université Libre de Bruxelles, 1050 Brussels, Belgium*

Received 9 March 2007; accepted for publication 10 July 2007

The present study reports a case where the survey of morphological and mitochondrial DNA variation among populations of a species complex of leaf beetle, the *Gonioctena variabilis* complex, has lead to the identification of a hybrid zone between two species of the complex in Southern Spain. The complex is divided into four species distributed around the western Mediterranean region. The four species, *G. variabilis*, *Gonioctena aegrota*, *Gonioctena gobanzi*, and *Gonioctena pseudogobanzi*, are traditionally determined by differences in the morphology of the male genitalia (aedeagus). To gain insight into the history of the speciation process within this species complex, we sampled populations in Portugal, Spain, Southern France, and Northern Italy. We sequenced a portion of the mitochondrial control region of each individual collected. A haplotype network of these sequences was found to comprise four distinct groups of sequence types, separated by a relatively large number of mutations. Moreover, in most of the samples for which morphological and molecular variation is available, there is a one-to-one correspondence between haplotype group, defined by mitochondrial sequence variation, and morphological groups defined on the basis of the aedeagus, showing evidence of four historically independent evolutionary units. This supports the use of the aedeagus morphology as a taxonomically informative trait in this species complex and a recent taxonomic revision upgrading four formerly subspecies, corresponding to the evolutionary units identified in the present study, to species status. However, some of the individuals from our samples in Southern Spain, morphologically identified as *G. aegrota*, were found to possess mitochondrial sequences typical of *G. pseudogobanzi*. The opposite case was also found. This suggests the presence of a zone of contact and hybridization between *G. aegrota* and *G. pseudogobanzi*. The location of this hybrid zone appears to be unusual. We identify historical scenarios that may explain our observations. *Biological Journal of the Linnean Society*, 2008, **94**, 105–114.

ADDITIONAL KEYWORDS: AT-rich region – *Gonioctena variabilis* – morphology – phylogeography – secondary contact zone.

INTRODUCTION

For many animals and plants in Europe, the presence of hybrid zones (i.e. areas in which genetically distinct groups of organisms meet and reproduce) is thought to be a consequence of the climatic oscillations that occurred during the Quaternary period (Hewitt, 2003; Hofreiter *et al*., 2004). During the last glacial maximum, all populations of many temperate species were restricted to a few small isolated refuges (Taberlet *et al*., 1998). As a result, populations of the same species but in different refuges have accumulated a large number of differences, possibly enough to induce speciation, before they were brought again into contact. At the end of the last ice age, when ice sheets retreated to the north, many species have re-colonized Central and Northern Europe. Populations that had been separated for a long period of time

