

Allopatry as a Gordian Knot for Taxonomists: Patterns of DNA Barcode Divergence in Arctic-Alpine Lepidoptera

Marko Mutanen^{1*}, Axel Hausmann², Paul D. N. Hebert³, Jean-François Landry⁴, Jeremy R. de Waard⁵, Peter Huemer⁶

1 Zoological Museum, Department of Biology, University of Oulu, Oulu, Finland, **2** Zoologische Staatssammlung München, Sektion Lepidoptera, München, Germany, **3** Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, Canada, **4** Agriculture and Agri-Food Canada, Eastern Cereal and Oilseed Research Centre, Ottawa, Ontario, Canada, **5** Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, Canada, **6** Tiroler Landesmuseen Betriebsges.m.b.H., Innsbruck, Austria

Abstract

Many cold adapted species occur in both montane settings and in the subarctic. Their disjunct distributions create taxonomic complexity because there is no standardized method to establish whether their allopatric populations represent single or different species. This study employs DNA barcoding to gain new perspectives on the levels and patterns of sequence divergence among populations of 122 arctic-alpine species of Lepidoptera from the Alps, Fennoscandia and North America. It reveals intraspecific variability in the barcode region ranging from 0.00–10.08%. Eleven supposedly different species pairs or groups show close genetic similarity, suggesting possible synonymy in many cases. However, a total of 33 species show evidence of cryptic diversity as evidenced by the presence of lineages with over 2% maximum barcode divergence in Europe, in North America or between the two continents. Our study also reveals cases where taxonomic names have been used inconsistently between regions and exposes misidentifications. Overall, DNA barcodes have great potential to both increase taxonomic resolution and to make decisions concerning the taxonomic status of allopatric populations more objective.

Citation: Mutanen M, Hausmann A, Hebert PDN, Landry J-F, de Waard JR, et al. (2012) Allopatry as a Gordian Knot for Taxonomists: Patterns of DNA Barcode Divergence in Arctic-Alpine Lepidoptera. PLoS ONE 7(10): e47214. doi:10.1371/journal.pone.0047214

Editor: Anna Tramontano, University of Rome, Italy

Received: June 15, 2012; **Accepted:** September 10, 2012; **Published:** October 11, 2012

Copyright: © 2012 Mutanen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by: Genome Canada (<http://www.genomecanada.ca/>), Ontario Genomics Institute (<http://www.ontariogenomics.ca/>), Ontario Ministry of Economic Development and Innovation (http://www.ontariocanada.com/ontcan/1medt/en/home_en.jsp), NSERC (<http://www.nserc-crsng.gc.ca/>), Emil Aaltonen Foundation (<http://www.emilaaaltonen.fi/>), and Kone Foundation (<http://www.koneensaatio.fi/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: marko.mutanen@oulu.fi

Introduction

Species delimitation is not straightforward. Young species pairs often show limited morphological and genetic divergence and decisions on their status are complicated because the acquisition of diagnostic characters does not always happen in the same order or at the same rate. As well, different species concepts emphasize different properties [1–2]. For example, the biological species concept highlights the importance of reproductive incompatibility, while the phylogenetic species concept only requires diagnosability. In practice, most species owe their description to the study of morphological variation, but some are based on the analysis of both molecular and morphological characters. The coupling of differing species concepts with variation in the characters examined has created an undesirable level of subjectivity in species delineation, particularly for taxa with allopatric ranges. Many species found in both alpine and arctic habitats fall into the latter category because their ranges are fragmented, reflecting the discontinuous distribution of the habitats that they occupy. Their disjunct distributions were gained through range shifts following deglaciation as rising temperatures provoked both the northward movement of populations and the shift of southern populations to higher elevations on mountains. Separated by broad zones of deciduous and boreal forest, gene flow between populations from the Alps and Scandinavia has now been halted for at least 10,000

years. Gene flow between North American and Eurasian conspecifics across the Beringian land bridge [3] was interrupted at about the same time, reflecting the postglacial flooding of the Bering Strait.

Few prior studies have examined the levels of sequence divergence across broad geographical areas in a large number of taxa. Past phylogeographic studies on arctic-alpine species have targeted single taxa, such as the moth *Zygaena exulans* [4] and the butterfly *Erebia ephron* [5]. Prior barcoding studies have examined many species, but they have usually focused on relatively small geographic regions. Work on Central Asian butterflies [6] is exceptional as it examined patterns of sequence diversity on a larger geographic scale. It established that the performance of DNA barcodes in differentiating species was not significantly reduced as geographic coverage expanded. It also showed a high correspondence in levels of morphological and genetic differentiation between allopatric populations, suggesting that decisions concerning the status of allopatric species can be made in a more standardized way than in the past.

In this study, we examine levels and patterns of barcode divergence among 122 species of Alpine, Fennoscandian and Nearctic populations of Lepidoptera that have almost certainly experienced very limited or no gene flow for more than 10K years. We have also tested the potential value of DNA barcode data to aid the delineation of species.

Materials and Methods

Taxon Sampling and Nomenclature

We initially targeted about 170 arctic-alpine species of Lepidoptera shared by Fennoscandia and the Alps, but we could only obtain representatives from 116 of these species, belonging to 26 families, from both areas (845 specimens in total, see Table S1 and Table S2 for details). We subsequently examined potential conspecifics from North America, analyzing 29 species with representatives from all three areas and 2 with arctic-alpine representatives, but absent from Fennoscandia (Table S1). We did not examine specimens from the subarctic or montane regions of Asia. Species assignments followed current taxonomy, which is mainly based on external morphology and genitalia. In total, we examined 1424 specimens with DNA barcode sequence data belonging to 122 species according to the nomenclature of Fauna Europaea [7] and Hodges *et al.* [8]. 1331 of these records are previously unpublished. Table S1 provides taxonomic authorities for all taxa.

DNA Sequencing and Analysis

Sequences for the barcode region were obtained at the Canadian Centre for DNA Barcoding (CCDB) using standard protocols [9]. A full barcode sequence (658 bp) was recovered from 1180 specimens (118 species), sequences greater than 500 bp from 1399 specimens (121 species) and shorter sequences from 25 specimens (one species). Full length barcodes were recovered from at least one specimen for all species except *Stenoptilia alpinalis*, *S. buwati*, *Xestia rhaetica*, and *X. fennica*. Barcode records for the European specimens are available in the BOLD [10] dataset “DATASET-AALE1”, accessed at <http://dx.doi.org/10.5883/DATASET-AALE1>. Data for the North American specimens are available in the BOLD dataset “DATASET-ALNA1”, accessed at <http://dx.doi.org/10.5883/DATASET-ALNA1>. All sequences are also available on GenBank under the accession numbers provided in Table S1.

