Lees et al Frontiers 2011 Main and Supp

DATASET · JUNE 2013

READS
165

8 AUTHORS, INCLUDING:

David C. Lees
Natural History Museum, London
100 PUBLICATIONS **3,177** CITATIONS
SEE PROFILE

Rodolphe Rougerie
Muséum National d'Histoire Naturelle
95 PUBLICATIONS **939** CITATIONS
SEE PROFILE

Sylvie Augustin
French National Institute for Agricultural Rese…
54 PUBLICATIONS **1,361** CITATIONS
SEE PROFILE

Carlos Lopez Vaamonde
French National Institute for Agricultural Rese…
92 PUBLICATIONS **1,796** CITATIONS
SEE PROFILE

Tracking origins of invasive herbivores through herbaria and archival DNA: the case of the horse-chestnut leaf miner

David C Lees[1,2*], H Walter Lack[3], Rodolphe Rougerie[4,5], Antonio Hernandez-Lopez[1], Thomas Raus[3], Nikolaos D Avtzis[6], Sylvie Augustin[1†], and Carlos Lopez-Vaamonde[1†]

Determining the native geographic range or origin of alien invasive species is crucial to developing invasive species management strategies. However, the necessary historical dimension is often lacking. The origin of the highly invasive horse-chestnut leaf-mining moth *Cameraria ohridella* has been controversial since the insect was first described in 1986 in Europe. Here, we reveal that herbarium collections across Europe indicate a Balkan origin for *C ohridella*. We successfully amplified nuclear DNA and mitochondrial DNA barcode fragments from larvae pressed within leaves of herbarium samples collected as early as 1879. These archival sequences confirm an identity of *C ohridella* and set back its history in Europe by more than a century. The herbarium samples uncovered previously unknown mitochondrial haplotypes and locally undocumented alleles, showing local outbreaks of *C ohridella* back to at least 1961 and dynamic frequency changes that may be associated with road development. This case history demonstrates that herbaria are greatly underutilized in studies of insect–plant interactions, herbivore biodiversity, and invasive species' origins.

Front Ecol Environ 2011; 9(6): 322–328, doi:10.1890/100098 (published online 21 Jun 2011)

Herbarium specimens have become pivotal in fields such as molecular phylogenetics (eg Chase *et al.* 1993) and plant invasion biology (eg Delisle *et al.* 2007; see also the bibliography provided in the Web-only materials). However, for studies of herbivore–plant interactions, this resource has remained largely untapped (Funk 2004, but see Abbott *et al.* [1999] and bibliography in the Web-only materials), despite the increasing interest in amplifying DNA from museum and herbarium material (eg Thomsen *et al.* 2009; Web-only materials). Here, we report a case study involving the spatiotemporal origin of a prominent invasive insect herbivore – one belonging to the feeding guild of "leaf miners" – in which evidence from herbaria played a critical role.

The origin of the leaf-mining moth *Cameraria ohridella* Deschka and Dimić 1986 (Lepidoptera: Gracillariidae), currently ravaging ornamental white-flowering horse-chestnut trees (*Aesculus hippocastanum* L, Sapindaceae) throughout Europe, has been debated since the moth's discovery in Macedonia in 1984 (Deschka and Dimić 1986; Grabenweger and Grill 2000; Hellrigl 2001). There have been many studies on *C ohridella*'s biology and control (for a list, see Lees *et al.* 2011); *C ohridella* belongs to a genus that is diverse in North America, although with few species in central and east Asia and which was previously unknown in Europe. Yet it seemed unlikely that the conspicuous leaf mines of this genus would have escaped the notice of entomologists working in the Balkans before 1984.

The horse-chestnut was first discovered growing wild in central Greece by English traveller John Hawkins in 1795 (Lack 2002). Fossil evidence has revealed that the horse-chestnut's present-day native distribution – spanning Albania, Greece, Macedonia, and debatably also Bulgaria (Adamović 1908) – represents a relict of a wider Tertiary range, including present-day southeast Germany and southern Poland during the Pliocene and southeastern France and northeastern Spain during the Pleistocene (Mädler 1939; Postigo-Mmijarra *et al.* 2008; Harris *et al.* 2009).

The genetic study done by Valade *et al.* (2009) transformed our understanding of the evolutionary and invasive history of *C ohridella* on horse-chestnut, revealing unexpected diversity for such a recently invasive herbivore. Twenty-five mitochondrial *cytochrome c oxidase subunit 1* [COI] haplotypes – each differing by at least one single nucleotide mutation – were found in this species. Valade *et al.* (2009) noted that the invasive diversity was highly restricted: only haplotypes "A", "B", and "C" have invaded western and eastern Europe in the past quarter century. These haplotypes appear closely related, if not ancestral. In the network of haplotypes reconstructed by Valade *et al.* (2009), the centrally positioned haplotype "A" was one mutational step distant from a haplogroup

[1]*INRA UR0633, Zoologie Forestière, Orléans, France;* [2]*Department of Entomology, Natural History Museum, London, UK* [*]*(dclees@gmail.com);* [3]*Botanischer Garten und Botanisches Museum Berlin-Dahlem, Freie Universität Berlin, Germany;* [4]*Biodiversity Institute of Ontario, University of Guelph, Ontario, Canada;* [5]*Université de Rouen, ECODIV, UPRES-EA 1293, Mont Saint Aignan, France;* [6]*Department of Forestry, Technological Educational Institute of Kavala, Drama, Greece;* [†]*contributed equally to the senior authorship*

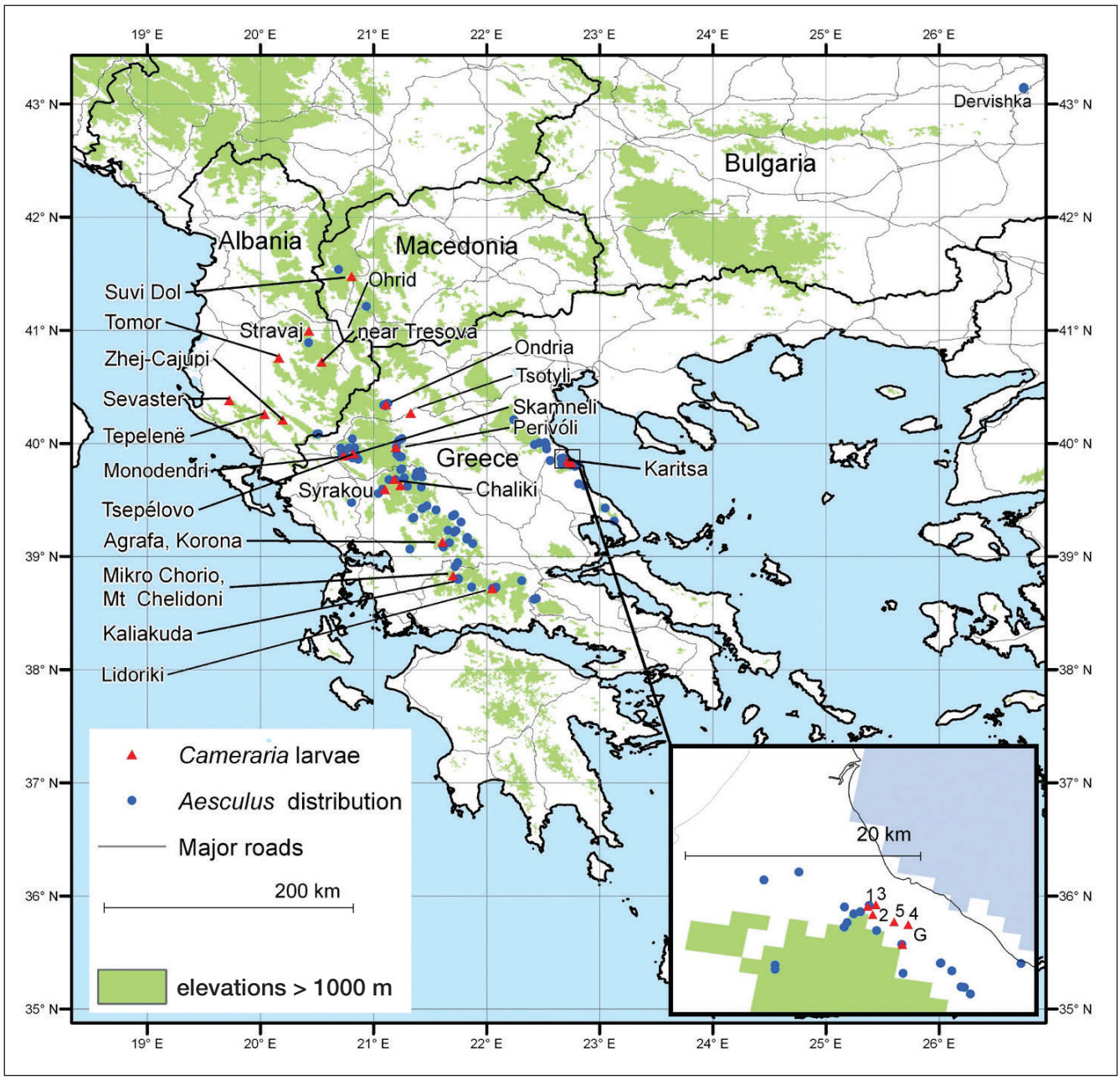


Figure 1. *Distribution of* Cameraria ohridella *larvae and pupae found in herbarium specimens of* Aesculus hippocastanum *(1879–1981; red triangles) as compared with or in addition to the known natural distribution of horse-chestnut (best known for Greece: Avtzis* et al. *2007; Valade* et al. *2009; Flora Hellenica database; blue circles). Many sites are remote from principal roads but one now serves Karitsa. At Karitsa (inset), recent* C ohridella *samples (2008: points 1, 2, and 3) are very close to archival samples (1974: point 4; 1981: point 5, Raus collections; 1936: point G, Grebenchikoff collection).*

comprising haplotypes "B" and "C" (together with haplotype "Y"). Moreover, only haplotype "A" is dominant (at a frequency of 67–100%), not only throughout its expanding range in Europe but, intriguingly, also in about 90% of Balkan relict horse-chestnut sites (47–91% frequency; Valade *et al.* 2009). Most COI haplotypes of C *ohridella* (not counting those found only or also in ornamental plantings in the Balkans) were either endemic (14) or very narrowly distributed (three) among natural horse-chestnut populations, which are often isolated in remote canyons, with up to 11 haplotypes reported from a single site (Perivóli, in central Greece; Figure 1).

Valade *et al.* (2009) also examined nuclear genetic variations using six microsatellite loci (short length-polymorphic tandem repeating regions) developed for C *ohridella* (Mari Mena *et al.* 2008). Consistent with the COI trend, these data showed a decrease in diversity away from the Balkans and from natural to ornamental populations. Also consistent with a Balkan origin, a higher frequency of rare or "private" alleles (here referring to population-specific or unique length variants of a single genetic locus) was found in natural host-plant populations.

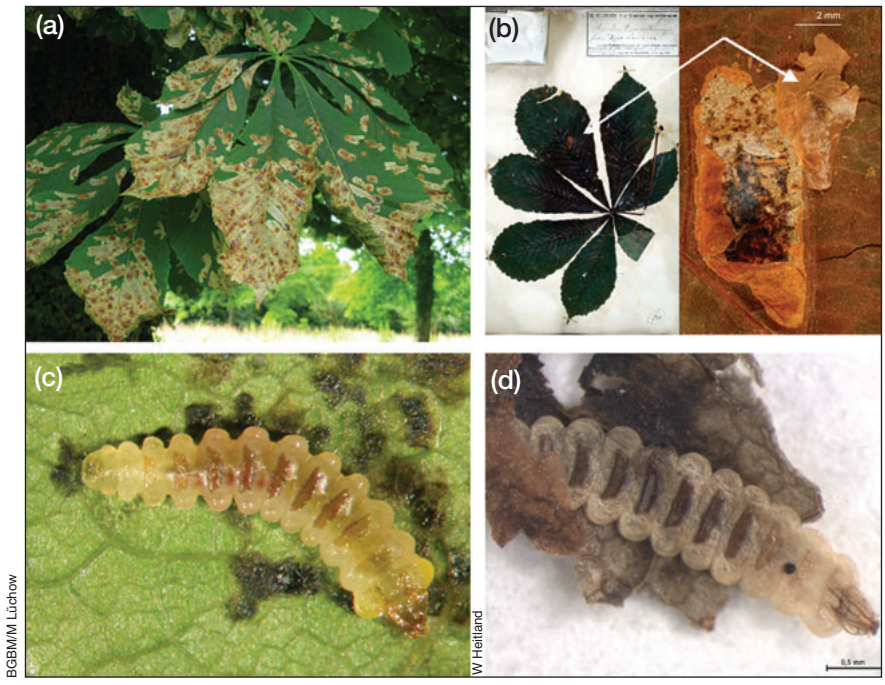


Figure 2. *(a) Typical late-summer damage by horse-chestnut leaf miner. (b) Archival herbarium specimen (Heldreich 11 Aug 1879, Kew; represented at three other herbaria) with mine and extracted "spinning-stage" larva of* C ohridella *(indicated by the white arrows); scale bar 2 mm. (c) Modern "tissue-feeding" L4-stage larva of* C ohridella*. (d) Pressed L4-stage larva of* C ohridella *extracted from archival leaf mine; scale bar 0.5 mm (Markgraf 28 Jun 1928, #1513, Berlin).*

The key question posed by Valade *et al.* (2009) was that, if C *ohridella* was Balkan in origin, why had the modern explosion apparently been delayed until two decades ago? On the basis of Valade *et al.*'s (2009) data, we suggest two spatiotemporally distinct hypotheses to explain why the mitochondrial DNA pattern of European invasion of C *ohridella* is so dominated by haplotype "A": (null hypothesis H_0) haplotype "A" was historically widespread and abundant among natural sites and is therefore most likely to have spread to ornamental plantings; (alternative hypothesis H_1) haplotype "A" was, like other haplotypes, rare and very localized in the Balkans but has only recently become highly invasive – even (re-)invading relict horse-chestnut sites. Modern distributional data (eg occurrence of haplotype "A" in the supposedly natural site of Dervishka in eastern Bulgaria; Figure 1) are consistent with either hypothesis. Ancient wide prevalence and abundances similar to modern levels would be consistent with H_0, whereas a sharp temporal frequency change among natural sites of haplotype "A" would be consistent with H_1.