*Corresponding author. E-mail: pmarduly@ulb.ac.be

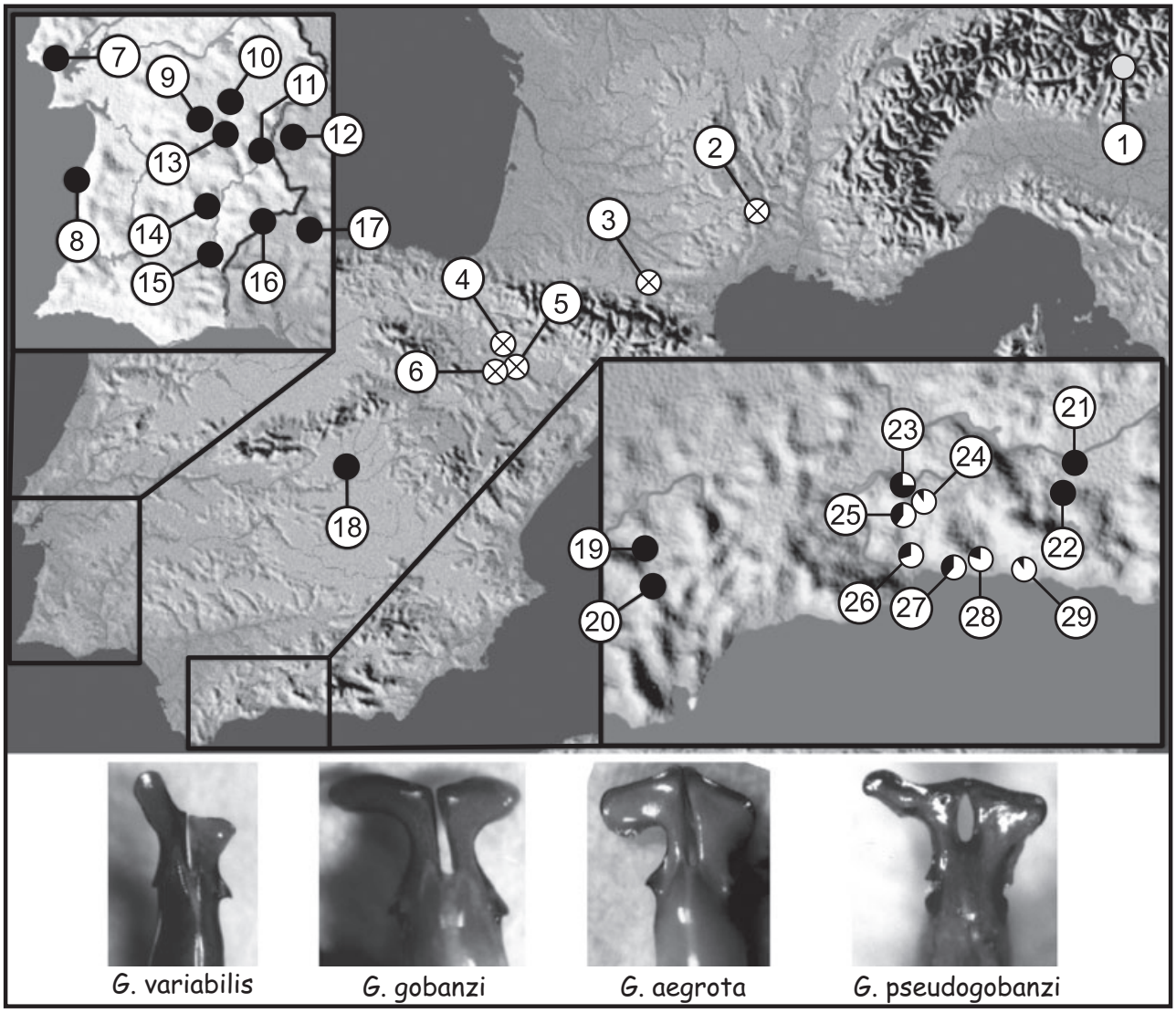

Figure 1. Sampling locations (for details, see Table 1) and haplogroup (hg) distribution: hgA (black), hgB (white), hgC (X), and hgD (grey). Photographs of the aedeagus for each of the four species, the main morphological character used for species identification, are displayed.

have expanded again and, in some cases, meet in so-called 'secondary contact zones', or 'hybrid zones' (if hybrid individuals are produced) (Taberlet *et al*., 1998; Hewitt, 2001).

Molecular markers have proven to be a powerful tool for identifying hybrid zones (Avise, 2004). In the present study, we report a case where the survey of morphological and mitochondrial DNA (mtDNA) variation among populations of a species complex of leaf beetle, the *Gonioctena variabilis* (Olivier) complex, has allowed the identification of a zone of contact and hybridization between two historically independent evolutionary units in Southern Spain. This complex is distributed around the Western Mediterranean region and, although previously considered as a single species divided into different subspecies, was recently divided into four separated species by Kippenberg (2001), each easily recognizable through the shape of the male genitalia (Fig. 1) and with specific host plant preferences.

To explore the history of the speciation process in this species complex, we assessed its mtDNA diversity across most of its geographical range and compared it to the geographical distribution of the observed morphological variation. More specifically: (1) we investigated the current geographical distribution of the morphological variation of the male genitalia (used by Kippenberg to define the four species of the complex);

Table 1. Sampling information: locations are grouped according to the haplogroup distributions

	Sampling location	Sample size	Geographical coordinates		Date
1	Pertica	3	45°53′00″N	11°03′00″E	August 1999
2	En Gauly	10	42°56′00″N	01°50′00″E	May 1999
3	Alès	7	44°08′00″N	04°05′00″W	June 1995
4	Monegrillo	2	41°38′00″N	00°38′00″W	March 2000
5	Osera de Ebro	3	41°35′00″N	00°35′00″W	March 2000
6	Pina de Ebro	5	41°34′00″N	00°36′00″W	March 2000
7	Lisboa	5	38°44′00″N	09°08′00″W	May 1995
8	Sines	9	37°58′00″N	08°52′00″W	May 1995
9	Evora	10	38°34′00″N	07°54′00″W	June 2000
10	Monte S Jage	8	38°41′00″N	07°44′00″W	June 1999
11	Mourao	10	38°22′00″N	07°20′00″W	June 1999
12	Alconchel	3	38°24′00″N	07°01′00″W	June 1999
13	N256	10	38°28′00″N	07°50′00″W	June 1999
14	Beja	10	38°01′00″N	07°52′00″W	June 1999
15	Mertola	10	37°38′00″N	07°40′00″W	June 1999
16	Rosa de la Frontera	10	37°57′50″N	07°12′50″W	June 2000
17	Aroche	10	37°57′00″N	06°55′00″W	June 1999
18	Aranjuez	2	40°02′00″N	03°37′00″W	May 2000
19	Ronda	9	36°45′00″N	05°10′00″W	May 2000
20	Col S Ronda	10	36°40′40″N	05°12′30″W	May 2000
21	Bracana	10	37°12′20″N	03°56′30″W	June 2000
22	Buenavista	10	37°05′40″N	03°59′00″W	June 2000
23	Antequera	8	37°00′00″N	04°31′20″W	May 2000
24	Torcal	10	36°58′00″N	04°31′50″W	June 2000
25	Villanueva	10	36°57′20″N	04°31′40″W	June 2000
26	Puerto de la Torre	10	36°44′20″N	04°28′30″W	May 2000
27	Chilces	13	36°44′00″N	04°14′00″W	May 1992
28	Velez-Malaga	10	36°47′00″N	04°06′00″W	June 2000
29	Torrox	10	36°47′00″N	03°57′40″W	May 2000

(2) we determined whether each species of the complex, as defined by Kippenberg, forms a monophyletic group based on mtDNA diversity; and (3) we tested whether the geographical and genetic isolation among all four defined species is complete, or whether hybridization is still occurring between individuals from different species.

MATERIAL AND METHODS

GEOGRAPHIC DISTRIBUTION, HOST PLANTS, AND SAMPLING OF POPULATIONS

Like most chrysomelid beetles, the studied *Gonioctena* are specialist feeders, whose diet is limited to a few plant species belonging to the *Fabaceae* family. The known geographical distribution and diet preferences of the four species defined by Kippenberg (2001) comprise: (1) *G. variabilis* (Oliver) in the French Mediteranean region and the North-East of Spain on *Genista scorpius*; (2) *Gonioctena gobanzi* (Reitter) in North Italy and Slovenia (Southern Alps) on *Genista radiata*; (3) *Gonioctena aegrota* (Fabricius) in Central and Southern Spain, Portugal and in North Africa (Bourdonné & Doguet, 1979; Kippenberg, 2001) on *Cytisus grandiflorus*, *Retama sphaerocarpa*, *Retama monosperma*, and *Genista umbellata*; and (4) *Gonioctena pseudogobanzi* Kippenberg whose geographical range appears to be restricted to Southern Spain (Kippenberg, 2001; J. M. Pasteels, pers. observ.), on *R. sphaerocarpa* and *G. umbellata*.