Sequence divergences were quantified using the Kimura 2-parameter model of nucleotide substitution calculated with the analytical tools on BOLD. We determined the maximum, mean and minimum intraspecific variation for each species and then separately for the three regions (Alps, Fennoscandia, North America) and between the regions. BOLD provides tools for calculating mean and maximum intraspecific variation values for each species, but analysis is based on species names so minimum and maximum distances within and between regions were obtained by assembling all pairwise values for each comparison type. For example, the mean divergence between Alpine and Fennoscandian populations involved the assembly of all pairwise distances for individuals from these two regions and then calculating their mean. Neighbor-joining (NJ) similarity trees were constructed with MEGA 5.05 [11] using the Kimura 2-parameter model of base substitution (with pairwise deletion of missing data).

Results

Molecular Divergence - Overview

Intraspecific variation averaged 1.08%, but ranged from 0.00%–10.08%. The mean value was increased by a few cases (discussed later) of probable cryptic species, most involving allopatric lineages. Intraspecific variation was considerably lower within single geographic regions, averaging 0.65% for North America, 0.39% for the Alps and 0.22% for Fennoscandia with levels of variation in the same rank order as the size of the regions (Figures 1 and 2). Mean intraspecific divergence was higher for

comparisons between regions, averaging 1.03% between Fennoscandia and the Alps, 2.59% between the Alps and North America and 2.43% between Fennoscandia and North America (Figures 1 and 2). These values were roughly proportional to the distance between the areas, but precision of the comparison is impeded by taxonomic uncertainty (see below). Table S3 provides exact values of intraspecific variation within and between regions.

Low Intraspecific Barcode Divergence

Sequence divergences higher than 2% in the barcode region often correspond to interspecific differences, while lower values are typical of intraspecific variation [12]. Of course, young sister-species may fall below the 2% threshold, while unusually variable species may exceed it. In our study areas, most species represented by multiple individuals followed the rule of less than 2% maximum divergence. In fact, 82 of 89 (92%) Fennoscandian species met this criterion and 78 of these species possessed less than 1% divergence. The same pattern was evident in Alpine specimens with 59 of 67 (88%) species falling below 2% divergence and 48 species falling below 1%. More intraspecific variation was apparent in North America as just 12 of 19 (63%) species fell below the 2% threshold. A comparison between the two regions in Europe showed that 66 of 86 (77%) species from both Fennoscandia and Alps had a maximum divergence below 2%, but the fraction of species with low divergence was much less for intercontinental comparisons. Only 6 of 18 (33%) species from Fennoscandia and North America and 6 of 19 (32%) species from the Alps and North America had a barcode divergence of less than 2%.

Low Barcode Divergence in Different Taxa – Conspecificity versus Barcode Similarity

Our analyses reveal 11 possible cases of overlooked synonymy (as indicated by close barcode similarity) including five species pairs and one quintet in Europe (Figure 3, Table S4). Five species pairs (*Erebria medusa* - *E. polaris*, *Oeneis glacialis* - *O. norma*, *Holarctia cervini* - *H. puengeleri*, *Apamea maillardi* - *A. schildei*, *Xestia fennica* - *X. rhaetica*) showed less than 1% maximum sequence divergence between the Alps and northern Europe (Figure 3A). As well, a group of five species, including the Fennoscandian *Stenoptilia islandica* and four Alpine endemics (*S. alpinalis*, *S. buwati*, *S. brigantiensis*, *S. mercantourica*) showed less than 1% divergence (Figure 3B). As all these taxa are all extremely similar in external morphology and genitalia, their species rank needs reconsideration. In fact, the status of most of these taxa is controversial. For example, *Xestia rhaetica* was treated as a subspecies of *X. fennica* by Kullberg *et al.* [13], but as a separate species by Fibiger *et al.* [14]. The latter authors viewed *Apamea schildei* as distinct from *A. maillardi*, reversing a decision on its subspecific status made just a few years earlier [15]. Similarly, *Holarctia cervini* and *H. puengeleri* have traditionally been viewed as subspecies, but more recently as separate species [16–17]. The taxonomic status of *Stenoptilia* is also uncertain as the four alpine taxa in our study have been treated both as junior synonyms of *S. pelidnodactyla* [18] and as distinct species [19]. Our data provide another option – their possible synonymy with *S. islandica* (Figure 3B). The species pairs of *Erebria* and *Oeneis* provide a final example of low genetic divergence and strong morphological similarity between allopatric lineages.

We encountered one triad and four species pairs where specimens from our three study areas showed low divergence. The group of three species included *Xanthorhoe incurvata* from the Alps, *X. annotinata* from Finland and *X. baffinensis* from North America (Figure 3D). They showed very low divergence (MIN = 0.08%, MAX = 0.66%), despite the fact that the two

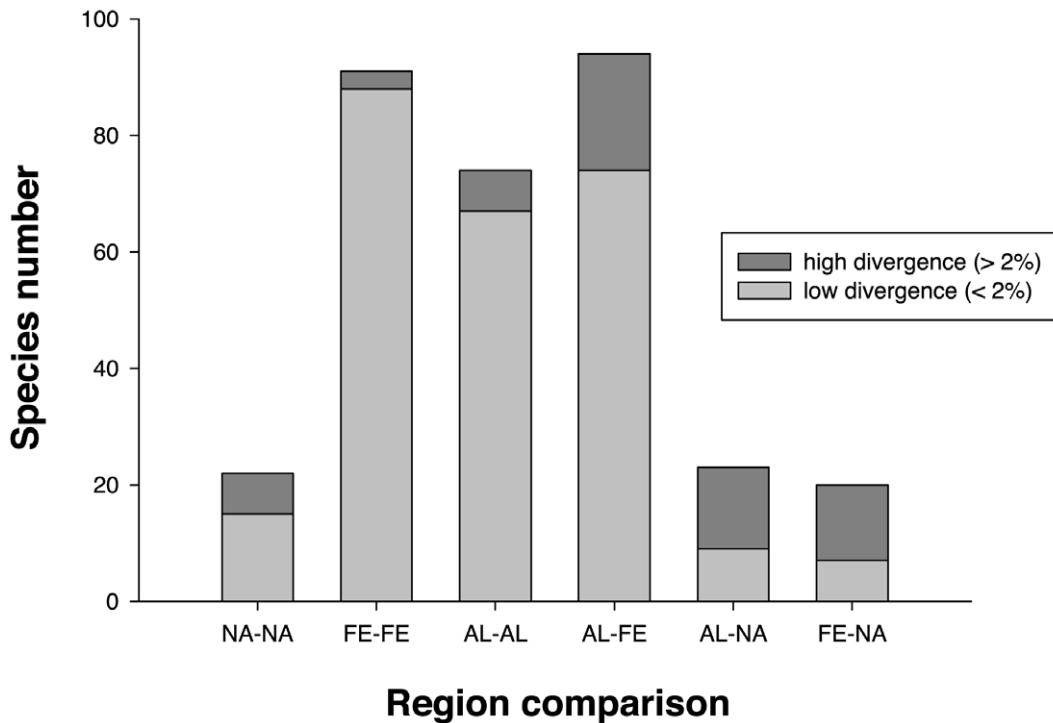


Figure 1. Proportion of high (>2%) intraspecific maximum divergences in DNA barcodes within and between regions. NA (North America), AL (Alps), FE (Fennoscandia). doi:10.1371/journal.pone.0047214.g001