Although the level of genetic diversity reported between remote Balkan mountain ravines would be difficult to explain if the moth had newly colonized this area, direct proof of a Balkan origin has been lacking. We know of no pre-1984 C *ohridella* specimens conserved in entomological collections. The puzzlingly rapid invasion (since 1989) calls for a fresh investigative approach.

Given a Balkan origin and a quarter-century-long record of outbreaks, a historical trace of C *ohridella* should exist in herbarium collections of *A hippocastanum*. For this study, we searched horse-chestnut collections of six European herbaria for leaf mines of C *ohridella*. Of particular relevance, we located a temporal series of three collections (1936–1981) from a single horse-chestnut locality (Karitsa, in eastern Greece; Figure 1) for which we already had field samples of larvae from 2008, as well as central Greek and Albanian horse-chestnut collections from the early 1960s in close proximity to populations examined by Valade *et al.* (2009). If archival C *ohridella* mines existed in such herbarium leaf samples, we wanted to know (1) how frequently such mines were preserved, (2) whether any pre-imaginal stages that could be extracted contained analyzable mitochondrial or nuclear DNA, and (3) whether these data might illuminate the colonization history of this invasive species. In particular, we wished to confirm the historical Balkan origin, as suggested by Valade *et al.* (2009), and to test the aforementioned hypotheses.

Materials and methods

Preserved remains of C *ohridella* larvae and pupae were carefully excised from mines in *A hippocastanum* leaf samples in herbaria (eg Figure 2b). Leaf-mine density on leaves was measured by surface-area imaging. We used standard methods for extracting archival samples and amplifying the DNA, either in short fragments (mitochondrial "mini-barcodes"; for example, see Lees *et al.* 2010) or as nuclear microsatellites, but using primer pairs developed specifically for C *ohridella*. Precautions were taken throughout against genetic contamination. Further details are provided in the Web-only materials.

Results

Herbarium analyses

Examination of leaf mines in different historical herbarium collections – at Kew, UK (Figure 2b; WebFigure 1, a–c; WebFigure 2, d–f); Berlin, Germany (Figure 2d; WebFigure 1, d–f); Paris, France (WebFigure 2, a–c); Vienna, Austria (http://herbarium.univie.ac.at/database/detail.php?ID=149323; WebTable 1); London, UK (Natural History Museum herbarium; notes, WebTable 1); and Jena, Germany (Haussknecht Herbarium; WebFigure 3) – as well as in material in one author's (TR's) collection) reveals that C *ohridella* has indeed long been a resi-

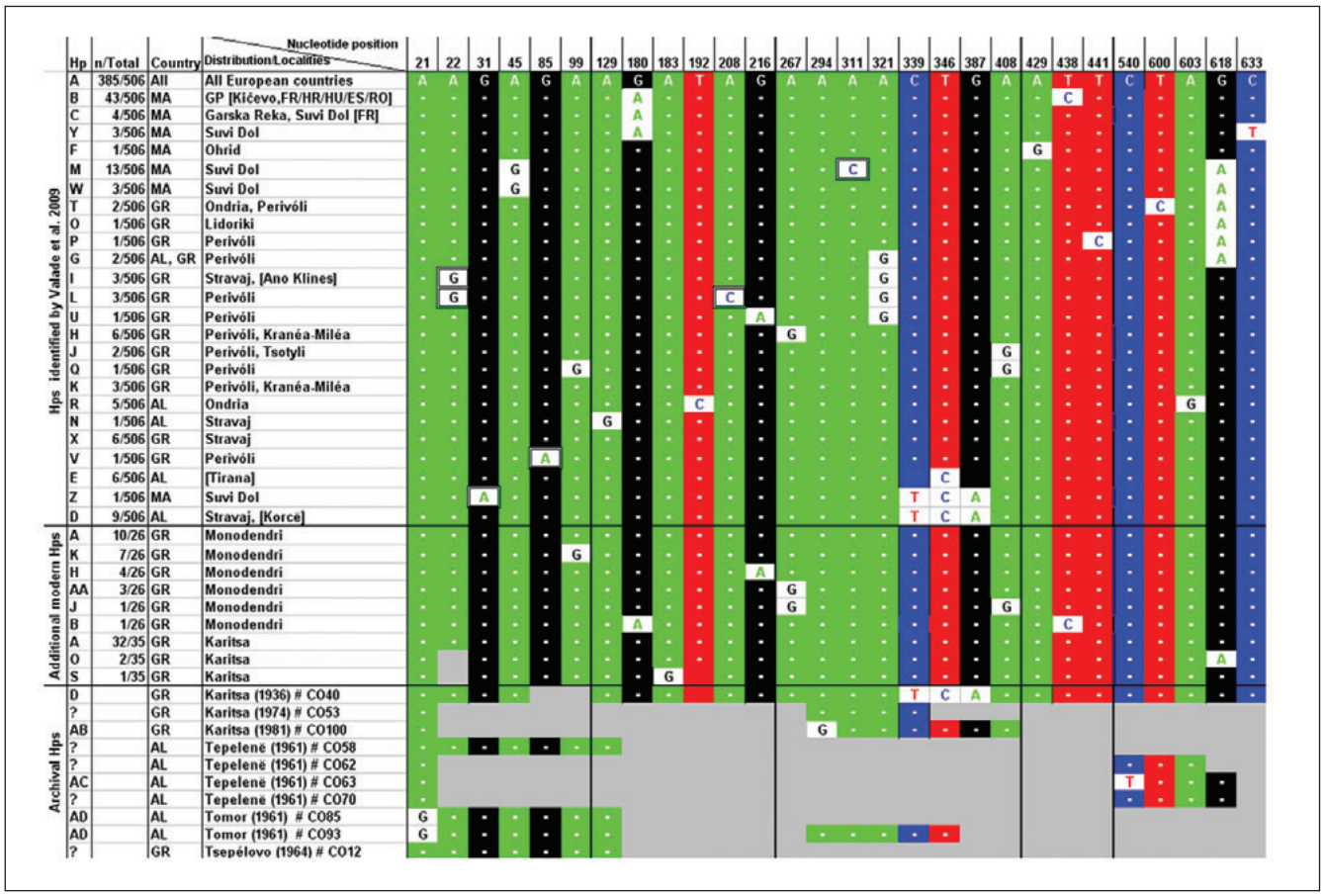


***Figure* 3.** *COI variability and geographic distribution of haplotypes among 577 sequenced individuals of* Cameraria ohridella. *The 25 haplotypes identified by Valade* et al. *(2009) (upper section, ordered by haplogroups for ease of comparison), plus haplotypes from additional modern sampling (middle section, Karitsa and Monodendri), and COI fragments from herbaria (lower section; bold vertical lines delineate the five sequenced regions), are shown for all variable nucleotides of 658-bp COI barcode, numbered from first complete codon of the barcode. Novel haplotypes (this study) are indicated in bold. Incomplete sequences are gray filled. All sequences are shown by white dot-plot relative to haplotype "A". Mutations implying amino-acid changes (for five endemic haplotypes) are outlined. Geographic abbreviations: AL = Albania; GR = Greece; MA = Macedonia; GP = Galicica National Park; FR = France; HR = Croatia; HU = Hungary; ES = Spain; RO = Romania. Samples from ornamental plantings are designated by square brackets. Hp = haplotype.*

dent among remote mountain populations of horse-chestnut in the southern Balkans (Figure 1). No archival mine was observed in horse-chestnut collections derived from ornamental plantings in any European herbarium outside the Balkans, dating back to at least 1737 (Lack 2002), nor from a 1928 herbarium specimen from the Dervishka area in eastern Bulgaria (Figure 1; WebTable 1). We examined historical horse-chestnut collections with leaf mines collected between 1879 and 1981 according to their extensive documentation (eg Heldreich 1879; Baldacci 1897; Markgraf 1931; Grebenchikoff 1938; Raus 1980; Figure 2b; WebTable 1; WebFigures 1–3). Our findings therefore set back the history of this moth in Europe, from its first reported collection in 1984 (Deschka and Dimić 1986) to at least 132 years ago.

Herbaria are not the most obvious source of leaf-mine density data. Surprisingly, within the 24 761 cm^2 of archival horse-chestnut specimen leaf–surface area examined, we counted >2499 mines, representing an overall density > 0.1 mines per cm^2 (2.7 mines per leaflet); from these, we excised 58 caterpillars or chrysalids (early developmental stages or instars) of *C ohridella* (WebTable 8). Not surprisingly, in most archival collections, mine densities were low, but a general increase in mine-containing leaf specimens collected between 23 June and 22 September (WebTables 1 and 7) showed that mine densities depend on the season of collection and possibly on tree-specific traits, such as dense, potentially mite-harboring hairs under leaves (eg J Mattfeld's collections from Kaliakuda that lacked mines in early August; WebTable 1). In some cases, however, as for FK Meyer's collections from September 1961 at Sevaster, Tomor, and Tepelenë, Albania (eg WebFigure 3), exceptional mine densities were found, which were hardly distinguishable from attack levels characteristic of late-summer foliage in 21st-century European parks. A single leaflet was found to contain as many as 21 mines (sample Baldacci-129; WebTable 1; WebFigure 2a) or more than 32 mines (sample Meyer-6380; WebFigure 3). Evidently, while botanists often deliberately tried to avoid collecting visibly damaged

leaves (GG Aymonin, collector of a 1964 specimen, pers comm) or actually disguised mines on herbarium sheets (Heldreich and Baldacci; Figure 2b; WebFigure 2, a and b), sometimes they had no choice (Meyer; WebFigure 3).

Molecular analyses

We amplified mitochondrial DNA from a time-series of geographically close specimens from Karitsa (Mount Ossa), eastern Greece (collected in 1936, 1974, 1981, and 2008; Figure 1 inset; WebTable 1). Genomic DNA from one of three larvae collected in 1936 (WebFigure 1c) was successfully amplified in five COI mini-barcode fragments totalling 600 base pairs (bp; Figure 3). Identifications based on morphology of archival mines and caterpillars (Figure 2) were confirmed by these DNA sequences (Figure 3). The near-complete barcode stitched together from 1936 is indistinguishable from haplotype "D" (Figure 3), previously unique to Stravaj, eastern Albania (Valade *et al.*'s Figures 2 and 3; locality label corrected in our Figure 1) – and nearby Korcë (southeastern Albania), where this haplotype is invasive. The mini-barcodes from the Karitsa herbarium collections revealed a novel haplotype "AB" (in the sample from 1981). We further sequenced mini-barcodes from six Albanian specimens from 1961 (Tomor, Tepelenë), and from a 1964 specimen from the northern Pindus (Tsepélovo) (Figure 1; WebTable 1). One Tepelenë sample belongs to a new haplotype, "AC", and two Tomor larvae to another new haplotype, "AD". For other Tepelenë and the Greek (Tsepélovo) sequences, we could not exclude haplotype "A" (Figure 3).

Full DNA barcodes from modern samples were newly sequenced to provide comparative haplotype frequency datasets to herbarium samples. Among 35 modern Karitsa larvae (from 2008), 91.4% were haplotype "A", 5.7% were haplotype "O" (previously detected at Lidoriki in the southern Pindus Mountains; Figure 1), and 2.9% (one larva) was of a previously undetected haplotype "S" (Figure 3). We had no new Albanian samples very close to Tepelenë or Tomor, but Stravaj – sequenced by Valade *et al.* (2009) with five haplotypes (including 48% "A" and 18.5% "D") – is located 35 km from Tomor. Among 26 larvae from another modern sample (Monodendri, 7.3 km west of Tsepélovo), we found haplotype "A" (38.5%), "K" (27%), "H" (15%), "J" and "B" (each 4%), along with another novel haplotype "AA" (11.5%; Figure 3).

Comparing the 30 haplotypes found among the 577 individuals of the combined (up to 658 bp) DNA barcode dataset, dot-plot analysis (eg for haplotype "A" in Figure 3) shows that haplotype "A" is the only one not characterized by any uniquely shared mutation among individuals, meaning that it is also the hardest to detect. There is substantial intraspecific variability in *C ohridella* (Figure 3), with pairwise divergences between COI sequences averaging 0.57% and up to 1.12% (haplotypes "Z"–"U").

Of the five different microsatellite markers tested by Mari Mena *et al.* (2008), only those generating the shortest amplicons (amplified DNA fragments; 93–108 bp for Ohrid2814 and 102–132 bp for Ohrid2782; WebTable 2) yielded data for herbarium samples. Among 54 genomic DNA extracts, 63% amplified for Ohrid2814, whereas 24% amplified for Ohrid2782 (WebTables 2 and 3). Amplicons were obtained even among the oldest samples we extracted (dating from 1879 to 1885). Ohrid2814 alleles common in modern Balkan samples (Valade *et al.*'s [2009] dataset; WebTable 3) were also common in archival material. Unfortunately, rarer or private alleles, constituting only 30% of the modern Karitsa population, were not detected. Just one possibly novel allele with an identical length-variant (homozygote), measuring 114 bp, was found for marker Ohrid2782 in the archival caterpillar CO51 from Karitsa (from 1974): only Stravaj specimen RV130C approximates its length among all Valade *et al.*'s 467 microsatellite-amplified samples (WebTable 3). By contrast, the only homozygotic alleles from the modern Karitsa sample (for 17% of samples amplified) measured 122–126 bp (WebTable 3). The Tsepélovo (1964) alleles, homozygotic for each marker, were also detected in one individual each at the nearby modern Monodendri population (indicated in bold in WebTable 3).

We conducted a population genetic analysis on the 14 amplifications we obtained from larvae extracted from the single Tepelenë leaflet as compared with samples from the nearest archival and modern populations. Genetic diversity was not significantly different between Stravaj and the two archival (Tepelenë and Tomor) populations (WebTable 4). However, there were significant ($P < 0.05$) differences in genetic structure between individuals of archival and modern populations (WebTable 5). This agrees with the results from Structure software analysis (WebTable 6), which group ancient individuals in a single cluster, as well as with the presence of locally private alleles (WebTable 7). Despite the modest sample sizes, these results are consistent with historical isolation of the Tomor and Tepelenë populations from that of Stravaj.

The Tepelenë leaflet, from which we successfully amplified microsatellites for 14 out of 20 extracted larvae, belonged to five sheets from the Jena herbarium containing an estimated 296 mines (eg WebFigure 3). Remarkably, five sheets from the Tomor collection contained an estimated 1365 mines, chiefly 1st and 2nd instar larvae (WebTable 8).