A total of 237 individuals belonging to the four species were collected across most of its European distribution. Table 1 lists the detailed collecting locations, collecting dates, and sample sizes. A map of the collecting sites is provided in Figure 1. The aedeagus of all collected male individuals was dissected for species identification.

DNA SEQUENCING

For each collected individual, we sequenced a portion of 356–376 bp of a nonrepetitive section (Mardulyn, Termonia & Milinkovitch, 2003) of the AT-rich region. Individual beetles were ground in homogenization

buffer and incubated in proteinase K (2 mg mL^{-1}) overnight at 40 °C. Successive phenol/chloroform extractions were performed. DNA was precipitated in ethanol, resuspended in TE buffer (10 mM Tris-HCl, 1 mM ethylenediaminetetraacetic acid, pH 8.0) and stored at –20 °C. The following primers were used for polymerase chain reaction amplification (35 cycles of 10 s at 95 °C, 60 s at 61 °C, and 60 s at 72 °C) of the considered portion of the AT-rich region: SR-J-14766 located in the ribosomal 12S gene (5′-TATAACC GCAACTGCTGGCAC-3′; Mardulyn *et al*., 2003) and ATR2 (5′-AATTAATTTTGGAACATCAATATTTGG-3′), located inside the control region. Sequencing reactions were performed with a Perkin Elmer Big Dye Terminator Kit, using the primers SR-J-14766 and ATR2b (5′-TTTTGGAACATCAATATTTGG-3′). The manufacturer's protocol was followed, except that an annealing temperature of 55 °C was used for the SR-J-14766 primer. Sequencing products were separated by vertical electrophoresis on a 4.75% acrylamide-urea gel on an ABI 377 Sequencer.

Molecular data analysis

Alignment of the nucleotide sequences was performed with the program PROALIGN (Loytynoja & Milinkovitch, 2003). Sites with a minimum posterior probability below 90% were removed. PAUP (Swofford, 2003) was used to characterize the aligned DNA sequences. A median-joining network was inferred from all AT-rich sequences using the computer program NETWORK (Bandelt, Forster & Röhl, 1999; downloaded from http://www.fluxus-engineering.com). Mean Kimura two-parameter distances within groups of haplotypes (see below) were computed using MEGA 3.1 (Kumar, Tamura & Nei, 2004).

Morphometric analysis

The genitalia of 45 males belonging to the four species were removed before DNA extraction for a morphometric analysis. Thirteen measurements, eight distances with a precision of 5 μm (*a*, *b*, *xy*, *x′y′*, *xz*, *x′z′*, *yz* and *y′z′*), and five angles with a precision of 0.5° (*xyz*, *x′y′z*, *yxz*, *y′x′z* and *xzx′*) were recorded per sample (Fig. 2).

The resulting dataset was subjected to a principal component analysis (PCA) to investigate whether the traits enable to statistically differentiate distinct morphological groups (corresponding here to the systematic units).

Time of divergence

Although there is no molecular clock calibration available for the control region in these beetles, we aimed to obtain an approximate estimate of the lower bound to the time of divergence (or time of splitting) among the four haplogroups. Molecular clock calibrations are available for some mitochondrial genes in a variety of insects (Zakharov, Caterino & Sperling, 2004), including some beetles, and range from 2.8×10^{-9} (carabid beetles: Su *et al*., 1998; Zakharov *et al*., 2004) to 49×10^{-9} (honey bees: Crozier, Crozier & Mackinlay, 1989) mutations per site and per generation. In addition, the AT-rich region is believed to be the most variable region of the mitochondrial genome, with the repetitive section of this region found to be 12 times more variable than cytochrome oxidase (CO) subunit *I* (CO *I*) in a related species of leaf beetle, *Gonioctena olivacea* (Mardulyn & Milinkovitch, 2005). At the same time, this repetitive section of the control region in *G. olivacea* is much more variable than any other part of the control region found in *G. olivacea* or in other *Gonioctena* or *Chrysomela* leaf beetles (Mardulyn *et al*., 2003), and we can reasonably assume that the DNA fragment sequenced here is not more variable than this repetitive section. If we use the highest estimated evolutionary rate for mtDNA in insects, 49×10^{-9} mutations per site and per generations, found in the CO *II* (versus 43×10^{-9} mutations per site and per generations for CO *I*) of *Apis mellifera* (Crozier *et al*., 1989), and multiply it by 12, to estimate the time of divergence among the different haplogroups identified, we probably have a very conservative estimate of the lower bound of these divergence times (i.e. the actual divergence times are likely to be higher than these values). Using this value, and assuming one generation per year, we have estimated the lowest possible divergence time among all four species that is compatible with the molecular data.

To estimate this lower bound, we used population modelling. Tavaré *et al*. (1997) suggested the use of an approximate method based on a summary statistic, the number of segregating sites in the sample (S_n). We used a similar approach, but with two summary statistics: the number of mutations on the genealogy and the number of allele types in the sample (K_n). Twenty-three thousand genealogies were simulated following a specified evolutionary model, each one being then accepted or rejected depending on the size of the differences between the simulated and real values (i.e. the values calculated from the DNA sequence data) of the two summary statistics considered (Pritchard *et al*., 1999). The resulting accepted genealogies are used to estimate the posterior probability density marginals for several variables (e.g. the time to the most recent common ancestor). As the time of splitting (the parameter we seek to estimate) varies between simulations [sampled from a flat prior on the inter-