European taxa have very different genital morphology. The four species pairs included:

- 1) European lineages of *Xanthorhoe decoloraria* and North American *X. alticolata* had little divergence (MIN = 0.77%) (Figure 3C). Interestingly, European and North American lineages of *X. decoloraria* showed more divergence (MIN = 1.55%).
- 2) *S. nigrita* from the Alps had substantial divergence (MIN = 2.34%) from Fennoscandian lineages of this species, while North American *Sympistis zetterstedtii* were very close (MIN = 1.08%) to Finnish *S. nigrita* (Figure 3E). The taxonomy of these species has been controversial. Skou [20] treated *S. zetterstedtii* and *S. nigrita* as distinct species, but Ronkay & Ronkay [21] noted the genitalic similarity of *S. nigrita* and *S. zetterstedtii* versus the marked morphological difference between alpine and Fennoscandian *S. nigrita*. On this basis they recognized four subspecies: *S. nigrita nigrita* (Alps), *S. nigrita zetterstedtii* (northern Eurasia), *S. nigrita sibirica* (Central Asia) and *S. nigrita labradoris* (Nearctic). Lafontaine & Schmidt [22] supported this decision, but suggested that *S. nigrita zetterstedtii* was also present in the far north of Canada.
- 3) *Coenophila subrosea* in Europe and its North American sister taxon, *C. opacifrons*, show little barcode divergence (MIN = 1.29%) and prior taxonomic publications have treated them both as subspecies and as distinct species [23].
- 4) The taxonomy of *Apamea zeta* is complex. Based on deep barcode divergences (MIN = 3.0%) between populations from the Alps and North America, Zilli *et al.* [15] partitioned the *zeta* complex into four species with *A. zeta* restricted to the mountains of southern and central Europe, while *A. exulis* was found only in Canada. By contrast, Mikkola *et al.* [24] recognized five subspecies of *A. zeta* in North America,

treating *A. exulis* as one of these. They did, however, note that *A. zeta* was the most polymorphic species of *Apamea*, and that the status of its various forms and populations was uncertain. Fibiger *et al.* [14] came to a different conclusion, recognizing several Eurasian species with *A. exulis* as a circumboreal species with more than 2% barcode divergence from *A. zeta*.

High Barcode Divergences in Single Taxa – Potential Cryptic Species

Prior studies have shown that barcode divergences of more than 2% are often an indication of overlooked species or misidentifications. For example, *Gnorimoschema valesiella* and *G. alaskense* were formerly treated as synonyms, but were recently separated because of their differing genitalic morphology [25]. Barcode data support this decision as these taxa show a bit over 2% mean divergence (MIN = 1.62%) (Figure 3F). For brevity, in the balance of this section, we only discuss taxa with more than 4% maximum divergence, where the likelihood of cryptic taxa is high. However, additional overlooked species are likely among other species, especially those with maximum divergences in the 2–4% range.

Divergence Within Regions

Fennoscandia. One of the 91 species from Fennoscandia possessed a maximum divergence greater than 4% - *Rheumaptera subhastata* (MAX = 5.07%, Figure 4A). It split into two barcode clusters, both shared with alpine populations in Europe, but only one of these lineages was detected in North America.

Alps. Two of the 74 species from the Alps had a barcode divergence greater than 4%. *Rheumaptera subhastata* (MAX = 4.75%) was divided into the same two clusters found in Fennoscandia, while *Eudonia sudevica* (MAX = 6.32%, Figure 4B) included two

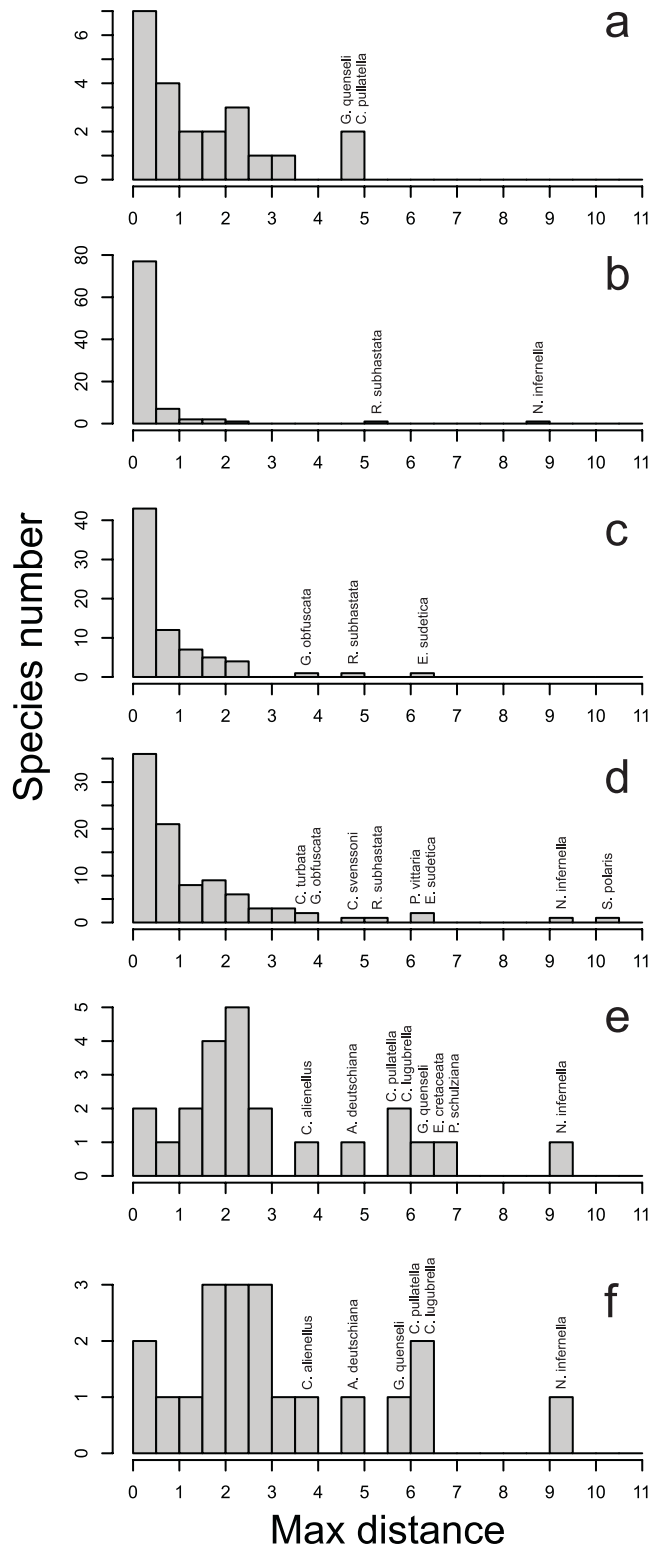


Figure 2. Distributions of maximum intraspecific barcode divergences (MAX) within and between the study regions. The x-axis shows MAX values in percentage. 2A. within North America, 2B. within Fennoscandia, 2C. within Alps, 2D. between Alps and Fennoscandia, 2E. between Alps and North America, 2F. between Fennoscandia and North America. Species showing over 4% intraspecific divergence are indicated above the bars. doi:10.1371/journal.pone.0047214.g002

barcode clusters with allopatric distributions. Most alpine specimens belonged to the same barcode cluster as those in Fennoscandia, but specimens from Central Italy belonged to a second cluster that is likely an undescribed cryptic species.