Discussion

We have demonstrated the importance of herbarium samples in tracking the colonization history of an invasive herbivore. The conspicuous mines of Europe's only known *Cameraria* species had been overlooked in botanical collections, not only during the quarter-century since the moth's formal discovery but also from the earliest extant collections of naturally occurring horse-chestnut trees in central Greece in 1879.

Archival herbarium specimen data from 1879 to 1981 add a historical dimension to a previous molecular phylogeographic study (Valade *et al.* 2009). The ancient presence of *C ohridella* in the southern Balkans revealed by well-documented herbarium samples has been confirmed by DNA analysis of early-stage specimens inadvertently pressed by botanists. The herbaria samples also contained three novel haplotypes undetected in modern sampling at Karitsa, Tepelenë, and Tomor, as well as at least locally private alleles, in the latter two cases, providing a potential baseline for modern sampling. Presence of mitochondrial haplotype "D" at Karitsa (from a sample collected in 1936), considering its modern localized distribution, suggests either ancient connectivity – or historical transport by humans – between the *C ohridella* populations of Karitsa and Stravaj, currently isolated by the ~235 km separating the western (Pindus) and eastern (Vermio) montane distribution ranges of *Aesculus* (Figure 1).

Herbaria also confirm the historical absence (implicit also from the chronology of the moth's spread) of an earlier invasion outside the Balkans and illuminate the apparent mystery of *C ohridella*' s sudden modern expansion. As evident from sparse historical collections of wild horse-chestnut trees, much of this region was highly inaccessible (Adamović 1908; Markgraf 1931; Valade *et al.* 2009). However, the construction of modern roads (Figure 1), such as the one into Karitsa in the early 1970s (WebFigure 4), could well have disrupted topographic and biogeographic isolation of moth/host-plant populations, allowing rapid vehicular transport of adult females, larvae, and overwintering pupae by means of fallen leaves containing living stages (Gilbert *et al.* 2005).

Although the historical data are scarce, at most one (from a 1974 sample) out of three (from samples collected between 1936 and 1981) archival amplifications was haplotype "A", as compared with 32 out of 35 from field samples from 2008 (Figure 3). These data favor H_1, suggesting that saturation with "A" (and possibly the arrival of "O") probably occurred after road surface construction in the 1970s. Moreover, during our modern sampling, we had suspected that the horse-chestnut itself was introduced to Karitsa. Instead, unique *C ohridella* haplotypes (if not also genotypes) suggest a natural host-plant site, as already evident from the occurrence of horse-chestnut within plant communities rich in typically relict species (Raus 1980).

However, our new data are insufficient to reject H_0 (ie that haplotype "A" was historically present at all natural sites surveyed, including Karitsa). Even in the isolated Greek canyon of Monodendri, the frequency of haplotype "A" was 38.5%. The sole natural site where Valade *et al.* (2009) did not detect haplotype "A" was Ondria, an isolated karstic plateau about 7.5 km from the nearest road, but here the sample size (n = 6) was low. The single, mine-less archival specimen from the supposedly natural site in eastern Bulgaria (collected in 1928) provides limited negative evidence; Valade *et al.* (2009) found only "A" there, with no endemic haplotypes. We also cannot determine whether sites in Albania (Tepelenë) and Greece (Tsepélovo) had haplotype "A" at the time of historical sampling, given that the COI fragments of six specimens were too short to exclude this possibility (Figure 3). Although available data do not yet rule out H_0, H_1 seems at least locally validated, and further herbarium and modern sampling in the Balkans should address the degree of genetic frequency changes to clarify the extent to which particular haplotypes of *C ohridella* may be invading native moth populations.

These herbarium-derived data underscore the need for a temporally explicit model of *C ohridella* phylogeography and range dynamics for putative southern Balkan refugia, and call for detailed local ecological studies. Highly variable densities (from zero to more than about 70 mines per 100 cm^2 of leaf area; WebTable 8) – the uppermost of which are characteristic of ornamental trees during summer across Europe, in archival herbarium collections of *C ohridella* so far examined – appear not to reflect mere botanical collecting bias. Instead, judging from the mine-infested leaves in late-summer, low-elevation specimens from Albania in 1961 (WebFigure 3), either the biological invasion started much earlier than previously thought or "outbreaks" of this moth species are a natural phenomenon. Studies on parasitoid communities and the effects of climate/elevation/season and plant defenses on moth population dynamics are therefore needed to understand how *C ohridella* irrupts and becomes invasive. The archival evidence from Albania (eg WebFigure 3) shows that outbreaks occurred decades earlier than previously thought, without certainly implicating the three recently "invasive" haplotypes, and apparently without precipitating invasion along major transport routes.

Additional molecular studies on genetics and population biology of both the moth and its host plant in the Balkans would complement ecological studies, by comparing adaptive traits of different *C ohridella* populations/haplogroups; this would provide vital data that could help in conserving relictual diversity, identifying resistant cultivars of white-flowering horse-chestnut, and seeking biological controls of one of the most invasive herbivores in Europe. With its high rate and spatial scale of invasion, *C ohridella* now poses a major threat to *Aesculus*, and possibly also *Acer* cultivars/species (Hellrigl 2001; Péré *et al.* 2010) and their native communities elsewhere, notably in North America. We highlight here the unique role of herbaria in providing genetic information from highly inaccessible or even extirpated populations. Native populations of horse-chestnut are endangered in several areas in Albania (Vangjeli *et al.* 1997). Added to the risk of potential genetic homogenization of *C ohridella* populations in remote sites now served by roads, this implies that substantial genetic diversity, not only of the moth (Valade *et al.* 2009) but also of its host plant, may be rapidly disappearing. The perspective changes radically from a highly invasive moth pest damaging a widely cultivated tree, to ancient relic populations of conserva-

tion importance. In such cases, a handful of historical herbarium specimens may represent the only remaining trace of genetic and spatiotemporal information.

Indeed, herbaria – far from fusty archival collections – have broad relevance in current research, emerging as a premium resource for documenting spatiotemporal changes in biodiversity. Tracing the invasion routes of insects through haplotypes and genotypes found in organisms preserved even inadvertently in herbarium or museum collections represents an important application for archival DNA studies (Condon and Whalen 1983; Funk 2004). Genetic sequencing of pathogens (including microorganisms) from herbarium samples – to understand the origin and epidemiology of plant mining pests and diseases or even to partially reconstruct past ecosystem conditions or interactions – is an emerging field of inquiry, especially with the advent of next-generation sequencing technology. Analysis of historical samples in herbaria can assist with investigating contemporary problems of bioinvasions and decreased biological diversity. Entomologists and ecologists have too long ignored spatiotemporal information available from historical herbarium specimens, and we urge an integrated approach in the use of such data to address questions, in particular about the origin of invasive species.

Acknowledgements

We thank C Loup for permission to extract MNHN specimens and for help; D Mabberley for permission to sample; W Milliken and R de Kok for help at RBG Kew; R Valade for dataset/field assistance; G Aymonin for sending a map; B Courtial, E Magnoux, and C Courtin for laboratory assistance; M Lüchow for help with imaging; N Saby for drawing the map; A Strid for providing *Aesculus* distribution data from the Flora Hellenica database; H Manitz and FK Meyer for information on the Meyer specimens; W Heitland for Figure 2c; N Kirichenko for an outgroup sample; and J Casas, A Roques, and H Zetsche for helpful comments. Funding was provided by STUDIUM (DCL, CLV, and SA), INRA-EFPA (DCL), SEE-ERA.NET Pilot Joint Call Western Balkan Region (Nr 06-1000031- 10627; CLV, NDA, and SA), and NSERC and Genome Canada (RR).

References

Abbott I, Wills A, and Burbidge T. 1999. Historical incidence of *Perthida* leafminer species (Lepidoptera) in southwest Western Australia based on herbarium specimens. *Aust J Ecol* **24**: 144–50.

Adamović L. 1908. Die Rosskastanie im Balkan. *Bot Jahrb Syst* **41**: Beibl 94.

Avtzis ND, Avtzis DN, Vergos SG, *et al.* 2007. A contribution to the natural distribution of *Aesculus hippocastanum* (Hippocastanaceae) in Greece. *Phyt Balcanica* **13**: 183–87.

Baldacci A. 1897. Rivista della collezione botanica fatta nel 1895 in Albania. *Nuov Giorn Bot Italiano* **4**: 386–419.

Chase MW, Soltis DE, Olmstead RG, *et al.* 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbc*L. *Ann Missouri Bot Gard* **80**: 528–80.

Condon M and Whalen MD. 1983. A plea for collection and preservation of herbivore and pathogen damaged plant materials. *Taxon* **32**: 105–07.

Delisle F, Lavoie C, Jean MA, *et al.* 2007. Reconstructing the spread of invasive plants: taking into account biases associated with herbarium specimens. *J Biogeography* **30**: 1033–42.

Deschka G and Dimić N. 1986. *Cameraria ohridella* sp n (Lep, Lithocolletidae) aus Mazedonien, Jugoslawien. *Acta Entom Jugosl* **22**: 11–22.

Funk VA. 2004. 100 uses for an herbarium (well at least 72). *ASPT Newsletter* **17**: 17–19.

Gilbert M, Guichard S, Freise J, *et al.* 2005. Forecasting *Cameraria ohridella* invasion dynamics in recently invaded countries: from validation to prediction. *J Appl Ecol* **42**: 805–13.

Grabenweger G and Grill R. 2000. On the place of origin of *Cameraria ohridella* Deschka & Dimić (Lepidoptera: Gracillariidae). *Beitr Ent* **1**: 9–17.

Grebenchikoff O. 1938. On the occurrence of *Fagus orientalis* in Greece. *Bull Misc Inf* **1938**: 38–45.

Harris AJ, Xiang Q-YJ, and Thomas DT. 2009. Phylogeny, origin, and biogeographic history of *Aesculus*. *Taxon* **58**: 108–26.

Heldreich T von. 1879. Beiträge zur Kenntniss des Vaterlandes und der geographischen Verbreitung der Rosskastanie, des Nussbaums und der Buche. *Verh Bot Vereins Prov Brandenburg* **1879**: 139–53.

Hellrigl K. 2001. Neue Erkenntnisse und Untersuchungen über die Rosskastanien-Miniermotte *Cameraria ohridella* Deschka & Dimić, 1986 (Lepidoptera, Gracillariidae). *Gredleriana* **1**: 9–81.

Lack HW. 2002. The discovery and rediscovery of the horse chestnut. *Arnoldia* **61**: 15–19.

Lees DC, Lopez-Vaamonde C, and Augustin S. 2011. Taxon page for *Cameraria ohridella* Deschka & Dimić 1986. http://eolspecies.lifedesks.org/pages/8675. Viewed 31 May 2011.

Lees DC, Rougerie R, Zeller-Lukashort HC, *et al.* 2010. DNA minibarcodes in taxonomic assignment: a morphologically unique new homoneurous moth clade from the Indian Himalayas described in *Micropterix* (Lepidoptera, Micropterigidae). *Zool Scr* **39**: 642–61.

Mädler K. 1939. Die pliozäne Flora von Frankfurt am Main. *Abh Senck Naturf Ges* **446**: 1–202.

Mari Mena N, Valade R, Magnoux E, *et al.* 2008. Microsatellites for the highly invasive *Cameraria ohridella* (Lepidoptera: Gracillariidae) and their amplification in related species. *Eur J Entomol* **105**: 945–48.

Markgraf F. 1931. Pflanzen aus Albanien 1928. *Denkschr Akad Wiss Wien, Math Naturwiss Kl* **102**: 317–60.

Péré C, Augustin S, Turlings TCJ, and Kenis M. 2010. The invasive alien leaf miner, *Cameraria ohridella*, and the native tree, *Acer pseudoplatanus*: a fatal attraction? *Agric For Entomol* **12**: 151–59.

Postigo-Mmijarra JM, Gomez-Manzaneque F, and Morla C. 2008. Survival and long-term maintenance of tertiary trees in the Iberian Peninsula during the Pleistocene: first record of *Aesculus* L (Hippocastanaceae) in Spain. *Veg Hist Archaeobot* **17**: 351–64.

Raus T. 1980. Die Vegetation Ostthessaliens (Griechenland). III. Querco-Fagetea und azonale Gehölzgesellschaften. *Bot Jahrb Syst* **101**: 17–82.

Thomsen PF, Elias S, Gilbert MTP, *et al.* 2009. Non-destructive sampling of ancient insect DNA. *PLoS ONE* **4**: e5048; doi:10.1371/journal.pone.0005048.

Valade R, Kenis M, Hernandez Lopez A, *et al.* 2009. Mitochondrial and microsatellite DNA markers reveal a Balkan origin for the highly invasive horse-chestnut leaf miner *Cameraria ohridella* (Lepidoptera, Gracillariidae). *Mol Ecol* **18**: 3458–70.

Vangjeli J, Misja K, Gjiknuri L, *et al.* 1997. Libri i Kuq (Bimë, shoqërime bimore dhe kafshë të rrezikuara). Tirana, Albania: The Regional Environmental Center for Central and Eastern Europe.

WebPanel 1. Materials and methods

Samples

All horse-chestnut specimens that were examined representing natural populations, and *Cameraria ohridella* specimens that were excised from archival leaf mines along with their main collection attributes, are documented in WebTable 1. Archival leaf mines found in herbaria are documented in more detail in WebTable 8. Larvae from modern populations extracted for comparison with archival ones (from localities used for microsatellite work in Valade *et al.* 2009) are also itemized in WebTable 1. We visited the herbaria at the Royal Botanical Gardens (RBG), Kew, UK, the Muséum National d'Histoire Naturelle (MNHN), Paris, France, Natural History Museum, London, UK, and Botanischer Garten und Botanisches Museum Berlin-Dahlem (BGBM), Berlin, Germany, because these herbaria were particularly likely or known to have historical collections of horse-chestnut from the Balkans, and we asked for high-resolution scans to be made of collections at Vienna, Austria, and Jena, Germany. We personally examined a total of 61 sheets of *Aesculus hippocastanum*, representing 18 localities (Figure 1) from potential natural populations (including some distributed duplicates) as well as many specimens from planted individuals (in parks, etc) held at these herbaria. Excised specimens or lysed remnants obtained by DNA extraction from archival herbarium specimens are archived in the collection of the National Institute for Agricultural Research (INRA), Orléans, France (WebTable 1). Larvae extracted from herbarium samples that were processed at Guelph, Canada (sample IDs given in WebTable 1) are also documented in the project COARC of the Barcode of Life Datasystems (BOLD: www.boldsystems.org/views/login.php).