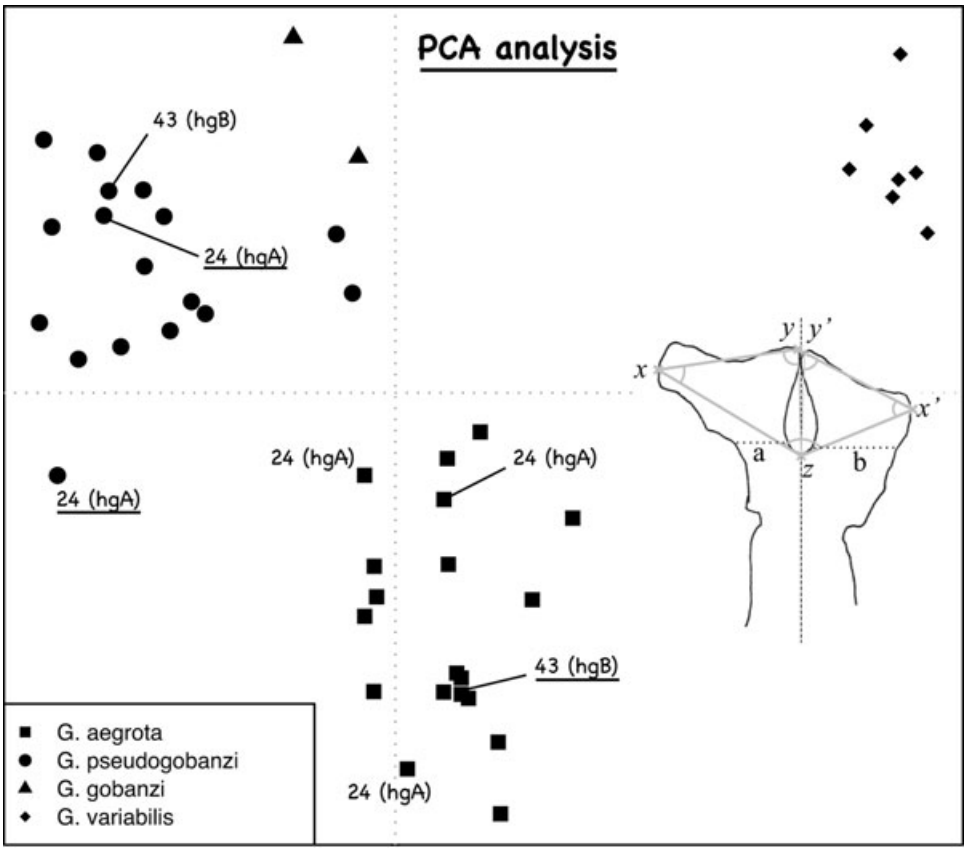

Figure 2. Result of the principal component analysis (PCA) analysis on the 13 morphological measurements (shown within the PCA analysis box). Unless otherwise indicated, *Gonioctena variabilis* males belong to haplogroup (hg) C, *Gonioctena gobanzi* males to hgD, *Gonioctena aegrota* males to hgA and *Gonioctena pseudogobanzi* males to hgB. Three specimens (underlined) out of the 45 males included in the morphometric analysis and belonging to two different haplotypes did not conform to this general pattern and are thus possible hybrids. For comparison, all other individuals of these two haplotypes (24 and 43) are also marked.

val (10^4, 10^6) years], an estimate of the probability density function of this parameter can be derived. The model used for the coalescence simulations is as follows: four populations have diverged separately (no migration) since time x (where x varies randomly between 10^4 and 10^6 years between each simulation), at which point, going backward in time, they are merged into a single ancestral population. All populations are of equal effective size, the value of this parameter being randomly chosen for each simulation between 10^3 and 5×10^5 individuals. The size of the ancestral population is chosen independently from the size of the daughter populations. Population size increase/decrease at the time of splitting is assumed instantaneous. The program TREES SIFTER 1.0 (Mardulyn, 2007; http://ueg.ulb.ac.be/treesSifter/) was used to perform these analyses. More details about the model used can be found in the program manual.

RESULTS

MORPHOLOGICAL DATA

The results of the PCA of the morphometric measurements (Fig. 2) clearly show three distinct groups of individuals. In addition, two individuals (Fig. 2, triangles) collected in Northern Italy, and thus belonging to a geographically isolated population, can easily be separated from the rest. Four distinct groups can therefore be identified based on morphological variation of the aedeagus, referred to hereafter as 'morphogroups'. The alternative hypotheses, of no distinction, or that the different forms are distributed along a morphological gradient, can safely be rejected.

The geographical distribution of the four morphogroups is as follows. Male individuals identified as belonging to morphogroup *variabilis* were found in France (location 3) and in the North of Spain (locations 4–6). Morphogroup *gobanzi* was found in our samples