North America. Two of 22 species from North America showed more than 4% barcode divergence: *Caryocolum pullatella* (MAX = 4.56%) and *Grammia quenseli* (MAX = 4.59%, Figure 4C). *C. pullatella* has been recognized as a probable cryptic species complex (Huemer [26]) and our barcode studies revealed two barcode lineages in North America and another two in Europe. North American specimens of *G. quenseli* also split into two divergent barcode clusters distinct from those in Europe, but their status is uncertain as Schmidt [27] found that mitochondrial markers were of limited taxonomic value in *Grammia* due to frequent hybridization and introgression.

Divergence Between Regions

Fennoscandia – Alps. Four of 94 species analyzed from Fennoscandia and the Alps showed greater than 4% divergence: *Coleophora svenssoni* (MAX = 4.67%), *Synanthedon polaris* (MAX = 10.08%), *Eudonia sudetica* (MAX = 6.32%), and *Elophos vittaria* (MAX = 6.47%, Figure 4D). We expect that most, if not all, of these cases involve overlooked species. For example, intraspecific divergence is reduced below 2% in *E. sudetica* once the divergent populations from Central Italy are excluded. Specimens of *S. polaris* from the Alps show close barcode similarity to those from Sweden and Norway, but not with specimens from Finland. This case involves two previously recognized species which are currently accepted as synonyms.

Alps/Fennoscandia – North America. Five of 23 species found in both the Alps and North America showed barcode divergences greater than 4%: *Caryocolum pullatella* (MAX = 5.63%), *Chionodes lugubrella* (MAX = 5.95%), *Aethes deuschiana* (MAX = 4.93%, Figure 4E), *Grammia quenseli* (MAX = 6.12%, Figure 4C) and *Eupithecia cretacea* (MAX = 6.83%). The same pattern of deep divergences was also seen in comparisons involving these species from Fennoscandia and North America, excepting *E. cretacea* which was not collected in Fennoscandia.

Corrected Misidentifications

Our barcode studies revealed three cases of deep intraspecific barcode divergence and one case of barcode sharing that were later found to arise from misidentifications.

- 1) Specimens of *Olethreutes schulziana* from North America fell into two barcode clusters with a maximum divergence of 7.68%. Subsequent dissection and literature comparison [28,29] revealed that one of the groups was actually *O. inquietana*. After this adjustment, specimens of *O. schulziana* from North America show low divergence (MAX = 0.16%), but considerable divergence from their European counterparts (MIN = 1.55%).
- 2) Specimens of *Olethreutes septentrionana* from Finland showed marked barcode divergence from those in the Alps (MAX = 8.57%). The northern European specimens formed a novel barcode cluster, but the sole specimen from the Alps shared its barcode with *O. palustrana*. Detailed analysis, including the dissection of additional specimens and literature survey [29,30], revealed that alpine specimens identified as *O. septentrionana* are actually *O. palustrana*. *O. septentrionana* is known from montane sites in Poland [29], but it seems absent from the Alps.
- 3) *Neofaculta infernella* included two barcode clusters (MAX = 8.88%) in Fennoscandia, while just one occurred

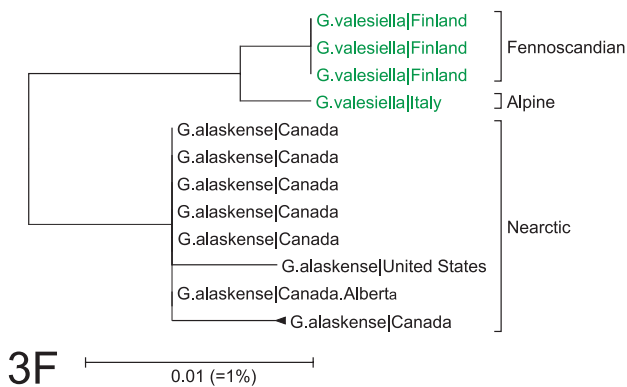
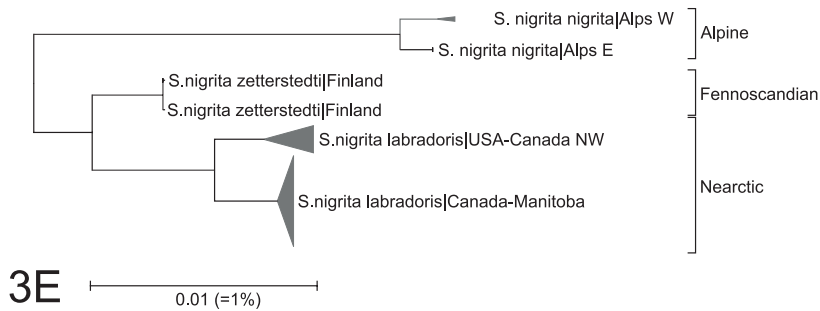
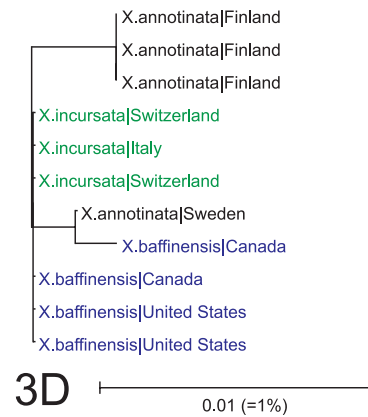
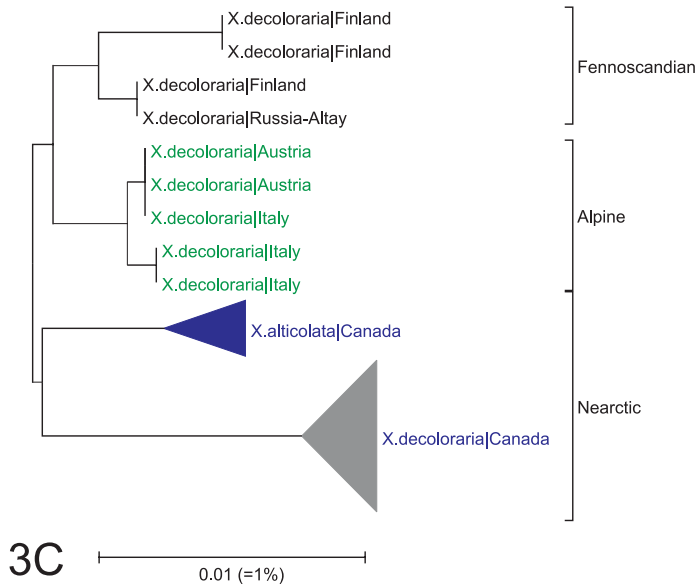
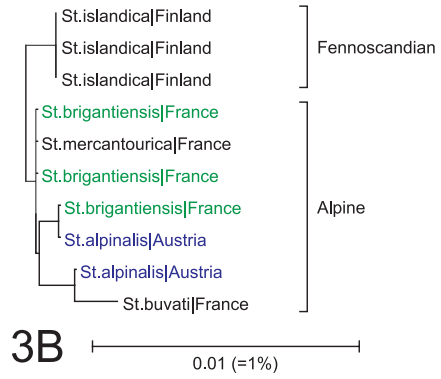
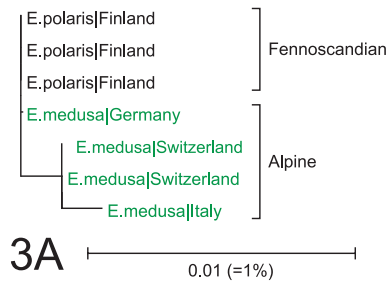