Mine density estimation

To see if mine densities were comparable to those known for modern populations of *C ohridella* and especially for the haplotypes undergoing rapid range expansion, we calculated the total examined leaf surface area by polygon in the program ImageJ 1.42q (http://rsb.Info.nih.gov./Ij), after calibrating to a 20-cm ruler included in the photograph. Except for Jena and Vienna, where we used high-resolution scans, we used a compound microscope to count all mines, because early stages are easy to miss when searching with the naked eye, and also to distinguish them from spots such as those made by the fungal pathogen *Guignardia aesculi.* Data are given in WebTable 8.

Genomic DNA extraction of larvae and quality control

So as to minimize damage to herbarium specimens, we excised a disk of the leaf upper epidermis around the perimeter of the mine, using a fine-bladed sterile scalpel, folded this back (as in Figure 2b), and transferred any larva, pupa, or parasitoid found therein to a small vial containing absolute alcohol using ultraviolet-sterilized forceps that were also carefully cleaned between each handling. All larval extractions were done in the respective herbaria under a microscope after photography of mines and larvae in situ.

To prevent contamination, we attempted extraction and conducted polymerase chain reaction (PCR) analyses in two separate laboratories, working on two separate sets of larvae from a range of herbaria, rather than dividing the larvae individually. Seven archival samples of larval *C ohridella* (CO1/CO3/CO12/CO16/CO21/CO25/CO34; WebTable 1) were processed at the Canadian Centre for DNA Barcoding (CCDB), while aliquots from 54 specimens (WebTable 1, indicated in bold) – along later with aliquots of the CCDB-processed specimens – were processed at INRA in France. Other CO numbers listed in WebTable 1 were not extracted, given that they represented exuviae or very small larval specimens, even though they were *C ohridella*. In addition, six "fresh" larvae from Karitsa (CCDB-02228-H06-H11) – from pairs of trees at three localities about 0.5 km apart, whose positions approximated closely (within a few kilometers, according to details in Grebenchikoff 1938 and Grebenščikov 1954; see also Figure 1) to the collection site of mine CO40 – were processed at the CCDB, targeting the full DNA barcode region of the mitochondrial DNA gene COI. DNA extracts of 29 additional samples from Karitsa that were used for microsatellite analysis in Valade *et al.* (2009) and 26 from Monodendri (Figure 1) were subsequently amplified for the full DNA barcode at INRA. We used negative controls (ultra-pure water) at all stages of PCR and, as a positive control, we ran a fresh genomic DNA extract for *Phyllonorycter populifoliella* (Treitschke, 1833), which was known to be significantly different from *C ohridella* for all examined regions (five small fragments) of the DNA barcodes and also amplified for microsatellites Ohrid2782 and Ohrid2814.

Tissue samples of fresh larvae were shipped to CCDB in wells of a micro-plate, and processed using the manual silica-based 96-well extraction protocol described in Ivanova *et al.* (2006), with a final elution volume of 30 µL. The extraction method is non-destructive and specimens were recovered after the lysis step. Entire archival samples (a larva, pupa, or exuvium) were placed individually into sterile tubes containing 100% ethanol and shipped to CCDB while genomic DNA extracted from other samples was processed at INRA using the NucleoSpin Tissue Kit (Macherley-Nagel, Hoerd, France) following the manufacturer's instructions, except that two elution steps using 200 µL of ultra-pure water were used, with 1-min incubation on the bench, then combined and spun dry in a Speedvac. Genomic DNA was solubilized in 10 µL of double purified water and 1 µL of all extracted genomic DNA then analyzed using a Nanodrop (Thermo-scientific, Wilmington, DE). Twelve of 13 samples first extracted at INRA showed a fairly flat profile with a 260/280 ratio below 1.60, while CO40 showed a curve and a 260/280 ratio ~ 1.91 indicative of reasonable DNA quality (the positive control measured 2.1). The Tepelenë and Tomor samples later extracted also had good quality profiles with 260/280 ratios ranging from 1.01–1.67 (or, in a single case, 2.28). An appropriate dilution was carried out to arrive at a final concentration of around 12.5 ng/µL.

PCR amplification and sequencing of the COI barcode region

Primers

For archival specimens, regular barcoding PCR amplification protocols targeting fragments of ca 300–650 bp are unlikely to yield

WebPanel 1. Materials and methods – *continued*

amplicons (Zimmermann *et al.* 2008). Thus, we attempted to amplify shorter fragments ("mini-barcodes": Hajibabaei *et al.* 2006a, b; Meusnier *et al.* 2008) of 90–140 bp. At Guelph, we used six different primer sets, which can successfully assemble full-length DNA barcodes (658 bp) for archival specimens, that were developed for sphingid moths (Lees *et al.* 2010; Rougerie, Meusnier *et al.* unpublished). At INRA, we used a combination of universal (LCO/HCO; Uni-minibar) and specially designed primers (LepF1/CamR1a, CamF2a/CamR2a, CamF3a/ CamR3a, CamF5a/ CamR5a, and CamF6a/LepR1; WebTable 2) using Oligo 3 (Molecular Biology Insights Inc), to cover informative regions of the *C ohridella* barcode dataset (Valade *et al.* 2009), guided by a consensus file made from all sequences (Figure 3). At Guelph, a first attempt was made to get the full barcode with LepF1/LepR1 and, for failures, and a second attempt was made with LepF1/MLepR1 and MLepF1/LepR1 (Hajibabaei *et al.* 2006a) targeting fragments of 307–407 bp. We then used the mini-barcode primers developed for Sphingidae (which sometimes work in other families; eg Hausmann *et al.* 2009; Lees *et al.* 2010).

PCR and sequencing

At INRA, PCR reactions were carried out in a 25 μL reaction volume containing 2.5 μL 10X buffer (10 mM TrisHCl [pH 8.3], 50 mM KCl, and 0.01% gelatine), 1 mM of each deoxyribonucleotide triphosphate (dNTP), 2 mM $MgCl_2$, 0.4 μM of forward and reverse primer, 0.5 μg betaine, 1 U of REDTaq® Genomic DNA Polymerase (Sigma-Aldrich) and between 12.5 ng (for sample CO40) and 25 ng (for other samples of genomic DNA). For a second batch of samples (Karitsa, 1974; 1981; Tsepélevo, 1964; all 1961 samples from Albania; WebTable 1), we substituted Platinum Taq (Invitrogen, Burlington, Canada) for REDTaq for higher performance. At CCDB, PCR reactions were carried out in 12.5 μL reaction volumes containing: 2.5 mM $MgCl_2$, 1.25 pM of each primer, 50 μM dNTPs, 10 mM TrisHCl (pH 8.3), 50 mM KCl, 10–20 ng (1–2 μL) of genomic DNA, and 0.3 U of Taq DNA polymerase (Platinum Taq DNA polymerase; Invitrogen).

At CCDB, the thermocycling profile consisted of one initial denaturation step of 1 min at 94°C, followed by five cycles of 40 s at 94°C, 40 s at 45°C, and 1 min at 72°C, followed by 35 cycles of 40 s at 94°C, 40 s at 51°C, and 1 min at 72°C, with a final extension of 5 min at 72°C. For mini-barcode fragments, we used the following touch-up profile: a hot start for 2 min at 94°C, followed by denaturation (40 s at 94°C), annealing for 1 min at 46°C, extension for 30 s at 72°C, the last three steps cycled five times, then denaturation for 40 s at 94°C, annealing for 1 min at 53°C, extension for 30 s at 72°C, the last three steps cycled 35 times, followed by a final extension for 30 s at 72°C. At INRA, a similar protocol was used on a 2720 Thermal Cycler (Applied Biosystems, Foster City, CA), except that the two annealing steps were set at 1 min at 48°C and 1 min at 51°C following optimization for the primer pairs used (WebTable 2).

PCR products were cleaned and visualized on a 2% agarose gel after 1 hour of ethidium bromide immersion (INRA) or E-Gel 96-well system (Invitrogen) at CCDB. Unpurified samples revealing faint to strong bands were cycle sequenced bidirectionally or unidirectionally (with the same primers used for the PCR reactions) in 10 μL reaction volumes containing: (at Guelph) 0.25 μL of BigDye v3.1, 1.875 μL of 5X ABI sequencing buffer, 5 μL of 10% trehalose, 1 μL of 10 μM primer, 0.875 μL of ultra-pure water, and 1 μL of PCR product, or (at INRA) in a 20 μL reaction volume containing 10 μL of PCR product and 2 μL of BigDye v3.1, 3 μL of 5X ABI sequencing buffer, 1 μL of 10 μM primer, and 4 μL of ultra-pure water. The following thermocycling profile was used for all products: initial denaturation at 96°C for 2 min, followed by 30 cycles of 96°C for 30 s, annealing at 55°C for 15 s, and extension at 60°C for 4 min. Sequence reads were generated (at CCDB) on a ATBI 3730xl DNA Analyser or (at INRA) on a ATBI 3000 DNA Analyser (Applied Biosystems), after clean-up at CCDB with Sephadex (Sigma-Aldrich, Oakville, Canada) or at INRA with the NucleoSpin Extraction Kit (Macherley-Nagel, Hoerd, France) for old samples or Sephadex G-50 (GE Healthcare, US, formerly Amersham Biosciences) for modern samples.

We obtained archival COI DNA data from larvae and compared their haplotype with that of modern samples (Figure 3). Three of the DNA fragments from CO40 were sequenced unambiguously for both strands and two in one direction only and the sequences were assembled in BioEdit 7.01 (Hall 1999). Other amplicons (for CO12, CO53, CO58, CO62-CO63, CO70, CO93, and CO100; Figure 3) were sequenced in the reverse direction only for one or more of primer pairs LepF1/CamR1a, Cam3aF/Cam3aR, and Cam6aF/LepF1. Editing was done using CodonCode Aligner v 3.0.1 (CodonCode Co 2009). The alignment using Bioedit was straightforward and non-ambiguous, and we converted to the correct reading frame using a corrected Lepidopteran codon table (as in Linares *et al.* 2009). For CO40, with one exception, we were able to verify all variable sites in both directions on the trace files. The sequences were straightforward to align by eye. We also used Bioedit in conservation plot mode to view dot plots against a reference sequence (eg Figure 3). Pairwise distances were calculated by K2P in Mega v 4 (Tamura *et al.* 2007), with pairwise deletion option for missing data, on sequences trimmed to 630 bp to minimize any effect of missing data.

For amplification of microsatellites, we used 1–2 μL (20–40 ng of genomic DNA). The total reaction volume was otherwise 10 μL, containing 20 ng of genomic DNA, 0.4 units of Taq DNA polymerase (Sigma), 1X buffer (100 mM TrisHCl, 500 mM KCl, and 0.1% gelatine), 2.5 mM $MgCl_2$, 20 mg/L of BSA, 250 μM of each dNTP, and 0.4 μM of each primer. Forward primers were 5'-labelled with either 6-FAM, PET, HEX (Sigma), or NED (Applied Biosystems) fluorescent dyes, using exactly the same labels and sequencing machine as Valade *et al.* (2009) to avoid dye-slip differences between the two studies. PCR conditions were 3 min at 95°C followed by 35 cycles of 50 s at 95°C, 1 min at 52°C (for primers 2753, 2759, and 2762) and at 50.6°C (for primers pair Ohrid2814), 45 s at 72°C, and 15 min at 72°C. The amplified products were detected on an ABI-3100 automated sequencer and their sizes estimated and compared to modern samples from the Karitsa and other sites from the Valade *et al.* (2009) study with

WebPanel 1. Materials and methods – *continued*

consistent rounding, using GeneMapper v 3.7 (Applied Biosystems). Observed and expected heterozygosities and pairwise population differentiation were calculated using ARLEQUIN 3.1 (Excoffier *et al.* 2005).

Using the software Structure (v2.3; Pritchard *et al.* 2000), we explored the distribution of nuclear genetic variation (microsatellites) by estimating the number of clusters represented by the three sample locations (two ancient from herbaria specimens and one recent from field collections). This approach uses a Bayesian, Monte Carlo Markov chain (MCMC) approach to cluster individuals into groups while minimizing Hardy-Weinberg disequilibrium and gametic phase disequilibrium between loci within groups. The optimal number of populations (K) represented by the data was calculated by comparing the estimated log probability of the data for different values of K (Pritchard *et al.* 2000). We ran two independent runs with K values of 1 and 2, a burn-in period of 40 000 MCMC iterations, and a data collection period of 1 million MCMC iterations. The independent runs produced consistent results for the same value of K.

WebPanel 2. Contributions

DCL was mainly responsible for designing and writing this study, drafting figures, and obtaining and counting early stages from herbarium specimens (this part jointly with SA), conducting analysis, and submitting the manuscript. SA redrafted Figure 3. CLV and SA secured funding for and jointly steered this project and share joint senior authorship. HWL provided access to herbarium samples at Berlin and contributed substantially to the botanical research and access to various herbaria, including private collections. AHL amplified and analyzed microsatellites and drafted that part of the study. RR developed the barcoding project at Guelph and wrote part of the methodology. All the above authors contributed substantially to revisions of the manuscript. NA helped critically in the field to provide the modern samples from remote sites, and TR provided the archival Karitsa samples.

WebReferences

Excoffier L, Laval G, and Schneider S. 2005. ARLEQUIN version 3.0: an integrated software package for population genetics data analysis. *Evol Bioinform* **1**: 47–50.

Folmer O, Black M, Hoeh W, *et al.* 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotech* **3**: 294–99.

Grebenščikov [Grebenchikoff] O. 1954. O pôvodných (prirodzených) lokalitách a stanovištiach pagaštana konského (*Aesculus hippocastanum* L) v Európe. *Biol (Bratislava)* **9**: 516–30.

Hajibabaei M, Janzen DH, Burns JM, *et al.* 2006a. DNA barcodes distinguish species of tropical Lepidoptera. *P Natl Acad Sci USA* **103**: 968–71.

Hajibabaei M, Smith A, Janzen DH, *et al.* 2006b. A minimalist barcode can identify a specimen whose DNA is degraded. *Mol Ecol Notes* **6**: 959–64.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acid S* **41**: 95–98.