Table 2. Haplotypes and morphogroups sampled in the 29 locations

Sampling location	Haplogroup	Haplotypes	Morphogroup
(1) Pertica	hgD	52 (*3*)	*Gg* (*2*)
(2) En Gauly	hgC	49 (*10*)	
(3) Al'es		51 (*1*) 50 (*2*) 49 (*4*)	*Gv* (*1*)
(4) Monegrillo		48 (*2*)	*Gv* (*1*)
(5) Osera de Ebdro		50 (*1*) 48 (*2*)	*Gv* (*2*)
(6) Pina de Ebro		50 (*2*) 48 (*3*)	*Gv* (*3*)
(7) Lisboa hgA		1 (*3*) 19 (*1*) 24 (*1*)	*Ga* (*1*)
(8) Sines		1 (*7*) 2 (*1*) 24 (*1*)	*Ga* (*2*)
(9) Evora		1 (*5*) 3 (*3*) 4 (*1*) 9 (*1*)	*Ga* (*7*)
(10) Monte S Jage		1 (*8*)	*Ga* (*1*)
(11) Mourao		1 (*3*) 5 (*1*) 13 (*2*) 22 (*1*) 23 (*1*) 24 (*2*)	*Ga* (*6*)
(12) Alconchel		24 (*3*)	*Ga* (*1*)
(13) N256		1 (*7*) 12 (*1*) 21 (*1*) 22 (*1*)	*Ga* (*3*)
(14) Beja		1 (*7*) 19 (*3*)	*Ga* (*4*)
(15) Mertola		1 (*4*) 8 (*1*) 17 (*3*) 19 (*1*) 32 (*1*)	*Ga* (*7*)
(16) Rosa de la Frontera		1 (*3*) 10 (*1*) 11 (*1*) 16 (*2*) 18 (*1*) 24 (*2*)	*Ga* (*5*)
(17) Aroche		1 (*3*) 15 (*1*) 18 (*1*) 19 (*1*) 27 (*3*) 31 (*1*)	*Ga* (*4*)
(18) Aranjuez		1 (*1*) 7 (*1*)	*Ga* (*2*)
(19) Ronda		1 (*4*) 6 (*1*) 12 (*1*) 14 (*1*) 16 (*2*) 26 (*1*)	*Ga* (*7*)
(20) Col S Ronda		1 (*7*) 19 (*1*) 20 (*1*)	*Ga* (*8*)
(21) Bracana		1 (*7*) 24 (*2*) 30 (*1*)	
(22) Buenavista		1 (*1*) 12 (*1*) 23 (*1*) 24 (*6*) 28 (*1*)	
(23) Antequera	hgA	1 (*1*) 12 (*3*) 16 (*2*) 33 (*2*)	*Ga* (*6*)
(24) Torcal	hgB	1 (*1*) 33 (*7*) 38 (*1*) 45 (*1*)	*Gp* (*9*)
(25) Villanueva		1 (*2*) 16 (*2*) 33 (*4*) 34 (*2*)	
(26) Puerto de la Torre		12 (*1*) 24 (*2*) 33 (*3*) 35 (*1*) 37 (*1*) 43 (*1*) 46 (*1*)	*Ga* (*2*), *Gp* (*8*)
(27) Chilces		1 (*1*) 24 (*1*) 25 (*1*) 29 (*2*) 33 (*5*) 35 (*1*) 41 (*1*) 43 (*1*)	*Gp* (*4*)
(28) Velez-Malaga		24 (*2*) 33 (*4*) 39 (*1*) 42 (*1*) 44 (*1*) 47 (*1*)	*Gp* (*10*)
(29) Torrox		24 (*1*) 33 (*6*) 36 (*1*) 40 (*1*) 43 (*1*)	*Gp* (*1*)

Italics indicate the number of sampled haplotypes and morphogroup individuals (identified only for males). Underlined haplotypes belong to the hgB haplogroup. *Gg*, *Gonioctena gobanzi*; *Gv*, *Gonioctena variabilis*; *Ga*, *Gonioctena aegrota*; *Gp*, *Gonioctena pseudogobanzi*.

from Italy (location 1). Morphogroup *aegrota* occurred in our samples from Portugal and Southern Spain (locations 7–20, 23, 26). The fourth morphogroup *pseudogobanzi* was found in some localities in Southern Spain (locations 24, 26–29) where *aegrota* was also found. In the present study, it is the only observed case of geographical overlap between morphogroups.

Molecular data

Sequences of the AT-rich region from all 237 individuals can be retrieved from GenBank under accession numbers EU296009–EU296245. Aligning these sequences was relatively straightforward and resulted in a data set of 384 nucleotides (including gaps) per individual. Removing sites with a minimum posterior probability below 90% generated a final data set of 292 nucleotides. Of these 292 sites, 61 were variable, among which 35 were parsimony informative. Two gaps, 1 and 2 nucleotides in length, are included. The mean base composition of the sequences is: A: 36.6%; C: 3.6%; G: 13.2%; T: 46.6%. Merging identical sequences yielded 52 unique haplotypes. Table 2 shows the number of individuals per haplotype and per sampling location.

The haplotype network inferred from the 52 haplotypic sequences is shown in Figure 3. It shows four distinct groups of sequence types that are separated from each other by a large number of substitutions. These four groups, hereafter referred to as haplogroups (hg), are: (1) hgD, formed by three individuals and comprising the only haplotype in the sample from Italy; (2) hgC (mean Kimura 2-parameters distance $d = 0.007$), formed by four haplotypes from 27 individuals collected in five localities in the North of Spain and in the South of France; (3) hgB ($d = 0.004$), comprising 15 haplotypes from 49 individuals from samples collected in Southern Spain; and (4) hgA ($d = 0.005$),