Figure 3. Examples of taxonomic findings. Taxa showing low interspecific divergence (3A. *Erebia polaris* - *E. medusa*, 3B. *Stenoptilia islandica* - *S. brigantiensis* - *S. mercantourica* - *S. alpinalis*, 3C. *Xanthorhoe decoloraria* - *X. alticolata*); taxonomic incongruence (3D. *Xanthorhoe annotinata* - *X. incurata* - *X. baffinensis*, 3E. *Sympistis nigrita* subspecies); support for the recent taxonomic revisions (3F. *Gnorimoschema valesiella* - *G. alaskense*). doi:10.1371/journal.pone.0047214.g003

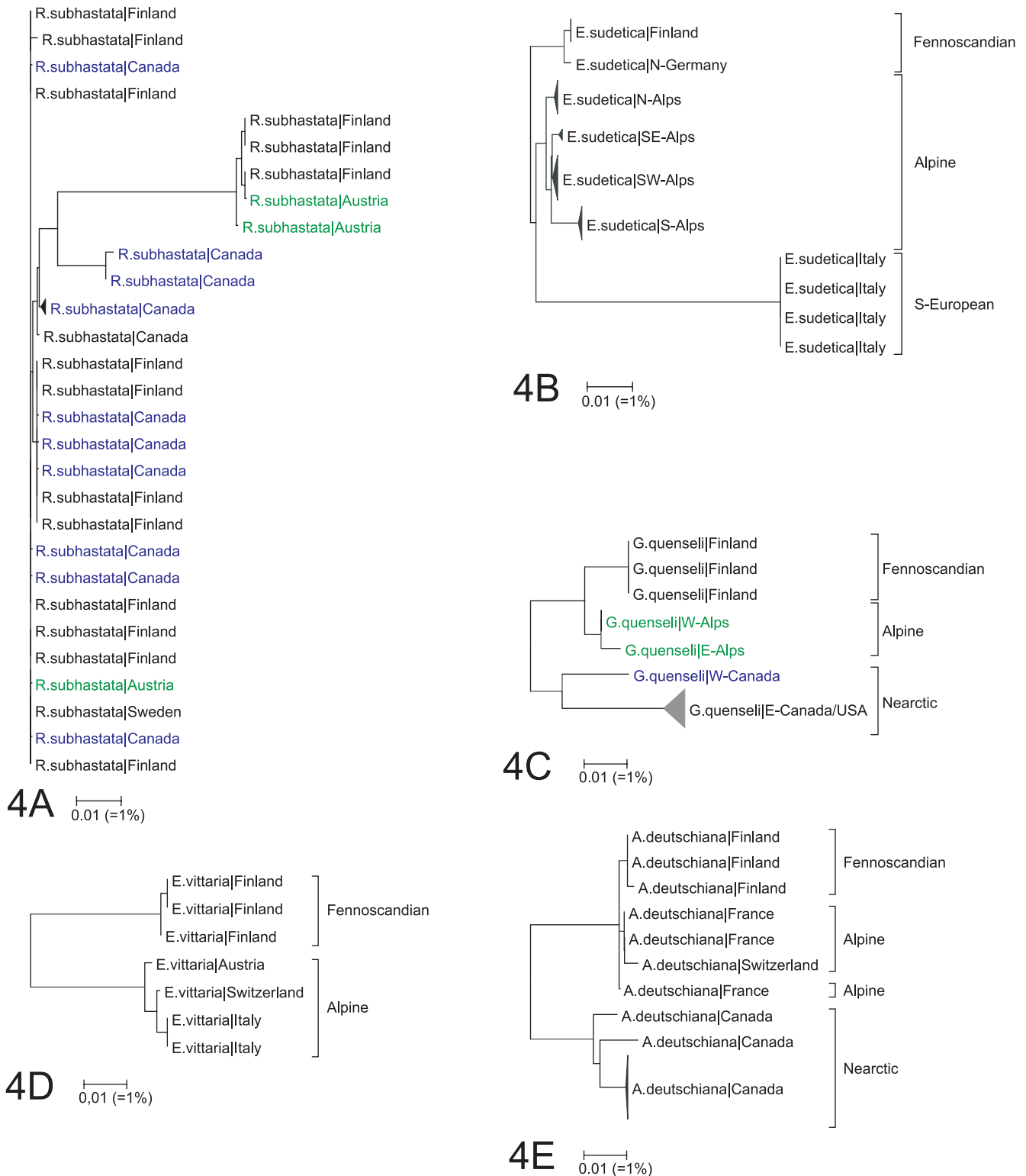


Figure 4. Species showing deep intraspecific splits within or between regions. 4A. *Rheumaptera subhastata*, 4B. *Eudonia sudetica*, 4C. *Grammia quenseli*, 4D. *Elophos vittaria*, 4E. *Aethes deutschiana*. doi:10.1371/journal.pone.0047214.g004

in the Alps. Barcoding of Siberian *N. taigana* material, including type material (unpublished records) and examination of the original description of *N. taigana* [31] suggested that specimens in the cluster from Fennoscandia shared with North America actually belong to another species - *N. taigana*.

- 4) Specimens from Finland thought to represent *Ancylis rhenana* shared identical barcodes with a specimen of *A. habeleri* from the Alps. These two species are best identified by external morphology, and based on forewing coloration and patterns [29–32] we conclude that supposed *A. rhenana* from Finland are actually *A. habeleri*.

Discussion

DNA barcoding has proven an efficient tool for both differentiating animal species and revealing cryptic diversity [12,33–37]. In groups with well-established taxonomy, such as European Lepidoptera and North American birds, identification success has been strong [12,34,38]. Most reports of low success have involved groups which were known to be taxonomically problematic [39]. Other apparent cases of failure have arisen from flawed taxonomy or from misidentification [40–42]. Our study revealed four cases of misidentification which would have led, if not corrected, to apparent barcoding failures. As our work involved one of the best known groups of insects and the resolution of each case required substantial effort, it is probable that many other cases of misidentification have been overlooked.