Hausmann A, Hebert PDN, Mitchell A, *et al.* 2009. Revision of the Australian *Oenochroma vinaria* Guenee, 1858 species-complex (Lepidoptera: Geometridae, Oenochrominae): DNA barcoding reveals cryptic diversity and assesses status of type specimen without dissection. *Zootaxa* **2239**: 1–21.

Ivanova NV, deWaard JR, and Hebert PDN. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Mol Ecol Notes* **6**: 998–1002.

Lees DC, Rougerie R, Zeller-Lukashort HC, *et al.* 2010. DNA minibarcodes in taxonomic assignment: a morphologically unique new homoneurous moth clade from the Indian Himalayas described in *Micropterix* (Lepidoptera, Micropterigidae). *Zool Scr* **39**: 642–61.

Linares MC, Soto-Calderón ID, Lees DC, *et al.* 2009. High mitochondrial diversity in geographically widespread butterflies of Madagascar: a test of the DNA barcoding approach. *Mol Phylogenet Evol* **50**: 485–95.

Meusnier I, Singer GAC, Landry J-F, *et al.* 2008. A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics* **9**: 214; doi:10.1186/1471-2164-9-214.

Pritchard JK, Stephens M, and Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–59.

Tamura K, Dudley J, Nei M, and Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**: 1596–99.

Zimmermann J, Hajibabaei M, Blackburn DC, *et al.* 2008. DNA damage in preserved specimens and tissue samples: a molecular assessment. *Front Zool* **5**: 8; doi:10.1186/1742-9994-5-18.

Bibliography

(supplementary references supporting main text and for further reading)

Botanical origins:

Barnes P and Hoda P. 2001. Plant exploration in Albania. *Curtis's Bot Mag* **18**: 170–80.

Bergmeier E. 1990. Wälder und Gebüsche des Niederen Olymp (Káto Olimbos, NO-Thessalien). *Phytocoenologia* **18**: 161–342.

Lack HW. 2000. Lilac and horse-chestnut: discovery and rediscovery. *Curtis's Bot Mag* **17**: 109–41.

Markgraf F. 1932. Pflanzengeographie von Albanien. Ihre Bedeutung für Vegetation und Flora der Mittelmeerländer. *Biblioth Botanica* **5**: 1–132.

Pschorn-Walcher H. 1997. Zur Biologie und Populationsentwicklung der eingeschleppten Roßkastanien-Miniermotte, *Cameraria ohridella*. *Forstschutz-Aktuell, Wien* **21**: 7–10.

Xiang Q-Y, Crawford DJ, Wolfe A, *et al.* 1998. Origin and biogeography of *Aesculus* L (Hippocastanaceae): a molecular phylogenetic perspective. *Evolution* **52**: 988–97.

Use for phylogenetics:

Li W, Song Q, Brlansky RH, *et al.* 2007. Genetic diversity of citrus bacterial canker pathogens preserved in herbarium specimens. *P Natl Acad Sci USA* **104**: 18427–32.

Roy MS, Girman DJ, Taylor AC, *et al.* 1994. The use of museum specimens to reconstruct the genetic variability and relationships of extinct populations. *Experientia* **50**: 5517.

Savolainen V, Cuénoud P, Spichiger R, *et al.* 1995. The use of herbarium specimens in DNA phylogenetics: evaluation and improvement. *Plant Syst Evol* **197**: 87–98.

Use for plant invasion biology:

Cozzolino S, Cafasso D, Pellegrino G, *et al.* 2007. Genetic variation in time and space: the use of herbarium specimens to reconstruct patterns of genetic variation in the endangered orchid *Anacamptis palustris*. *Conserv Genetics* **8**: 629–39.

Provan J, Booth D, Todd NP, *et al.* 2007. Tracking biological invasions in space and time: elucidating the invasive history of the green alga *Codium fragile* using old DNA. *Divers Distrib* **14**: 343–54.

Saltonstall K. 2002. Cryptic invasion by a non-native genotype of the common reed, *Phragmites australis*, into North America. *P Natl Acad Sci USA* **99**: 2445–49.

Use of herbaria for herbivore–plant interactions:

Alario SM and Alamá JAG. 2010. *Mariposas diurnas de la communitat. Valenciana (Papilionoidea & Hesperioidea) [p271, Aricia morronensis]*. Museu Valencià d'Història Natural, Fundación Entomológica Torres Sala.

Bachelard P and Chabrol L. 2010. Témoignages de la présence de *Phengaris alcon alcon* (Lep. Lycaenidae) dans les herbiers anciens. *Oreina* **Avril 2010** (9): 4–7.

Berenbaum MR and Zangerl AR. 1998. Chemical phenotype matching between a plant and its insect herbivore. *P Natl Acad Sci USA* **95**: 13743–48.

Booister M, Eurlings M, Gravendeel B, *et al.* 2009. Microsatellite analysis of *Maculinea alcon* (Lepidoptera, Lycaenidae) eggs and legs using a new method for DNA extraction from tiny amounts of tissue (poster). Abstracts 16th European Congress of Lepidopterology, Cluj-Napoca. 25–31 May 2009: 34. http://science.naturalis.nl/media/176708/booisterposter_final.pdf. Viewed 6 Jun 2010.

Lees DC and Smith NG. 1991. Foodplant associations of the Uraniinae (Uraniidae) and their systematic, evolutionary and ecological significance. *J Lepid Soc* **45**: 296–347.

Morrow PA and Fox LR. 1989. Estimates of pre-settlement insect damage in Australian and North American forests. *Ecology* **70**: 1055–60.

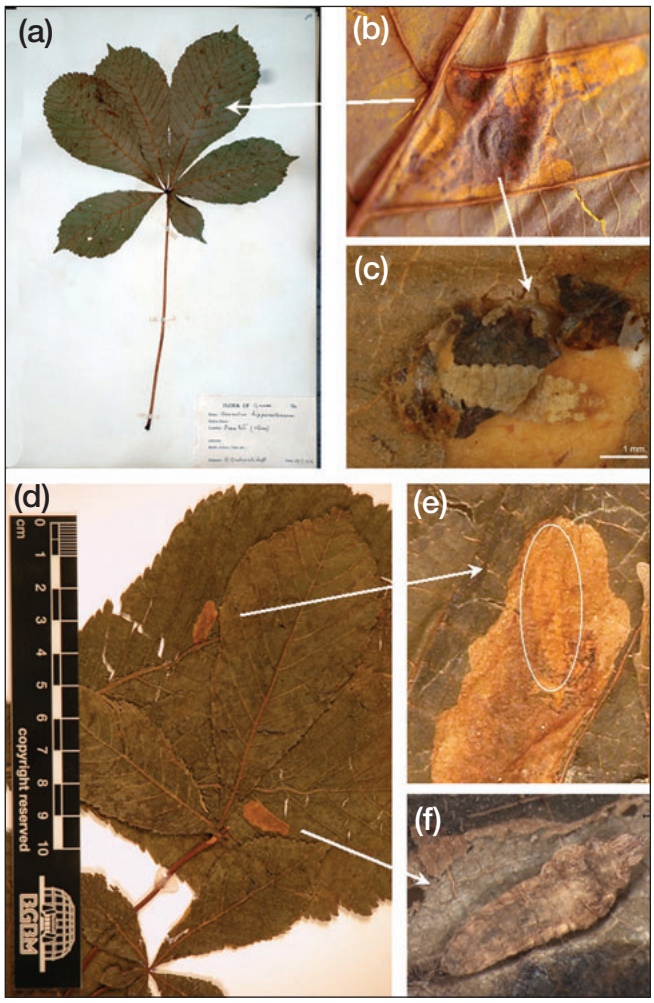
van Nieukerken E and Koster S. 1999. De valkruidmineervlinder *Digitivalva arnicella* in nederland: herontdekking en behoud (Lepidoptera: Plutellidae: Acrolepiinae). *Ned Faun Med* **9**: 15–28.

van Nieukerken EJ and de Berg C van. 2003. A new *Stigmella* feeding on Urticaceae from Guam: first records of Nepticulidae (Lepidoptera) from Micronesia and Polynesia. *Invertebr Syst* **17**: 27–37.

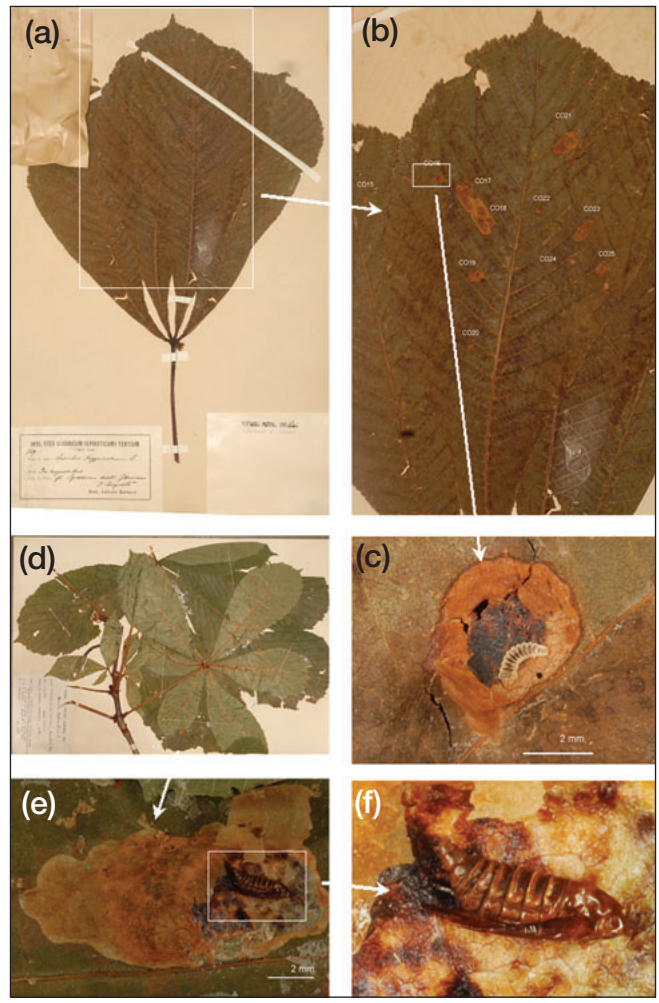
Ancient DNA studies:

Gilbert M, Bandelt H, Hofreiter M, *et al.* 2005. Assessing ancient DNA studies. *Trends Ecol Evol* **20**: 541–44.

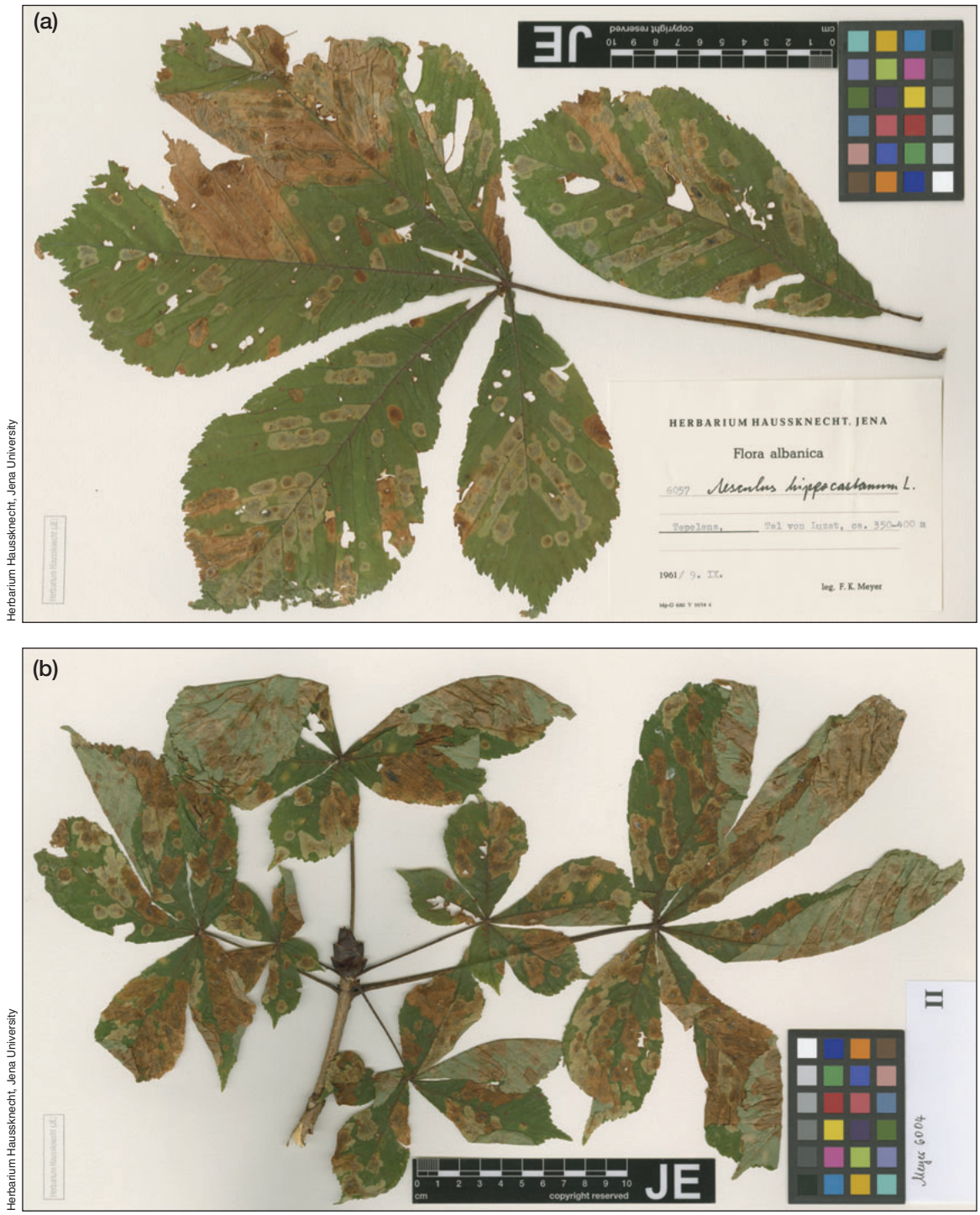
Pääbo S, Poinar H, Serre D, *et al.* 2004. Genetic analyses from ancient DNA. *Ann Rev Genetics* **38**: 645–79.