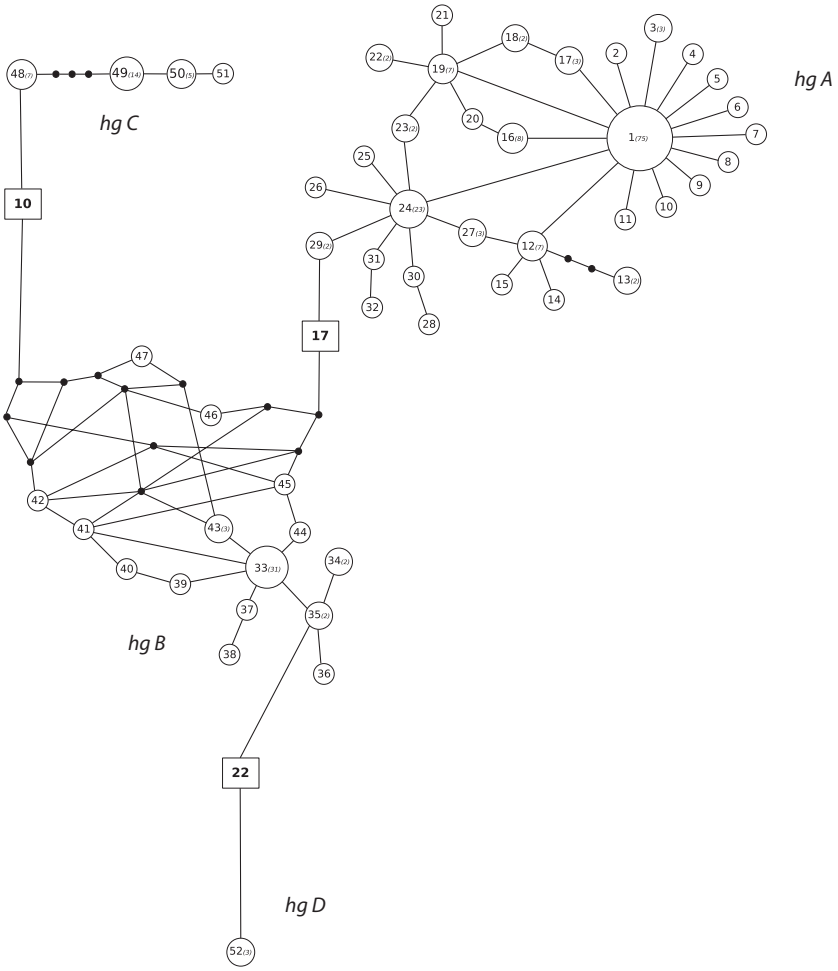

Figure 3. Median-joining network for the 52 haplotypes. Separation of the haplotypes in four haplogroups (hg), A, B, C, and D, is indicated (see text). The numbers of substitutions separating the four haplogroups from each other are indicated inside boxes. Each haplotype is identified by a number. When a haplotype is found in more than one sampled individual, the number of sampled individuals with this haplotype is shown in parentheses.

comprising 32 haplotypes from 150 individuals collected in Portugal and in Spain (Table 2).

Coalescent simulations of genealogies following the model defined a priori, coupled with the rejection algorithm described in the previous section, resulted in 1259 accepted genealogies. Less than 2% of these accepted genealogies were associated with a time of divergence below 50×10^3 years, suggesting that the four haplogroups have diverged for longer than this period. As noted earlier, this value is a very conservative estimate of the lower bound of the population splitting time, and the actual value could easily be ten-fold larger.

DISCUSSION

The morphometric measurements of the male genitalia have confirmed earlier work (Kippenberg, 2001) suggestiing that, in populations of the leaf beetle *G. variabilis s.l.* in the Mediterranean region, four well-separated morphogroups can be defined (Fig. 2). In terms of geographical distribution, our sampling suggests that the morphogroup *gobanzi* occurs in Northern Italy, the morphogroup *variabilis* (*s.s.*) occurs in Southern France and Northern Spain, and they both appear now geographically isolated from each other by the alpine mountain range in which suitable host plants are absent. By contrast, we found that in Southern Spain individuals of morphogroup *pseudogobanzi* occur in the same location as individuals of the morphogroup *aegrota*. Individuals from this last morphogroup were also found in samples collected in Portugal (Table 2).

From sequence comparisons of the AT-rich region, four distinct haplogroups can be clearly defined: hgA, hgB, hgC, and hgD (Fig. 3). The fact that these hap-

logroups are separated by a large number of substitutions suggests that they have been isolated from each other for a long period of time, either geographically or because they feed on different host plants. The highly conservative estimate of 50 000 years places their isolation well before the last glacial maximum.

Haplogroups hgC and hgD display a one-to-one correspondence with the morphogroups *variabilis* and *gobanzi*, respectively (Table 2), which appear to be geographically isolated. Haplogroups hgA and hgB on the other hand, separated by 19 mutations, display a sympatric distribution: individuals characterized as hgA were found from Portugal to Southern Spain, and individuals from hgB were found in Southern Spain (Table 2), in some of the same locations where hgA individuals were collected. This area where the two haplogroups occur in sympatry appears isolated as it is surrounded by sea to the South and East, although delimited by sierras (also called 'Sistemas Beticos', in which these beetles are absent) to the North and West. Such a pattern exhibited by hgA and hgB of a 'deep gene tree with major lineages broadly sympatric', is usually considered the result of a secondary contact between populations that have diverged previously in allopatry for a long period of time (Avise, 2000).

Moreover, in this area where both hgA and hgB occur in sympatry, there was no one-to-one correspondence between morphogroups and haplogroups. Indeed, all four possible combinations of haplogroup and morphogroup were observed: hgA-*aegrota* ($N = 65$), hgB-*aegrota* ($N = 1$), hgA-*pseudogobanzi* ($N = 7$), and hgB-*pseudogobanzi* ($N = 25$). Thus, morphogroup *aegrota* is mainly associated with hgA and *pseudogobanzi* with hgB. We interpret this as an indication that these are the ancestral associations, established before the two divergent populations or species were brought back into contact. Therefore, these data present evidence of four historically independent evolutionary units, each with a characteristic association between mitochondrial lineage and morphology of the aedeagus. The data support the use of the aedeagus morphology as a taxonomically informative trait to study the *G. variabilis* complex, as well as the taxonomic revision upgrading subspecies, corresponding to the identified evolutionary units, to species status by Kippenberg. The presence of the hgB-*aegrota* and hgA-*pseudogobanzi* associations suggests the possibility of hybridization events occurring between the two species *G. aegrota* and *G. pseudogobanzi* (i.e. between the hgA-*aegrota* and hgB-*pseudogobanzi* evolutionary units). Although another hypothesis would be that an ancestral polymorphism for hgA and hgB has been retained in the two species, it is a much less probable explanation. Indeed, this scenario would imply that most of the DNA sequence variation observed today was already present before these two species started to diverge from each other, which is unlikely.

We have observed three haplotypes (12, 24, and 43) that are associated with both the *aegrota* and *pseudogobanzi* morphogroups. Haplotype 12 belongs to hgA and has been sampled both in the west and the south of the Iberian Peninsula (locations 13, 19, 22, 23, and 26). The *G. pseudogobanzi* form characterized as haplotype 12 was found in location 26 and the *G. aegrota* form was collected at locations 19 and 23. Haplotype 24 also belongs to hgA and has been found in central and southern Spain (locations 7, 8, 11, 12, 16, 21, 22, 26, 27, 28, and 29). Again, the only *G. pseudogobanzi* form of this haplotype was collected at location 26 and the *G. aegrota* form was found in locations 11, 12, 16, 26, and 28. Two males of haplotype 43, characterized as belonging to hgB, were collected at location 26 (*G. aegrota* form) and at location 27 (*G. pseudogobanzi* form). The geographical location of these observations would appear to support the hypothesis of recent hybridization between the two species.