Few studies have examined patterns of barcode variability over large geographic areas, and fewer still have examined divergences between allopatric populations or sister species. Past studies have shown that wider geographic sampling usually increases the amount of intraspecific variation, but that it often has little effect on identification success because the incremental variation typically erodes the barcode gap rather than producing overlap between species [6]. However, there are exceptions. Broader geographic sampling reduced both identification success and the number of monophyletic species in a group of closely related beetles [43]. Further studies of this sort are needed because taxonomic complexity rises as the scale of geographic coverage expands. Species with discontinuous ranges present a particular challenge because some species concepts cannot be employed. For example, the biological species concept can only be applied to sympatric taxa because range overlap is required to assess the presence or absence of gene flow. Other species concepts, such as the differential fitness concept [2], are theoretically applicable, but would necessitate breeding experiments with allopatric populations. The phylogenetic species concept has the advantage of being easily applied to allopatric lineages as it only requires that species be diagnosable clusters of individuals with shared ancestry [44]. However, few taxonomists favour the recognition of species that can only be discriminated by genetic markers [45]. Moreover, strict application of the phylogenetic species concept would lead to a tremendous increase in species numbers [46]. For example, many of the allopatric populations that we examined possessed diagnostic barcodes, qualifying them for recognition as separate species under the phylogenetic species concept.

In practice, diversity in morphological traits underlies most taxonomic systems with genitalic characters playing a decisive role in much insect taxonomy [47], reflecting the view that such divergence acts as a prezygotic isolating mechanism [48]. Sauer & Hausdorf [49] support this priority, arguing that copulatory organs should be assigned special value in taxonomy because they are

directly involved in speciation. It has also been suggested that genitalia show rapid divergence as a result of selection against hybrids when populations with partial reproductive isolation come into secondary contact [50]. These views have often led genitalic variation to be rated as decisive in considerations of taxonomic status. For example, alpine and Fennoscandian populations of *Xestia lorezi* are considered as subspecies because they lack diagnostic genital characteristics [23,51–52] although they possess differing wing patterns and colouration, differing larval characters [53] and substantial COI divergence (MIN = 2.82%). Conversely, *Xestia rhaetica* and *X. femica* are recognized as distinct species [13–14] because of small genital differences [14,51], despite their morphological similarity and lack of barcode divergence (MIN = 0%). The arctic-alpine butterflies, *Erebia polaris* and *E. medusa*, provide another example of the subjectivity of taxonomic decisions based solely on morphology. While noting differences in size, wing colour and markings, Warren [54] treated *Erebia polaris* as a subspecies of *Erebia medusa* because of their similar genitalia, but other authors [55–56] have accepted them as valid species because of their differing external morphology. These examples not only demonstrate that morphological and genetic data do not always provide coincident signals in relation to species status, but also that there is no invariant rule for prioritizing external versus genitalic morphology in determinations of species status. There is certainly no basis for considering morphological features more “valuable” than genetic data and no reason to assume that species should always be morphologically distinct. *Neofaculta infernella* may represent a case of the latter situation as its Finnish populations include two barcode clusters with deep divergence (MIN = 8.4%), but no obvious morphological differences. If this barcode split is correlated to divergence in a nuclear marker or in ecological traits, recognition of two species should follow [45].

The idea that genitalic differentiation (or the lack of it) can indicate the presence or absence of reproductive isolation is unconvincing. For example, sperm transfer was frequently successful among species in the moth genus *Euxoa* despite their genitalic differences [57]. However, because genitalic morphology evolves rapidly [48,58] it remains taxonomically informative. Our results further demonstrate a close correspondence between genitalic and barcode divergences. For example, five of six cases of barcode sharing that we detected between Fennoscandian and alpine taxa involved species with little or no divergence in their genital morphology. Conversely, species with clearly different genitalia nearly always possessed substantial barcode divergence.

Based on their review of many case studies, Funk & Omland [59] concluded that the transfer of mitochondrial genomes was relatively common among closely related species. It now seems that their conclusion reflected the fact that species selected for genetic analysis were not chosen randomly - they often represent taxonomically complex situations. Although Funk & Omland's study thus probably overestimated the incidence of mitochondrial introgression, this complication does occur in a few arctic-alpine taxa. For example, Schmidt [27] found that Nearctic specimens of *Grammia quenseli* possessed five COI clusters with a maximum divergence of 11%. Because there was no evidence for their linkage to morphology, Schmidt & Sperling [60] concluded that hybridization and introgression were widespread in this genus. Our study revealed an apparent case of paraphyly in *Rheumaptera subhastata*. This species included two barcode clusters, one closer to *R. hastata* than to its conspecifics. However, such cases of paraphyly are rare; this is the sole example that we detected among more than 3000 barcoded species of European Lepidoptera. Before speculating further on the origin of this case, studies are needed on

nuclear markers to rule out the possibility that *R. subhastata* is actually two species.

The current subjectivity in decisions concerning the taxonomic status of allopatric populations is unsatisfactory. We believe that objectivity can be enhanced by assigning priority to molecular rather than morphological traits. Under this approach, lineages would be flagged for consideration as distinct species if their allopatric populations exceeded a threshold value (e.g. 2%) of sequence divergence in the DNA barcode region. Although no threshold can act as a species diagnostic in all situations, the same criticism applies to any other criterion. On a positive side, the use of a molecular yardstick as an initial screening tool has a primary advantage; it can be applied in a standardized way across all species. Furthermore, analytical models (ABGD [61], GMYC [62]) can provide an objective way to operationalize the flagging process. Many morphological taxonomists would impose a supplemental requirement; allopatric lineages meeting a sequence threshold should also possess diagnostic morphological difference(s) before gaining recognition as distinct species. This view places molecular characters as subservient to morphological characters since the converse situation of recognizing morphologically distinct, but genetically similar populations as separate species is established practice. A compromise solution to delineate allopatric populations might reserve species status for cases where divergence is apparent for two or more independent characters. Under this approach, species status would be granted to lineages which not only exceed the barcode threshold, but that also show correlated differentiation in any ecological or morphological trait, or in an unlinked molecular marker. The latter criterion will play an increasingly important role in aiding the recognition of species in structurally simple groups of eukaryotes [63], just as it has in bacteria, archaea and fungi.