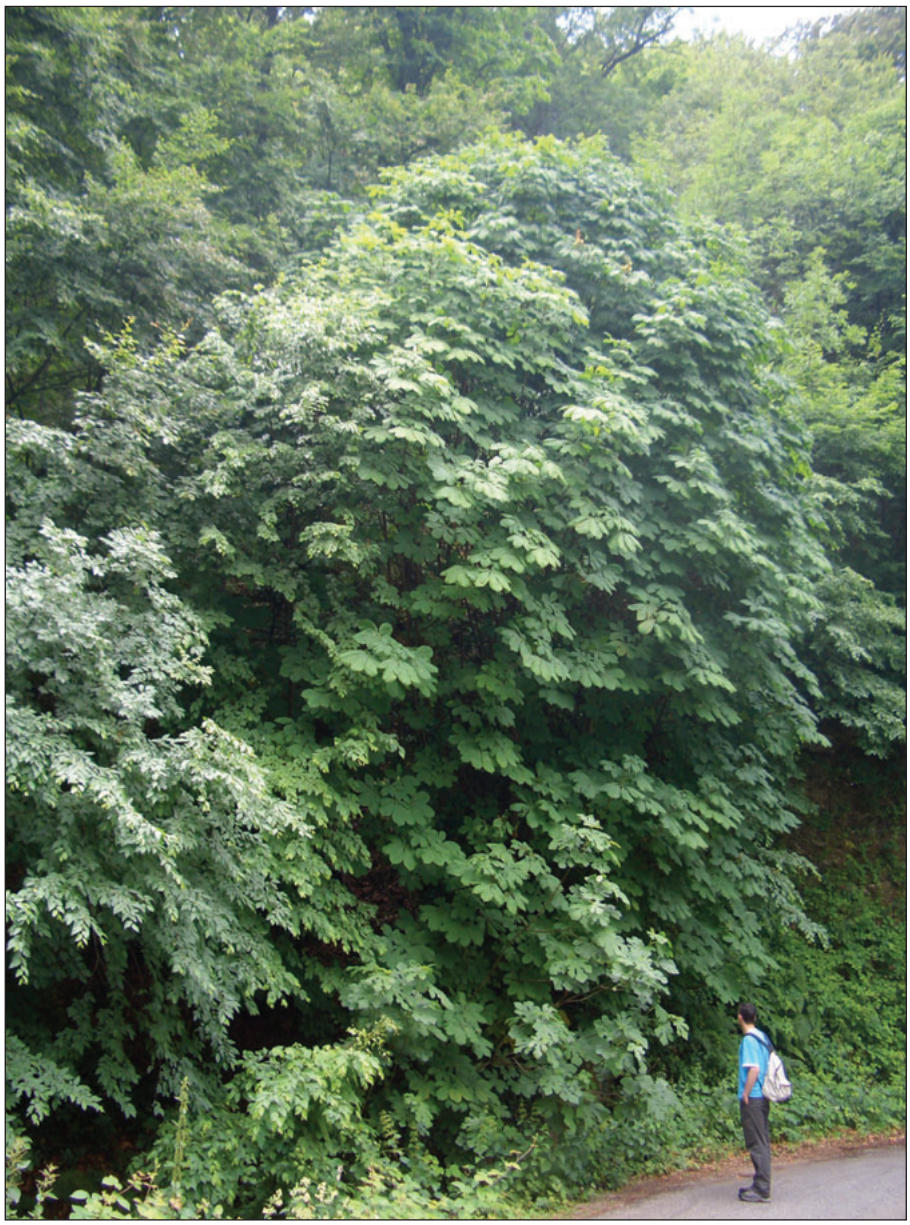
WebFigure 1. *Archival herbarium specimens of* Aesculus hippocastanum *from the southern Balkans containing mines and larval voucher specimens of* Cameraria ohridella. *(a) Herbarium specimen Grebenchikoff – collected at Karitsa, eastern Greece, on 29 Jul 1936 – from Kew, showing (b) a mine and (c) extracted spinning stage larva that was successfully sequenced, CO40; scale bar, 1 mm. (d) Herbarium specimen Markgraf-1513 – collected at Tresova, southeast Albania, on 28 Jun 1928 – from Berlin, showing two mines: (e) mine with well-pressed larva (L4 CO1) visible inside, as shown in detail in Figure 1d, and (f) mine of CO2, showing spinning stage larva inside.*



WebFigure 2. *Herbarium specimens of* Aesculus hippocastanum *containing* Cameraria ohridella *mines at Paris and Kew. (a) Herbarium specimen Baldacci-129 at Paris, collected at Syrakou, Ioannina, central Greece, on 3 Aug 1895. (b) Eleven mines in Baldacci-129 that were hidden by the preparator by the topmost leaflet (note: botanists prefer to display "perfect" specimens of leaves). (c) Specimen CO16 from the same leaflet showing L3 stage larva; scale bar, 2 mm. (d) Herbarium specimen Alston and Sandwith-2345 at Kew, collected at Çajup[i]-Zhej, southern Albania, on 8 Aug 1935. (e) Specimen CO43, a pupa of* C ohridella *(enlarged in [f]), from the herbarium specimen in (d); scale bar, 2 mm.*



WebFigure 3. *Representative unmounted herbarium sheets of* Aesculus hippocastanum *at Jena herbarium from Tepelenë and Sevaster, Albania, showing heavy late-season attack ("outbreaks") of* Cameraria ohridella*, 28 years before the start of the previously documented biological invasion of the species and 23 years before the first recorded outbreak. (a) Herbarium specimen Meyer-6057, labelled "Tepelena [Tepelenë], Tal von Luzat, 350–400 m, 1961/9/IX", leg FK Meyer (probably natural site; specimen containing at least 126 mines). (b) Herbarium specimen Meyer-6004, companion herbarium sheet labelled "Griba, Tälchen südlich Sevaster an der Strasse [Griba, small valley south of Sevaster on the road], ca 700 m. 1961/8. IX, leg FK Meyer (probably natural site in small valley; specimen containing at least 306 mines).*



WebFigure 4. *Probably natural* Aesculus hippocastanum *tree at Karitsa, next to a road, June 2008, examined by R Valade, from which* Cameria ohridella *larvae were collected. Archival herbarium collections by Grebenchikoff from this area in 1936 and by Raus and Royl in 1981 reveal the presence of* C ohridella *haplotypes ("D" and "AB", respectively) different from the predominant "A" (and, at low frequency, "O" and "S") that we found at three Karitsa sites (six trees) in 2008. We could not rule out haplotype "A" in the case of the short 1974 archival sequence. Modern roads were only built into Karitsa in the mid-1970s, suggesting that they have allowed haplotype "A" and possibly "S" to invade or largely wipe out the original population of the area.*

WebTable 1. Herbarium specimens of *Aesculus hippocastanum* examined (see also Baldacci 1897; Heldreich 1879; Grebenchikoff 1938; Grebenščvikov 1954; Markgraf 1931 [listing Markgraf #1513]), Raus 1980 [listing Raus 2567], and details of the six trees from which modern Karitsa samples were derived, together with sample and BOLD/GenBank references for all early stages of *Cameraria ohridella* excised or sequenced from leaf mines.

Sample source (KEW; BER= Berlin; PAR= MNHN Paris; WIE = Vienna; JEN= Jena; NEW= 2008)	Date of collection	Country (AL=Albania; GR=Greece; BG= Bulgaria)	Locality	Est latitude (decimal)	Est longitude (decimal)	Est Geog error (km)	Est elevation (m)	Elevation error (m)	Natural/ornamental population	Closest known site (in Avtzis et *al.* 2007, Valade et *al.* 2009)	Distance to closest known site (in Avtzis et *al.* 2007, Valade et *al.* 2009)	# herbarium sheets	# larvae/pupae C *ohridella* excised	Sample reference number	Institutional voucher number (INRA-URZF-), BOLD accession number (CCDB-; CO-ARC-) GenBank Accession number (JF-, HM-)
KEW - T.Heldreich	11/08/1879	GR	Mikro Chorio, Mt. Chelidoni	38.829	21.705	1	1294	380	N	Pano-Kalesmenou	10 km SSW	2	4	CO37-CO39: CO46:	INRA-URZF-6466-INRA-URZF-6468; INRA-URZF-6469
WIE - T.Heldreich-W 0023011; W 1889-0315748	11/08/1879	GR	Mikro Chorio, Mt. Chelidoni	38.829	21.705	1	1294	380	N	Pano-Kalesmenou	10 km SSW	2	?	[Has mines; specimen not closely examined]	
PAR - T.Heldreich	11/08/1879	GR	Mikro Chorio, Mt. Chelidoni	38.829	21.705	1	1294	380	N	Pano-Kalesmenou	10 km SSW	2	2	CO8-CO9:	INRA-URZF-6455-NRA-URZF-6456
PAR - T.Heldreich	11/08/1879	GR	Mikro Chorio, Mt. Chelidoni	38.829	21.705	1	1294	380	N	Pano-Kalesmenou	10 km SSW	2	1	CO34:	COARC006-09, INRA-URZF-6465
PAR - T.Heldreich	24/06/1885	GR	Agrafa, Korona	39.13	21.612	2.5	1100	50	N	Marathos	5.4 km N	2	0		
KEW - C.Haussknecht	07/1885	GR	inter Chaliki et Krania*	39.63	21.24	10	1000	-	N	Kalliroi-Koukouffi	5 km NW	1	0		
KEW - A.Baldacci-129	03/08/1895	GR	Syrakou, dist. Janina [Ioannina]	9.594	21.104	2.5	1100	150	N	Kalliroi-Koukouffi	17 km WNW	1	0		
PAR - A.Baldacci-129	03/08/1895	GR	Syrakou, dist. Janina	39.594	21.104	2.5	1100	150	N	Kalliroi-Koukouffi	17 km WNW	4	6	CO16: CO17: CO21: CO22: CO24: CO25:	COARC005-09, INRA-URZF-6459 INRA-URZF-6460 COARC003-09, INRA-URZF-6461 INRA-URZF-6462 INRA-URZF-6463 COARC004-09, INRA-URZF-6464
KEW-P.Sintenis-692	18/06/1896	GR	Chaliki*: / in subalp.Turnara	39.682	21.189	-	1150	-	N	Malakasi	10 km SW	1	0		
PAR-P.Sintenis-692, 693	18/06/1896	GR	Chaliki*: in valle Negerli/ in subalp.Turnara	39.682	21.189	-	1150	-	N	Malakasi	10 km SW	3	0		
KEW-J.Mattfeld-2562	01/08/1926	GR	Megalochorio, Kaliakuda	38.8146	21.7532	0.25	1050	50	N	Pano-Kalesmenou	10.5 km SSE	3	0		
WIE-J.Mattfeld	01/08/1926	GR	Megalochorio, Kaliakuda	38.8146	21.7532	0.25	1050	50	N	Pano-Kalesmenou	10.5 km SSE	1	0		
PAR-J.Mattfeld	01/08/1926	GR	Megalochorio, Kaliakuda	38.8146	21.7532	0.25	1050	50	N	Pano-Kalesmenou	10.5 km SSE	1	0		
KEW-N. Stojanoff	11/05/1928	BG	Preslavsk[a] Balkan	43.1521	26.7542	7	300	150	?	Dervishka	0 km.	1	0		
BER-Markgraf-1513	28/06/1928	AL	Devollbridge, Schlucht [gorge] near Tresova, Albania	40.7342	20.5786	25	750	100	N	Rozhan/ Korçe (A,D)	within 33 km	1	3	CO1: CO2: CO3:	COARC001-09, INRA-URZF-6452; INRA-URZF-6453 COARC002-09, INRA-URZF-6454
KEW-A.H.G.Sandwith and N.Y.Alston-2365	08/08/1935	AL	Çajup[i}-Zhej, Lunxheriës range	40.2072	20.1994	0.5	975	50	N	Nemercke (Gjirokaster)	within 15 km	1	2	CO43: CO44:	INRA-URZF-6472 INRA-URZF-6473
KEW-O.Grebenchikoff	29/07/1936	GR	Karit[s]a, Ossa Mt. [Fig. 1 inset: "G"]	39.826	22.743	0.3	1000	50	N	Karitsa	ca. 1-2 km.	1	2	CO40: CO41	JF746742, INRA-URZF-6470 INRA-URZF-6471
KEW-K.H.Rechinger-21509	17/07/1958	GR	Montes Timphi, *supra pagum* [above village] Skamneli	39.9138	20.8512	1	1200	-	N	Monodendri-Petrino Davos	10 km W	1	-		
BER-K.H.Rechinger 21509	17/07/1958	GR	Montes Timphi, *supra pagum*	39.9138	20.8512	1	1200	-	N	Monodendri-Petrino Davos	10 km W	1	-		