It is interesting, given the relatively large differences in shape of the male genitalia, that hybridization should occur at all. It has been shown that morphological differences in the aedeagus between two species of beetles can act as a strong selective agent against hybridization (Sota & Kubota, 1998). This is probably facilitated by the species having two host plants in common, *G. umbellata* and *R. sphaerocarpa*. The location of the identified hybrid zone, in South-east Spain, appears to be unusual because it does not correspond to any of the main suture zones previously identified in Europe (Taberlet *et al*., 1998). It therefore supports the view that, in Europe, each taxon has responded in its own way to the Quaternary cold periods (Taberlet *et al*., 1998).

If the four species have diverged from each other for such a long period that large genetic distances now separate their mitochondrial alleles, what barriers have isolated them from one another? The geographical distribution of *G. gobanzi* suggest that it was restricted to Italian refuges during the last ice age, and that its subsequent expansion out of the Italian Peninsula was barred by the alpine mountain range. Species *G. variabilis* has been shown to occur on both sides of the Pyrenees, in France and in Northern Spain, implying that it has crossed the Pyrenees at least once (in either direction). Because one of the four haplotypes (50) of hgC is present on both sides of this mountain range, it is likely that one crossing has occurred relatively recently. Species *G. aegrota*, which is found in North Africa, in Portugal, in Southern Spain, and also in the North of Spain (Kippenberg,

2001), has the widest geographical distribution. It is more difficult to imagine what barrier has constrained its geographical range in the past. One possible scenario, although by no means the only one, is that *G. aegrota* has recently invaded Europe from Africa, crossing the Gibraltar strait, expanding its range northward, and coming into secondary contact with at least one other species, *G. pseudogobanzi*. The alternative scenario is that *G. aegrota* was already present in the Iberian Peninsula during the last glacial period, and that it survived there in a refuge somehow isolated from the refuge of *G. pseudogobanzi*. Such hypotheses, implying long periods of physical separation between the groups, is necessary to explain our sequencing data, which show that divergence in this *Gonioctena* complex is well advanced, even though the comparison between the mitochondrial sequences and morphological variation reveals some unexpected hybridization. Finally, it is conceivable that the two evolutionary units hgA-*aegrota* and hgB-*pseudogobanzi* have diverged in sympatry, but on different host plant species. *Gonioctena* species, like most leaf beetles, are characterized by a highly specialized plant diet (Mardulyn, Milinkovitch & Pasteels, 1997). Each species within the *G. variabilis* complex is associated with a small number of different host plant species (see above). A strong difference in host plant preference is likely to be at the base of several speciation processes in this genus, and could have played a role in the isolation of the four evolutionary units studied here. To achieve the pattern of haplotype diversity observed between hgA and hgB, the specialization on different host plants would need to be strong enough to prevent any exchange between the two sympatrically isolated populations. Under this hypothesis, after a long period during which both evolutionary units would subsequently have been specialized on different host plants, at least one of them (hgA-*aegrota*) would have switched to a more generalized diet, which would explain that they share two host plant species today.

Note that this possible shift to a less restricted diet in hgA-*aegrota* is not necessarily associated with a hypothesis of sympatric divergence. It could also explain the observed recent hybridization in the case of an allopatric divergence. The two host plants that are used only by *G. aegrota*, *R. monosperma*, and *Cytisus grandiflora*, have their geographical distribution within the Iberian Peninsula restricted to the South-West and the West (where *G. pseudogobanzi* is absent), respectively, whereas *Genista umbellata*, on which both sympatric species are found, is restricted to the South-East of the Iberian Peninsula (all three plant species are also found in North Africa) (Castroviejo *et al*., 1999). It is therefore possible that both *Gonioctena* evolutionary units were at some point specialized on different plant species, located in different parts of the Iberian Peninsula (e.g. hgA-*aegrota* on *R. monosperma* and *C. grandiflora* in the West and hgB-*pseudogobanzi* on *G. umbellata* to the South-East), and that, later, one or both has extended its diet to one or two additional plant species, including the more widespread *R. sphaerocarpa* (found more or less all over the Iberian Peninsula as well as in North Africa), leading to the overlap of their geographical distribution. To further explore the impact of the host plants on the evolution of the *G. variabilis* complex, a knowledge of the past geographical distribution of these plant species is needed. To the best of our knowledge, such historical information remains lacking.

It is worth noting that of 128 males sampled by us in the south of Spain (not all individuals sequenced), 117 belonged to morphogroup *pseudogobanzi*, and only 11 to morphogroup *aegrota* (including six found in single locality, location number 23). This small proportion of *aegrota* contrasts with the higher proportion (16 out of 63) of hgA haplotypes (the haplogroup associated with *aegrota* across most of its geographical distribution) found in the same area. It may reflect the difficulty of the *G. aegrota* to colonize this region, possibly because it is not sufficiently adapted to the available host plants, whereas the hgA mitochondrial genome is spreading more freely through introgression into *G. pseudogobanzi*. Similar cases of mtDNA introgression have already been documented in recent studies (e.g. Melo-Ferreira *et al*., 2005).