Supporting Information

Table S1 List of specimens with sequence and collection data information. (PDF)

References

- deQueiroz K (2007) Species concepts and species delimitation. *Systematic Biology* 56: 879–886.
- Hausdorf B. (2011) Progress toward a general species concept. *Evolution* 65: 923–931.
- Mikkola K, Lafontaine JD, Kónonenko VS (1991) Zoogeography of the Holarctic species of the Noctuidae (Lepidoptera): importance of the Beringian refuge. *Entomologica Fennica* 2: 157–173.
- Schmitt T, Hewitt GM (2004) Molecular biogeography of the arctic-alpine disjunct burnet moth species *Zygaena exulans* (Zygaenidae, Lepidoptera) in the Pyrenees and Alps. *Journal of Biogeography* 31: 885–893.
- Schmitt T, Hewitt GM, Müller P (2006) Disjunct distributions during glacial and interglacial periods in mountain butterflies: *Erebia epiphron* as an example. *Journal of Evolutionary Biology* 19: 108–113.
- Lukhtanov VA, Sourakov A, Zakharov EV, Hebert PDN (2009) DNA barcoding Central Asian butterflies: increasing geographical dimension does not significantly reduce the success of species identification. *Molecular Ecology Resources* 9: 1302–1310.
- de Jong YSDM (ed.) (2011) Fauna Europaea version 2.4. Web Service available online at <http://www.faunaeur.org>. Accessed 2012 Feb 1.
- Hodges RW, Dominick T, Davis DR, Ferguson DC, Franclemont JG, et al. (eds.) (1983) Check List of the Lepidoptera of America North of Mexico (Including Greenland). EW Classey Ltd. and The Wedge Entomological Research Foundation, London, 284 pp.
- deWaard JR, Ivanova NV, Hajibabaei M, Hebert PDN (2008) Assembling DNA Barcodes: Analytical Protocols. Pp. 275–293 In: Cristofre M, ed. *Methods in Molecular Biology: Environmental Genetics*. Humana Press Inc., Totowa, USA, 364 pp.
- Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes* 7: 355–364.

Table S2 Provisional check-list of arctic-alpine and boreo-montane species shared between the Alps and Fennoscandia but not barcoded from both major distribution areas in this study; taxonomic status of several taxa needs revision. (PDF)

Table S3 Minimum, mean and maximum intraspecific variation in DNA barcodes in study species within and between the study regions. NA refers to North America, AL to Alps and FE to Fennoscandia. Species showing over 2% intraspecific divergence within or between the regions in question are shown highlighted. (PDF)

Table S4 Minimum, mean and maximum variation in DNA barcodes between allopatric species, which morphologically are considered distinct taxa, but which genetically are closely similar. NA refers to North America, AL to Alps and FE to Fennoscandia. Superscript numbers after the species refer to the areas of occurrence as follows: 1 = Alps, 2 = Fennoscandia, 3 = North America. (PDF)

Acknowledgments

We are very grateful to staff at the Canadian Centre for DNA Barcoding for sequence analysis. Josef Jaros (Ceske Budejovice, Czech Republic) and Jürg Schmid (Illanz, Switzerland) kindly contributed specimens, while Niklas Wahlberg provided valuable advice on DNA substitution rates. We also thank Bengt Å. Bengtsson, Sami Haapala, Petri Hirvonen, Juhani Itämies, Lauri Kaila, Jari Kaitila, Ali Karhu, Erkki M. Laasonen, Reima Leinonen, Tomi Mutanen, Kari Nupponen, Nils Ryrholm, Ingvar Svensson (†), Kari Vaalamo and Panu Välimäki for providing specimens from northern Europe. Two anonymous referees provided useful comments that improved the paper, for which we are grateful. Finally, we thank Sami Kivelä and Hannes Kühtreiber for preparing the figures.

Author Contributions

Conceived and designed the experiments: MM PH. Performed the experiments: MM PH. Analyzed the data: MM JFL JRdW PH. Contributed reagents/materials/analysis tools: MM AH PDNH JFL JRdW PH. Wrote the paper: MM AH PDNH JFL JRdW PH.

- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* 28: 2731–2739.
- Hausmann A, Haszprunar G, Hebert PDN (2011) DNA barcoding the geometrid fauna of Bavaria (Lepidoptera): Successes, surprises, and questions. *PLoS ONE* 6: e17134. (doi:10.1371/journal.pone.0017134).
- Kullberg J, Albrecht A, Kaila L, Varis V (2002) Checklist of Finnish Lepidoptera – Suomen perhosten luettelo. *Sahlbergia* 6: 45–190.
- Fibiger M, Ronkay L, Yela JL, Zilli A (2010) Noctuidae Europaeae. Vol. 12. Rivulinae – Phytometrinae, and Micronoctuidae, including Supplement to Noctuidae Europaeae. Vols 1–11. Entomological Press Soro, 451 pp.
- Zilli A, Ronkay L, Fibiger M (2005) Noctuidae Europaeae. Vol.8. Apameini. Entomological Press Soro, 323 pp.
- Pöyry J, Kullberg J (1997) A taxonomic revision of the genus *Holarctia* Ferguson, 1984 (Arctiidae). *Nota lepidopterologica* 20: 45–56.
- Witt T, Ronkay L (eds.) (2012) Noctuidae Europaeae. Vol. 13. Lymantriinae – Arctiinae, including Phylogeny and Check List of the Quadrifid Noctuoidea of Europe. Entomological Press Soro, 448 pp.
- Gielis C (1996) Pterophoridae. In: Hucmer P, Karsholt O, Lyneborg L, eds. *Microlepidoptera of Europe 1*. Apollo Books, Stenstrup, 222 pp.
- Arenberger E (2005) Pterophoridae 3. Teilband Platyptilinae: Platyptilini: *Stenoptilia*. In: Gaedike R, ed. *Microlepidoptera Palaearctica 12*. Goecke & Evers, Keltern, 191 pp.
- Skou P (1991) Nordens Ugler. Danmarks Dyreliv, Bind 5. Apollo Books Stenstrup, 565 pp.
- Ronkay G, Ronkay L (1995) Noctuidae Europaeae. Vol. 7. Cucullinae II. Entomological Press Soro, 224 pp.
- Lafontaine JD, Schmidt BC (2010) Annotated check list of the Noctuoidea (Insecta, Lepidoptera) of North America north of Mexico. *ZooKeys* 40: 1–239.