continued

WebTable 1. – ***continued***

Sample source (KEW; BER= Berlin; PAR= MNHN Paris; WIE = Vienna; JEN= Jena; NEW= 2008)	Date of collection	Country (AL=Albania; GR=Greece; BG= Bulgaria)	Locality	Est latitude (decimal)	Est longitude (decimal)	Est Geog error (km)	Est elevation (m)	Elevation error (m)	Natural/ornamental population	Closest known site (in Avtzis et *al.* 2007, Valade et *al.* 2009)	Distance to closest known site (in Avtzis et *al.* 2007, Valade et *al.* 2009)	# herbarium sheets	# larvae/pupae *C ohridella* excised	Sample reference number	Institutional voucher number (INRA-URZF-), BOLD accession number (CCDB-; CO-ARC-) GenBank Accession number (JF-, HM-)
JEN- Meyer 3316	25/06/1959	AL	Mali i Gjer, Delvina-Palohori,	39.987	20.075	4.5	400	50	N	Tomor	31 km	3	0		
JEN- Meyer 5430	04/05/1960	AL	Devoll Bridge, Bei [near] Tresova	40.7342	20.5786	3.5	700	100	N	Rozhan/ Korçe (A,D)	within 33 km	5	0		
JEN- Meyer 6004	08/09/1961	AL	Griba, Südlich Sevaster	40.3818	19.7282	0.5	700	75	?	Tepelen[ë]	29 km	4	3	CO95-CO97:	INRA-URZF-6509- INRA-URZF-6511
JEN- Meyer 6057	09/09/1961	AL	Tepelen[ë], Tal von Luzat	40.2585	20.0378	3.4	375	25	N	Stravaj	88 km	4	18	CO56-CO67: CO58: CO62: CO63: CO69-CO72: CO70: CO80-CO82:	INRA-URZF-6478- INRA-URZF-6489 JF746746 JF746747 INRA-URZF-6490 INRA-URZF-6493 JF746748 INRA-URZF-6494- INRA-URZF-6496
JEN- Meyer 6380	22/09/1961	AL	[Mt] Tomor, Sotir[ë]	40.7552	20.1663	8.5	500	50	N	Stravaj	34 km	6	11	CO83-CO93: CO85: CO93:	INRA-URZF-6497- INRA-URZF-6507 JF746749 JF746750
PAR - G.G.Aymonin	23-26/06/1964	GR	Tsepélovo	39.911	20.821	0.5	1100	50	N	Monodendri-Petrino Davos	6.4 km WSW	1	1	CO12: CO13 [parasitoid]	COARC007-09; INRA-URZF-6457 INRA-URZF-6458
BER/Raus- T. Raus 2567	05/07/1974	GR	Kechriá, between Karitsa and Séloma Karitsa [Fig. 1 inset:"4"]	39.8380	22.7466	0.1	740	50	N	Centroid of samples below	within 1.5 km	2	4	CO51: CO52: CO53: CO54:	INRA-URZF-6474 INRA-URZF-6475 JF746743, INRA-URZF-6476 INRA-URZF-6477
BER –T. Raus & Royl 5356	12/09/1981	GR	Karitsa [Fig. 1 inset:"5"]	39.8397	22.7381	0.1	800	50	N	Centroid of samples below	within 2.2 km.	2	1	CO100:	JF746744, INRA-URZF-6512
NEW- C. Lopez-Vaamonde and S.Augustin	06/06/2008	GR	Karitsa, forest in canyon [Fig. 1 inset:"1"]	39.849	22.722	0.15	625	25	N				2	T43,T45	CCDB-02228-H06 CCDB-02228-H06- CCDB-02228-H07, HM379297-HM379298
NEW- C. Lopez-Vaamonde and S.Augustin	10/06/2008	GR	Karitsa, in forest [Fig. 1 inset:"3"]	39.85	22.727	0.15	550	25	N				2	T58,T59	CCDB-02228-H08- CCDB-02228-H09, HM379299-HM379300
NEW- C. Lopez-Vaamonde and S.Augustin	10/06/2008	GR	Karitsa, by road [Fig. 1 inset:"2"]	39.844	22.725	0.15	725	25	N				2	T60,T61	CCDB-02228-H10- CCDB-02228-H11, HM379301-HM379302

continued

WebTable 1. – ***continued***

Sample source (KEW; BER= Berlin; PAR= MNHN Paris; WIE = Vienna; JEN= Jena; NEW= 2008)	Date of collection	Country (AL=Albania; GR=Greece; BG= Bulgaria)	Locality	Est latitude (decimal)	Est longitude (decimal)	Est Geog error (km)	Est elevation (m)	Elevation error (m)	Natural/ornamental population	Closest known site (in Avtzis et *al.* 2007, Valade et *al.* 2009)	Distance to closest known site (in Avtzis et *al.* 2007, Valade et *al.* 2009)	# herbarium sheets	# larvae/pupae *C ohridella* excised	Sample reference number	Institutional voucher number (INRA-URZF-), BOLD accession number (CCDB-; CO-ARC-) GenBank Accession number (JF-, HM-)
NEW- C. Lopez-Vaamonde and S. Augustin	10/06/2008	GR	Karitsa	39.85	22.73	0.15	650	75	N				29	K1-K2: K3: K4-K7: K9-K10: K11 K12-K24: K25: K26-K30:	JF746752-JF746753 JF746778 JF746754-JF746757 JF746758-JF746759 JF746779 JF746760-JF746772 JF746780 JF746773-JF746777
NEW- C. Lopez-Vaamonde and S. Augustin	4/06/2008	GR	Monodendri	39.894	20.739	0.15	1300	50	N				26	M01: M02: M03: M04-M06: M07: M08: M09: M10: M11: M12-M13: M14: M15: M17: M18: M19: M20-M21: M22: M23: M24-M25: M26: M27:	JF746799 JF746781 JF746800 JF746782-JF746784 JF746801 JF746792 JF746785 JF746803 JF746791 JF746793-JF746794 JF746786 JF746802 JF746787 JF746795 JF746805 JF746788-JF746789 JF746796 JF746790 JF746797-JF746798 JF746804 JF746804

*The northernmost of the two red triangles corresponding to the "Chaliki" label in Figure 1.

Notes: A more complete list of leaf mines from these herbarium specimens and their frequency and density are given in WebTable 8. After this article was accepted, we located a further three horse-chestnut specimens from native populations that contained *C ohridella* mines in the herbarium of the Natural History Museum in London, whose collections had earlier been in the process of being moved. Two of these collections are duplicates of above documented specimens: A.H. Alston and N.Y. Sandwith 2365 and Baldacci 129. One of these collections represents an additional native horse-chestnut site in Greece with archival mines: Thess., in valle Aspri pr Pyrrha, 20.VIII. 1934, F. Guiol 2453 (BM) [Greece, Thessaly, Nomos and Eparchia Trikala, village of Pirra, W of Pertouli, in a tributary valley of the Acheloos river, c. 39°32'N/21°24'E].

WebTable 2. Primers used in this study and their amplification success.

Primer name	*F/R*	*Primer sequence 5' → 3'*	*bp*	*Reference*
LCO-1498	F	GGTCAACAAATCATAAAGATATTGG		Folmer *et al.* (1994)
HCO-2198	R	TAAACTTCAGGGTGACCAAAAAATCA	658	"
Uni-MinibarF1	F	TCCACTAATCACAARGATATTGGTAC		Meusnier *et al.* (2008)
Uni-MinibarR1	R	GAAAATCATAATGAAGGCATGAGC	126	"
LepF1	F	ATTCAACCAATCATAAAGATATTGG		Hajibabaei *et al.* (2006a)
CamR1a	R	GGAACTAATCARTTACCAAATCC	182	This paper
CamF2a	F	TTAGGAAATCCTRGATCTTTAATTGG		"
CamR2a	R	TGAAATTAAAAGTAATATTGAAGGTGG	170	"
CamF3a	F	ATAAGATTTTGATTATTRCCACC		"
CamR3a	R	GCTCCYAAAATAGAAGAAATTCC	116	"
CamF5a	F	ATTTTTTCATTACATTTRGCTGG		"
CamR5a	R	GTAATAGCTCCTGCTAAAACAGG	137	"
CamF6a	F	CCATTATTTGTTTGAGCTGTTGG		"
LepR1	R	TAAACTTCTGGATGTCCAAAAAATCA	140	Hajibabaei *et al.* (2006a)
Ohrid2753	F	FAM-AGAGGCCATAGGCGCTTAAC		
	R	AGTAGAGGACGCCCACGAAG	215–237	Mari Mena *et al.* (2008)
Ohrid2759	F	NED-AAGGAGTTGGCACAGGACAG		
	R	GGGTATCGGACAAGTTTTAACG	149–199	"
Ohrid2762	F	NED-TTGCTCGTCTTCCAAGTCCC		"
	R	TCCGACCAACCCCAACAC	117–137	"
Ohrid2782	F	HEX-TTCTTTATTGGCTTATCCGC		"
	R	CTGCATAATCTAAGTTTCCATGTC	**102–132**	
Ohrid2794*	F	PET-CTGCATAATCTAAGTTTCCATGTC		"
	R	TTCTTTATTGGCTTATCCGG	113–133	"
Ohrid2814	F	FAM-ACCGTAAAGATAATATTTAACCCG		"
	R	GTGAAAGTTTTTGTTTGAATTAGC	**93–108**	"

Notes: Primer pairs regularly providing amplicons for archival material, all working at annealing temperatures specified in WebPanel 1, shown in bold. Size of amplicon in base pairs (column heading "bp") does not include that of primers. Allele size range for two microsatellite primers updated from Mari Mena *et al.* (2008) using data from Valade *et al.* (2009). We obtained COI amplicons up to 182 bp and microsatellites up to 132 bp, for at least one *C ohridella* specimen in 12 of those 13 archival *A hippocastanum* specimens (representing half of the collection events) for which we extracted larvae/pupae (WebTable 1). For COI, primer pair CamF3a/CamR3a was the most informative (see also Figure 3). We got bands in one 1936 sample and 75% of 40 post-1960 samples for one or more primer pairs; we obtained clean sequences from 10 of these (Figure 3). Overall, microsatellite amplification succeeded only for markers Ohrid2814 and Ohrid2782 in 37/54 (68.5%) of pre-imaginal samples extracted from herbarium specimens: see WebTable 3. The 102–104 bp amplicons were obtained for 68% of specimens back to 1879, 10 of these from 1879 to 1936; three/eight 1879 samples (102–122 bp) and two/four 1895 samples (104–126 bp) provided peaks. Dedicated species-specific primers appeared crucial, with the Uni-Minibar pair performing poorly and Sphingidae-designed primers providing no amplicons (though see also Lees *et al.* 2010). *Identical to Ohrid 2782 (a fact not noted by Mari Mena *et al.* 2008) except for the fluorescent marker, so this primer pair is presumably redundant.

WebTable 3. Comparison of archival (1879–1981, CO sample numbers, this study) and modern (2002–2008, Valade *et al*. 2009) microsatellite data from *C ohridella* larvae from natural *Aesculus* localities in Albania (Stravaj) and Greece (other sites).

Site, year, sample number	*µsat-2782 Allele 1*	*µsat-2782 Allele 2*	*µsat-2814 Allele 1*	*µsat-2814 Allele 2*
Karitsa, 1936, CO40	-	-	102	104
Karitsa, 1974, CO51	**114**	**114**	102	104
Karitsa, 1974, CO52	-	-	**104**	**104**
Karitsa, 1974, CO53	-	-	**104**	**104**
Tsepélovo, 1964, CO12	124	124	102	102
Syrakou, 1895, CO21	120	122	-	-
Syrakou, 1895, CO22	-	-	102	104
Syrakou, 1895, CO25	-	-	104	104
Chelidoni, 1879, CO09	-	-	104	104
Chelidoni, 1879, CO09	120	124	-	-
Chelidoni, 1879, CO34	-	-	104	104
Chelidoni, 1879, CO39	124	126	104	104
Çajup[i]-Zhej, 1935, CO44	-	-	102	104
Tresova, 1928, CO01	-	-	104	104
Tresova, 1928, CO02	-	-	104	104
Tresova, 1928, CO03	**119**	**119**	104	104
Tepelenë, 1961, CO60	120	126	102	104
Tepelenë, 1961, CO61	120	120	102	102
Tepelenë, 1961, CO62	-	-	104	104
Tepelenë, 1961, CO64	-	-	102	102
Tepelenë, 1961, CO64	-	-	102	102
Tepelenë, 1961, CO65	120	**122**	102	104
Tepelenë, 1961, CO66	**119**	120	102	102
Tepelenë, 1961, CO67	120	124	102	104
Tepelenë, 1961, CO69	-	-	102	102
Tepelenë, 1961, CO70	-	-	102	102
Tepelenë, 1961, CO80	-	-	102	102
Tepelenë, 1961, CO81	-	-	102	102
Tepelenë, 1961, CO82	-	-	104	104
Tepelenë, 1961, CO84	-	-	102	102
Tomor, 1961, CO83	**128**	**128**	-	-
Tomor, 1961, CO85	118	124	104	104
Tomor, 1961, CO87	-	-	102	102
Tomor, 1961, CO88	-	-	102	102
Tomor, 1961, CO89	-	-	102	102
Tomor, 1961, CO92	-	-	102	102
Tomor, 1961, CO93	-	-	102	102
Stravaj, 2006, RV130c	**112**	**112**	102	102
Stravaj, 2006, EM72	117	124	104	104
Stravaj, 2006, EM67	117	128	104	104
Stravaj, 2006, RV130	118	120	104	104
Stravaj, 2006, RV243	118	124	104	104
Stravaj, 2006, RV244	118	126	104	104
Stravaj, 2006, EM68	118	126	104	104
Stravaj, 2006 EM73	120	120	-	-
Stravaj, 2006, RV241	120	120	102	102
Stravaj, 2006, EM69	120	124	104	104
Stravaj, 2006, EM70	120	126	104	104
Stravaj, 2006, EM74	120	126	104	104
Stravaj, 2006, RV130d	120	126	102	104
Stravaj, 2006, EM64	120	128	104	104

continued

WebTable 3. – *continued*

Site, year, sample number	μsat-2782 Allele 1	μsat-2782 Allele 2	μsat-2814 Allele 1	μsat-2814 Allele 2
Stravaj, 2006, EM66	122	122	104	104
Stravaj, 2006, EM65	122	124	104	104
Stravaj, 2006, RV240	122	124	102	104
Stravaj, 2006, RV130b	122	127	102	102
Stravaj, 2006, EM71	124	128	-	-
Stravaj, 2006, RV242	124	129	104	104
Stravaj, 2006, RV130a	126	129	101	101
Karitsa, 2008, Karitsa27	108	115	103	103
Karitsa, 2008, Karitsa15	111	119	103	103
Karitsa, 2008, Karitsa19	115	123	102	102
Karitsa, 2008, Karitsa17	115	123	103	103
Karitsa, 2008, Karitsa16	115	123	103	103
Karitsa, 2008, Karitsa30	115	123	103	103
Karitsa, 2008, Karitsa18	115	124	102	102
Karitsa, 2008, Karitsa26	115	124	103	105
Karitsa, 2008, Karitsa6	116	122	**102**	**104**
Karitsa, 2008, Karitsa20	116	124	102	105
Karitsa, 2008, Karitsa1	116	124	**102**	**104**
Karitsa, 2008, Karitsa2	116	124	**102**	**104**
Karitsa, 2008, Karitsa10	117	124	103	105
Karitsa, 2008, Karitsa23	117	125	103	103
Karitsa, 2008, Karitsa22	117	126	102	102
Karitsa, 2008, Karitsa13	117	126	103	105
Karitsa, 2008, Karitsa9	117	126	103	103
Karitsa, 2008, Karitsa29	120	124	103	103
Karitsa, 2008, Karitsa8	**122**	**122**	102	106
Karitsa, 2008, Karitsa12	122	128	103	103
Karitsa, 2008, Karitsa28	123	126	103	103
Karitsa, 2008, Karitsa11	**124**	**124**	103	103
Karitsa, 2008, Karitsa25	124	126	102	102
Karitsa, 2008, Karitsa4	124	126	**102**	**104**
Karitsa, 2008, Karitsa24	124	126	103	103
Karitsa, 2008, Karitsa7	124	126	104	106
Karitsa, 2008, Karitsa14	**126**	**126**	102	102
Karitsa, 2008, Karitsa21	**126**	**126**	103	104
Karitsa, 2008, Karitsa5	**126**	**126**	104	102
Karitsa, 2008, Karitsa3	126	128	104	106
Monodendri, 2008, Mono18	118	118	102	104
Monodendri, 2008, Mono10	118	118	102	104
Monodendri, 2008, Mono3	118	120	**102**	**102**
Monodendri, 2008, Mono8	118	124	102	104
Monodendri, 2008, Mono14	118	126	102	104
Monodendri, 2008, Mono21	118	126	102	104
Monodendri, 2008, Mono17	118	128	102	104
Monodendri, 2008, Mono28	120	120	102	104
Monodendri, 2008, Mono24	120	120	102	104
Monodendri, 2008, Mono22	120	120	102	104
Monodendri, 2008, Mono12	120	120	104	104
Monodendri, 2008, Mono26	120	120	104	106
Monodendri, 2008, Mono2	120	122	102	104
Monodendri, 2008, Mono23	120	122	102	104