In the present study, the combination of mtDNA sequences with morphological characters has allowed us to identify a hybrid zone between two historically independent evolutionary units of the *G. variabilis* species complex. Further work will be necessary to characterize this hybrid zone (e.g. are the hybrid individuals fertile and what is the geographical extent of the hybrid zone?) and to find out whether other such hybrid zones are present in other locations, for example in the north of Spain, where *G. variabilis* and *G. aegrota* could meet. One difficulty with the present system is that morphogroups can be identified only by dissecting male individuals, and that hybrids cannot be identified morphologically. For this reason, this hybrid zone may be more difficult to characterize than in other beetle hybridization studies (Sota & Kubota, 1998; Peterson *et al*., 2001).

ACKNOWLEDGEMENTS

We are grateful to J. Araujo, J. C. Bourdonné, M. Daccordi and J. M. Vela for their help in collecting the insects, as well as to S. J. E. Baird, R. C. Hardwick, and two anonymous reviewers for their help in sig-

nificantly improving previous versions of this manuscript. J.M.P. warmly thanks Mireille Vidal for her hospitality in Chilces during field work, and Jorge Araujo for his in Foros do Queimado. This research was supported by grants from the Belgian Fonds de la Recherche Fondamentale et Collective (2.4505.98 and 2.4543.03). L.G. is a Research Fellow and P.M. is a Research Associate, both at the Belgian Fund for Scientific Research (FNRS).

REFERENCES

Avise JC. 2000. *Phylogeographyraphy. The history and formation of species*. Cambridge, MA: Harvard University Press.

Avise JC. 2004. *Molecular markers, natural history, and evolution*. Sunderland, MA: Sinauer Associates.

Bandelt H-J, Forster P, Röhl A. 1999. Median-joining networks for inferring instraspecific phylogenies. *Molecular Biology and Evolution* **16:** 37–48.

Bourdonné JC, Doguet S. 1979. Contribution à l'étude des Gonioctena Chevr. (Phytodecta Kirby) d'Afrique du Nord. *Nouvelle Revue d'Entomologie* **IX:** 49–58.

Castroviejo S, Talavera S, Acedo C, Romero Zarco C, Sáez L, Salgueiro FJ, Velayos M. 1999. *Flora Ibérica: Plantas vasculares de la Península Ibérica e Islas Baleares Vol. VII (I) Leguminosae (partim).* Madrid: Real Jardín Botánico, CSIC.

Crozier RH, Crozier YC, Mackinlay AG. 1989. The CO-I and CO-II region of honeybee mitochondrial DNA: evidence for variation in insect mitochondrial evolutionary rates. *Molecular Biology and Evolution* **6:** 399–411.

Hewitt GM. 2001. Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology* **10:** 537–549.

Hewitt GM. 2003. Ice ages: species distribution and evolution. In: Rothschild L, Lister A, eds. *Evolution on planet earth*. London: Academic Press, 339–362.

Hofreiter M, Serre D, Rohland N, Rabeder G, Nagel D, Conard N, Munzel S, Paabo S. 2004. Lack of phylogeography in European mammals before the last glaciation. *Proceedings of the National Academy of Sciences of the United States of America* **101:** 12963–12968.

Kippenberg H. 2001. Neuordnung des *Gonioctena variabilis* – Gruppe (Coleoptera, Crysomelidae). *Entomology Bulletin* **97:** 13–34.

Kumar S, Tamura K, Nei M. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics* **5:** 150–163.

Loytynoja A, Milinkovitch MC. 2003. A hidden Markov model for progressive multiple alignment. *Bioinformatics* **19:** 1505–1513.

Mardulyn P. 2007. Trees Sifter 1.0: an approximate method to estimate the time to the most recent common ancestor of a sample of DNA sequences. *Molecular Ecology Notes* **7:** 418–421.

Mardulyn P, Milinkovitch MC. 2005. Inferring contemporary levels of gene flow and demographic history in a local population of the leaf beetle *Gonioctena olivacea* from mitochondrial DNA sequence variation. *Molecular Ecology* **14:** 1641–1653.

Mardulyn P, Milinkovitch MC, Pasteels JM. 1997. Phylogenetic analyses of DNA and allozyme data suggest that Gonioctena leaf beetles (Coleoptera; Chrysomelidae) experienced convergent evolution in their history of host-plant family shifts. *Systematic Biology* **46:** 722–747.

Mardulyn P, Termonia A, Milinkovitch M. 2003. Structure and evolution of the mitochondrial control region of leaf beetles (Coleoptera: Chrysomelidae): a hierarchical analysis of nucleotide sequence variation. *Journal of Molecular Evolution* **56:** 38–45.

Melo-Ferreira J, Boursot P, Suchentrunk F, Ferrand N, Alves PC. 2005. Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. *Molecular Ecology* **14:** 2459–2464.

Peterson MA, Dobler S, Holland J, Tantalo L, Locke S. 2001. Behavioral, molecular, and morphological evidence for a hybrid zone between *Chrysochus auratus* and *C. cobaltinus* (Coleoptera: Chrysomelidae). *Annals of the Entomological Society of America* **94:** 1–9.

Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW. 1999. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular Biology and Evolution* **16:** 1791–1798.

SotaT, Kubota K. 1998. Genital lock-and-key as a selective agent against hybridization. *Evolution* **52:** 1507–1513.

Su ZH, Tominaga O, Okamoto M, Osawa S. 1998. Origin and diversification of hindwingless Damaster ground beetles within the Japanese islands as deduced from mitochondrial ND5 gene sequences (Coleoptera, Carabidae). *Molecular Biology and Evolution* **15:** 1026–1039.

Swofford DL. 2003. *PAUP*, phylogenetic analysis using parsimony (*and other methods)*, Version 4b10. Sunderland, MA: Sinauer Associates.

Taberlet P, Fumagalli L, Wust-Saucy A-N, Cossons J-F. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7:** 453–464.

Tavaré S, Balding DJ, Griffiths RC, Donnelly P. 1997. Inferring coalescence times from DNA sequences. *Genetics* **145:** 505–518.

Zakharov EV, Caterino MS, Sperling FA. 2004. Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera: Papilionidae). *Systematic Biology* **53:** 193–215.