23. Lafontaine JD (1998) Noctuoidea Noctuidae (part) Noctuinae (part – Noctuini). In: Hodges RW, et al., eds. The Moths of America North of Mexico including Greenland, Fascicle 27.3. The Wedge Entomological Research Foundation, Washington, 348 pp.
24. Mikkola K, Lafontaine JD, Gill J (2009) Noctuoidea Noctuidae (part) Xyleninae (part) Apameini (part – *Apamea* group of genera). In: Hodges RW, et al., eds. The Moths of America North of Mexico including Greenland, Fascicle 26.9. The Wedge Entomological Research Foundation, Washington, 192 pp.
25. Huemer P, Karsholt O (2010) Gelechiidae II (Gelechiinae: Gnorimoschemini). Volume 2 In: Huemer P, Karsholt O, Nuss M, eds. Microlepidoptera of Europe 6. Apollo Books, Stenstrup, 586 pp.
26. Huemer P (1988) A taxonomic revision of *Caryocolum* (Lepidoptera, Gelechiidae). Bulletin of the British Museum (Natural History) (Entomology) 57: 439–571.
27. Schmidt BC (2009) Taxonomic revision of the genus *Grammia* Rambur (Lepidoptera: Noctuidae: Arctiinae). Zoological Journal of the Linnean Society 156: 507–597.
28. Heinrich C (1926) Revision of the North American moths of the subfamilies Laspeyresinae and Olethreutinae. Bulletin of the United States National Museum 132: 1–216.
29. Razowski J (2003) Tortricidae (Lepidoptera) of Europe. Volume 2. Olethreutinae. Bratislava, 301 pp.
30. Svensson I (2006) Nordens vecklare, The Nordic Tortricidae. Kristianstads Boktryckeri AB, Lund, 349 pp.
31. Ponomarenko MG (1998) New taxonomic data on Dichomeridinae (Lepidoptera, Gelechiidae) from the Russian Far East. Far Eastern Entomologist 67: 1–17.
32. Huemer P, Tarmann G (1997) Die *Ancylys badiana* (Denis & Schiffermüller), 1775)-Gruppe in Europa: ein verkannter Kleinschmetterlings-Artenkomplex (Lepidoptera, Tortricidae). Veröff. tirol. Landesmus. Ferdinandeum 77: 203–222.
33. Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN (2006) DNA barcodes distinguish species of tropical Lepidoptera. Proceedings of the National Academy of Sciences USA 103: 968–971.
34. Kerr KCR, Stoeckle MY, Dove CJ, Weigt LA, Francis CM, et al. (2007) Comprehensive DNA barcode coverage of North American birds. Molecular Ecology Notes 7: 535–543.
35. Footitt RG, Maw HEL, Von Dohlen CD, Hebert PDN (2008) Species identification of aphids (Insecta: Hemiptera: Aphididae) through DNA barcodes. Molecular Ecology Resources 8: 1189–1201.
36. Hebert PDN, deWaard JR, Landry JF (2009) DNA barcodes for 1/1000 of the animal kingdom. Biology Letters 6: 359–362.
37. Janzen DH, Hallwachs W, Blandin P, Burns JM, Cadiou JM, et al. (2009) Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity. Molecular Ecology Resources 9(s1): 1–26.
38. Dinca V, Zakharov EV, Hebert PDN, Vila R (2010) Complete DNA barcode reference library for a country's butterfly fauna reveals high performance for temperate Europe. Proceedings of the Royal Society of London B 278: 347–355.
39. Wiemers M, Fiedler K (2007) Does the DNA barcoding gap exist? – a case study in blue butterflies (Lepidoptera: Lycaenidae). Frontiers of Zoology 4: 8.
40. Harris JD (2003) Can you bank on GenBank? Trends in Ecology & Evolution 18: 317–319.
41. Nilsson RH, Ryberg M, Kristiansson E, Abarenkov K, Larsson KH, et al. (2006) Taxonomic reliability of DNA sequences in public sequence databases: a fungal perspective. PLoS ONE 1 (doi:10.1371/journal.pone.0000059).
42. Mitchell A (2008) DNA barcoding demystified. Australian Journal of Entomology 47: 169–173.
43. Bergsten J, Bilton DT, Fujisawa T, Elliott M, Monaghan MT, et al. (2012) The effect of geographic scale of sampling on DNA barcoding. Systematic Biology (doi: 10.1093/sysbio/sys037).
44. Eldredge N, Cracraft J (1980) Phylogenetic patterns and the evolutionary process. Columbia University Press, New York, 349 pp.
45. Cook LG, Edwards RD, Crisp MD, Hardy NB (2010) Need morphology always be required for new species descriptions? Invertebrate Systematics 24: 322–326.
46. Isaac NJB, Mallet J, Mace GM (2004) Taxonomic inflation: its influence on macroecology and conservation. Trends in Ecology & Evolution 19: 464–469.
47. Mutanen M (2005) Delimitation difficulties in species splits: a morphometric case study on the *Euxoa tritici* complex (Lepidoptera: Noctuidae). Systematic Entomology 30: 632–643.
48. Eberhard WG (1985) Sexual selection and animal genitalia. Harvard University Press, Cambridge MA, 244 pp.
49. Sauer J, Hausdorf B (2011) A comparison of DNA-based methods for delimiting species in a Cretan land snail radiation reveals shortcomings of exclusively molecular taxonomy. Cladistics (doi: 10.1111/j.1096-0031.2011.00382.x).
50. Shapiro AM, Porter AH (1989) The lock-and-key hypothesis: evolutionary and biosystematic interpretation of insect genitalia. Annual Review of Entomology 34: 231–245.
51. Fibiger M (1993) Noctuidae Europaeae. Vol. 2. Noctuinae II. Entomological Press Soro, 230 pp.
52. Lafontaine JD, Mikkola K, Kononenko VS (1987) A revision of the genus *Xestia* subg. *Pachnobia* (Lepidoptera: Noctuidae), with descriptions of two new subspecies. Entomologica Scandinavica 18: 305–331.
53. Ahola M, Lafontaine JD (1990) Larvae of *Xestia kolymae* (Herz) and *Xestia lorezi* (Staudinger) (Lepidoptera: Noctuidae), with notes on the geographical variation in the latter. Entomologica Scandinavica 21: 77–90.
54. Warren BCS (1936) Monograph of the genus *Erebia*. British Museum (Natural History), 407 pp., 107 pls.
55. Higgins LG (1975) The Classification of European Butterflies. Collins, London, 320 pp.
56. Eliasson CU, Ryrholm N, Holmer M, Jilg K, Gärdenfors U (2005) Nationalhyckeln till Sveriges flora och fauna. Fjärilar: Dagfjärilar. Hesperidae – Nymphalidae. ArtDatabanken, SLU, Uppsala.
57. Byers JR, Hinks CF (1978) Biosystematics of the genus *Euxoa* (Lepidoptera: Noctuidae) XI. Mating discrimination between three closely related species of the *declarata* group. Canadian Journal of Zoology 56: 1981–1987.
58. Arnqvist G (1998) Comparative evidence for the evolution of genitalia by sexual selection. Nature 393: 784–786.
59. Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology, Evolution, and Systematics 34: 397–423.
60. Schmidt BC, Sperling FAH (2008) Widespread decoupling of mtDNA variation and species integrity in *Grammia* tiger moths (Lepidoptera: Noctuidae). Systematic Entomology 33: 613–634.
61. Puillandre N, Lambert A, Brouillet S, Achaz G (2011) ABGD, Automatic barcode gap discovery for primary species delimitation. Molecular Ecology 21: 1864–1877.
62. Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, et al. (2006) Sequence based species delimitation for the DNA taxonomy of undescribed insects. Systematic Biology 55: 595–609.
63. Blaxter M, Mann J, Chapman T, Thomas F, Whitton C, et al. (2005) Defining operational taxonomic units using DNA barcode data. Philosophical Transactions of the Royal Society B 360: 1935–1943.