continued

WebTable 3. – *continued*

Site, year, sample number	*µsat-2782 Allele 1*	*µsat-2782 Allele 2*	*µsat-2814 Allele 1*	*µsat-2814 Allele 2*
Monodendri, 2008, Mono16	120	122	102	104
Monodendri, 2008, Mono1	120	122	102	104
Monodendri, 2008, Mono25	120	122	102	104
Monodendri, 2008, Mono5	120	124	104	104
Monodendri, 2008, Mono9	120	128	104	106
Monodendri, 2008, Mono27	120	130	102	104
Monodendri, 2008, Mono11	120	131	102	104
Monodendri, 2008, Mono6	120	132	102	104
Monodendri, 2008, Mono13	122	124	102	104
Monodendri, 2008, Mono15	122	124	104	106
Monodendri, 2008, Mono19	122	132	102	104
Monodendri, 2008, Mono7	**124**	**124**	102	104
Monodendri, 2008, Mono20	126	126	102	104
Monodendri, 2008, Mono4	126	126	102	104
Perivóli, 2002, RV153h	110	122	102	102
Perivóli, 2002, RV191	110	126	-	-
Perivóli, 2002, RV192	117	122	102	102
Perivoli, 2002, RV214	117	122	102	102
Perivóli, 2002, RV153i	117	126	104	108
Perivóli, 2002, RV153g	118	124	102	102
Perivóli, 2002, RV190	120	120	102	102
Perivóli, 2002, RV218	120	120	102	102
Perivóli, 2002, RV153c	120	122	102	104
Perivóli, 2002, RV184	120	122	102	104
Perivóli, 2002, RV196	120	122	102	102
Perivóli, 2002, RV188	120	122	104	104
Perivóli, 2002, RV153	120	124	102	102
Perivóli, 2002, RV215	120	124	104	104
Perivóli, 2002, RV185	120	126	-	-
Perivóli, 2002, RV189	120	126	102	102
Perivóli, 2002, RV153f	120	126	104	104
Perivóli, 2002, RV195	120	126	104	104
Perivóli, 2002, RV153b	120	130	102	102
Perivóli, 2002, RV198	120	131	102	102
Perivóli, 2002, RV186	122	124	-	-
Perivóli, 2002, RV187	122	124	102	102
Perivóli, 2002, RV153a	122	124	104	104
Perivóli, 2002, RV193	122	124	104	104
Perivóli, 2002, RV153d	122	126	102	102
Perivóli, 2002, RV153e	122	128	102	102
Perivóli, 2002, RV199	122	130	102	102
Perivóli, 2002, RV194	124	124	102	102
Perivóli, 2002, RV217	124	126	102	102
Perivóli, 2002, RV216	124	128	102	102
Perivóli, 2002, RV197	126	126	102	102
Ondria, 2002, AH154	118	120	102	104
Ondria, 2002, AH151	118	126	102	104
Ondria, 2002, AH170	120	120	102	104
Ondria, 2002, AH164	120	120	103	104
Ondria, 2002, AH159	120	122	102	104
Ondria, 2002, AH180	120	122	102	104

continued

WebTable 3. – *continued*

Site, year, sample number	*µsat-2782 Allele 1*	*µsat-2782 Allele 2*	*µsat-2814 Allele 1*	*µsat-2814 Allele 2*
Ondria, 2002, AH165	120	122	104	104
Ondria, 2002, AH176	120	122	104	108
Ondria, 2002, AH174	120	124	96	104
Ondria, 2002, AH155	120	124	102	104
Ondria, 2002, AH156	120	124	102	104
Ondria, 2002, AH160	120	126	102	104
Ondria, 2002, AH167	120	126	102	104
Ondria, 2002, AH177	120	126	102	104
Ondria, 2002, AH178	120	126	102	104
Ondria, 2002, AH169	122	122	96	104
Ondria, 2002, AH157	122	122	102	104
Ondria, 2002, AH175	122	124	102	104
Ondria, 2002, AH152	124	126	102	104
Ondria, 2002, AH161	124	126	102	104
Ondria, 2002, AH163	124	126	102	104
Ondria, 2002, AH171	124	126	102	104
Ondria, 2002, AH173	124	126	104	108
Ondria, 2002, AH168	126	126	93	100
Ondria, 2002, AH153	126	126	102	104
Ondria, 2002, AH158	126	126	102	104
Ondria, 2002, AH162	126	126	102	104
Ondria, 2002, AH166	126	126	102	104
Ondria, 2002, AH172	126	126	102	104
Ondria, 2002, AH179	126	126	102	104
Tsotyli, 2008, N64	-	-	102	104
Tsotyli, 2008, N63	118	120	102	104
Tsotyli, 2008, N61	120	124	102	102
Tsotyli, 2008, N47	120	126	102	102
Tsotyli, 2008, N65	120	126	102	104
Tsotyli, 2008, N50	121	122	102	104
Tsotyli, 2008, N48	121	126	104	104
Tsotyli, 2008, N44	122	122	102	102
Tsotyli, 2008, N38	122	126	102	102
Tsotyli, 2008, N60	122	126	102	104
Tsotyli, 2008, N66	122	126	102	104
Tsotyli, 2008, N37	124	126	102	104

Notes: Data rounded to the nearest bp. Locally private alleles for archival samples are indicated in bold. For example, those from Albania were previously detected, within rounding error, in the Valade *et al.* dataset at Tsotyli, Greece. By contrast, the allele for specimen CO51 from Karitsa (1974) for the marker Ohrid2782 was measured at 1–2 bp different (actual measurements were 113.88 versus 112.44) from the homozygous allele shown in bold, detected at Stravaj (RV130c), and might thus be novel.

WebTable 4. Genetic diversity (observed [Ho] and expected [He] heterozygosity) for *C ohridella* archival (Tepelenë, Tomor) and recent (Stravaj) populations from Albania.

	Stravaj	*Tepelenë*	*Tomor*
Ho	0.378	0.407	0.25
He	0.635	0.585	0.667

Notes: There is a lower than expected observed heterozygosity (as for Tomor) revealing deviation from the Hardy-Weinberg equilibrium, probably due to reduced gene flow and inbreeding, both common in isolated populations.

WebTable 5. Pairwise genetic differentiation for recent and archival Albanian populations of *C ohridella*.

Phi_st	*Stravaj*	*Tepelenë*
Stravaj (*n* = 21)	--	
Tepelenë (*n* =14)	0.072 **(0.015, *)**	--
Tomor (*n* = 7)	0.086 **(0.038, *)**	0.083 (0.122)

Notes: Significant *P* values are shown in bold. The allelic structure of the Stravaj population was significantly differentiated from both Tepelenë and Tomor (in bold), which were not significantly different from each other.

WebTable 6. Proportion of membership of individuals from each sample location in each of the three population clusters inferred by Structure 2.3, despite the moderate sample sizes, confirming the result in WebTable 5.

	Clusters		
Ancient populations	1	2	3
Tepelenë (*n* = 14)	0.269	0.271	**0.460**
Tomor (*n* = 7)	0.267	0.267	**0.466**
Recent population			
Stravaj (*n* = 21)	**0.373**	**0.373**	0.253

Notes: The probability that the individuals represent three groups was marginally higher (ln Likelihood = –151.3; K = 3; –152.4 for K = 2), with substantial differences among populations with respect to assignment of individuals to these three groups. Group 3 contains the two herbarium (ancient) populations, whereas the other two groups cluster all of the individuals from the recent population. The percentage of individuals assigned to each cluster is high for the herbaria populations, while individuals from the recent population were assigned in equal proportions to clusters 1 and 2. Proportions > 0.33 are in bold.

WebTable 7. Allele count for the three Albanian populations of *C ohridella* compared for each locus.

Ohrid2782											
Population	**112**	**116**	**117**	**118**	**119**	**120**	**122**	**124**	**126**	**127**	**128**
Stravaj	**2**	**6**	0	**6**	0	6	8	1	7	0	**4**
Tepelenë	0	0	0	0	**2**	5	1	1	1	0	0
Tomor	0	0	**1**	0	0	0	0	1	0	**2**	0

Ohrid2814			
Population	**100**	**102**	**104**
Stravaj	**9**	27	0
Tepelenë	0	21	7
Tomor	0	10	2

Notes: Locally private alleles are shown in bold. Interestingly, both Stravaj and Tomor/Tepelenë show alleles that are locally private in the context of this analysis (WebTable 6).

WebTable 8. Quantities of mines in relation to examined leaf area on herbarium sheets and details of early stages of *C ohridella* excised from herbarium specimens.

Sample source (KEW; BER= Berlin; PAR= MNHN Paris; WIE = Vienna; JENA= JEN; NEW= 2008	Date of collection	Locality	# herbarium sheets	# leaves	Est # leaflets (mature)	Leaf area (cm^2)	# mines *C ohridella*	# larvae/pupae *C ohridella* excised	Life stage of *C ohridella*	Mine reference number
KEW- T.Heldreich	11/08/1879	Mt. Chelidoni	2	5	35 (21)	1103	4	4	Pupal exuvium SP1-2 L2 SP1	CO37 CO38 CO39 CO46
WIE- T.Heldreich-W 0023011; W 1889-0315748	11/08/1879	Mt. Chelidoni	2	7	49 (35)	979	3	?	Not sampled	
PAR- T.Heldreich	11/08/1879	Mt. Chelidoni	2	4	28	2119	4	2	Not excised Pupal exuvium L3-4 Not excised	CO7 CO8 CO9 C10
PAR- T.Heldreich	11/08/1879	Mt. Chelidoni	2	4	26	1621	5	1	Not excised Not excised Pupa Not excised Not excised	CO32 CO33 CO34 CO35 CO36
PAR- T.Heldreich	24/06/1885	Agrafa	2	5	16	587	0	0		
KEW- C.Haussknecht	07/1885	Chaliki- Krania	1	3	21 (14)	645	0	-		
KEW- A.Baldacci-129	03/08/1895	Syrakou	1	3	15 (10)	433	0	-		
PAR - A.Baldacci-129	03/08/1895	Syrakou	4	8	35	1587	21	6	Not excised L4? L3? Not excised L3? SP1 Not excised SP1-2 L3 Not excised	CO14-CO15 CO16 CO17 CO18-CO20 CO21 CO22 CO23 CO24 CO25 CO26-CO31
KEW- P.Sintenis-692	18/06/1896	Chaliki	1	3	13 (5)	427	0	-		
PAR- P.Sintenis-692, 693	18/06/1896	Chaliki	3	6	34	966	0	-		
KEW- J.Mattfeld-2562	01/08/1926	Kaliakuda	3	20	70 (35)	1346	0	-		
WIE- J.Mattfeld	01/08/1926	Kaliakuda	1			622	0	-		
PAR- J.Mattfeld	01/08/1926	Kaliakuda	1	5	27	353	0	-		
KEW- N. Stojanoff	11/05/1928	Preslavsk[a] Balkan	1	3	18 (18)	369	0	-		
BER- Markgraf-1513	28/06/1928	Tresova	1	5	25 (15)	331	5	3	L4 SP2? L3? - Empty	

continued

WebTable 8. – *continued*

Sample source (KEW; BER= Berlin; PAR= MNHN Paris; WIE = Vienna; JENA= JEN; NEW= 2008	Date of collection	Locality	# herbarium sheets	# leaves	Est # leaflets (mature)	Leaf area (cm^2)	# mines *C ohridella*	# larvae/pupae *C ohridella* excised	Life stage of *C ohridella*	Mine reference number
KEW- A.H.G.Sandwith and N.Y.Alston-2365	08/08/1935	Çajup-Zhej	1	5	30 (18)	755	2	2	Pupa SP1	CO43 CO44
KEW-O.Grebenchikoff	29/07/1936	Karitsa	1	9	35 (10)	780	3	2	SP2 L3 Empty	CO40 CO41 CO42
KEW-K.H.Rechinger-1509	17/07/1958	Skamneli	1	1	7 (7)	380	0	-	-	
BER-K.H.Rechinger 21509	17/07/1958	Skamneli	1	6	42 (21)	294	0	-	-	
JEN- Meyer 3316	25/06/1959	Mali I Gjer	3	14	45	835	3	0	Empty	
JEN- Meyer 5430	04/05/1960	Bei [near] Tresova	5	19	94	1508	0	0	Empty	
JEN- Meyer 6004	08/09/1961	Griba, Sevaster	4	12	70	1146	>780	3	Larvae	CO95-CO97
JEN- Meyer 6057	09/09/1961	Tepelen[ë]	4	6	30	1156	>296	18	Larvae Larvae L2-SP2	CO57-CO67 CO69-CO72 CO80-CO82
JEN- Meyer 6380	22/09/1961	[Mt] Tomor	6	18	91	2572	>1365	11	Larva L2-SP2 Larvae(6)/ pupae (4)	CO84 CO83 CO85-CO93
PAR- G.G.Aymonin	23-26/06/1964	Tsipélovo	1	4	28	603	3	1	Empty L3 [*Pnigalio* pupal parasitoid]	CO11 CO12 CO13
BER/Raus- T. Raus	05/07/1974	Karitsa	12	5	60 (15)	568	5	4	L2	CO51-CO54
BER–T. Raus & Royl 5356	12/09/1981	Karitsa	2	6	30	676	4	1	L4-5	CO100