Now DNA-barcoded: the butterflies and larger moths of Germany

(Lepidoptera: Rhopalocera, Macroheterocera)

Axel Hausmann, Gerhard Haszprunar, Andreas H. Segerer, Wolfgang Speidel, Gottfried Behounek & Paul D. N. Hebert


This study provides a DNA barcode library for 1264 of the 1338 species of butterflies and larger moths (Rhopalocera and Macroheterocera) of Germany. These results arise from a research program established by the State of Bavaria which is constructing a DNA barcode library for all animal species within its territorial boundaries. Open access is provided to a data set that includes records for 3467 specimens (957 species) from Germany. An additional 307 species of the German fauna are represented by barcode data specimens collected in other European nations. Most (99 %) of the 957 species from the study area were found to possess diagnostic barcode sequences. A few taxa which apparently share DNA barcodes are discussed in detail. Deep intraspecific sequence divergences (> 2 %) were detected in 32 traditionally recognized species which are undergoing more detailed analysis to ascertain whether they represent cases of cryptic diversity. The study reinforces the effectiveness of DNA barcoding as a tool for monitoring biodiversity and for other applications reliant on species identification.

Axel Hausmann (corresponding author), Gerhard Haszprunar, Andreas H. Segerer & Wolfgang Speidel: Bavarian State Collection of Zoology, Münchhausenstr. 21, 81247 Munich, Germany; e-mail: Axel.Hausmann@zsm.mwn.de

Gottfried Behounek, Sudetenstr. 6, 85567 Grafing, Germany

Paul D. N. Hebert, Biodiversity Institute of Ontario, University of Guelph, Guelph, ON, N1G 2W1, Canada

Introduction

Although situated in the centre of Europe, Germany still lacks a comprehensive faunistic monograph for its more than 3600 Lepidoptera species. An updated faunal list can be inferred through the German faunal list (Gaedike & Heinicke 1999), but such checklists are no substitute for faunistic assessments. Furthermore, the German butterfly and ‘macromoth’ fauna includes many unresolved and controversial taxonomic questions. DNA barcoding offers a rapid, cost-effective alternative strategy for both the identification of described species and the discovery of new ones (Hebert et al. 2003, Savolainen et al. 2005, Mitchell 2008). For that reason the authors activated the ‘Barcoding Fauna Bavarica (BFB)’ project enabled by a 5-year grant from Bavarian State Government (Haszprunar 2009). Coordination of this project is mainly performed by curators at the ZSM and benefits from the involvement of a strong network of private collectors and entomological associations, Nature Reserve authorities and other cooperating institutions. The DNA Bank facility at the ZSM (see www.zsm.mwn.de/dnabank/ and Gemeinholzer et al. (2011) for details) holds extracts of barcoded specimens for spin-off research projects. Research activities involve close cooperation with the Biodiversity Institute of Ontario, which performs the sequence
analyses under the framework of the International Barcode of Life Project (iBOL). The BFB project represents the first effort to create a DNA barcode library for all animal species in a whole country. The project seeks to achieve coverage for at least 10,000 species by 2013 (Haszprunar 2009) out of the 45,000 Metazoa species known from Germany (Völkl & Blick 2004), of which 85% are Arthropoda (Haszprunar 2009).

By March 2011, barcode records were available for more than 14,000 specimens representing more than 5700 species (Balke et al. 2011), of which about 2250 are Lepidoptera (cf. Hausmann et al. 2011b, slightly updated). Thus, barcode coverage for the German fauna is currently at 13 per cent.

Reflecting a long entomological tradition starting with German lepidopterists like Hufnagel, Hübner, Esper and Herrich-Schäffer, the macrolepidopteran fauna of Germany is generally thought to be well known faunistically and taxonomically. This view has been tested recently by the assembly of a DNA library for Bavarian geometrids (Hausmann et al. 2011a), which provided an additional character set to challenge existing taxonomic concepts (species delimitations, possible synonymies, etc.) in an integrated taxonomic approach. In about 5% of the species, deep sequence divergences were detected reflecting cases where further research may reveal cryptic species. In the present paper, the DNA library is enlarged three-fold.

This paper provides open access to our data on German butterflies and larger moths (Rhopalocera and Macroheterocera). The sharing of sequence data accompanied by georeferenced information and images of its source specimen in public BOLD projects and on GenBank represent an important contribution to the democratization of biodiversity information (Ratnasingham & Hebert 2007, Janzen et al. 2009, Hebert et al. 2010).

**Materials and methods**

**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZSM</td>
<td>Zoological Collection of the State of Bavaria, Munich</td>
</tr>
<tr>
<td>CCDB</td>
<td>Canadian Centre for DNA Barcoding</td>
</tr>
<tr>
<td>BOLD</td>
<td>Barcode of Life Data System</td>
</tr>
<tr>
<td>iBOL</td>
<td>International Barcode of Life project</td>
</tr>
<tr>
<td>BFB</td>
<td>Barcoding Fauna Bavarica</td>
</tr>
</tbody>
</table>

**Sampling**

Taxon sampling was restricted to the 'obtectomeran' Lepidoptera (Nieukerken et al. 2011) excluding Thyrididae and Pyraloidea, i.e. all Rhopalocera and all Macroheterocera. DNA barcodes were obtained by sampling dry legs from specimens in the ZSM with additional material from some private collections (see acknowledgements). Since the projects arise from the 'Barcoding Fauna Bavarica' project, sampling was restricted to a few individuals per species, trying to include material from all four major Bavarian fauna regions as defined in Voith (2004). In a second step, sampling was extended beyond the boundaries of Bavaria, but the whole sampling program remained focused on southern Germany. Other studies have established that geographically focused studies provide a good basis for the construction of barcode libraries that allow the re-identification of species across a much broader area (Lukhtanov et al. 2009). By early 2011, tissue samples from 4208 German Rhopalocera and Macroheterocera had been submitted for DNA barcoding. All specimens were identified by the authors, with AH being responsible for Geometridae, AHS for Rhopalocera, GB for Noctuidae and Erebiidae excl. Arctiinae, Notodontinae and Lymantriinae, WS for the rest. Detail analyses and dissections of the vouchers were made in all difficult cases. Taxonomy and nomenclature (see Appendix S1) is based on the latest version of the Fauna of Europe database (Fauna Europaea 2011), barring a few updates reflecting more recent taxonomic decisions, e.g. taxa raised to species rank, new synonymies and new combinations. Higher-level taxonomy follows a recent paper that provides a consensus supported by leading Lepidoptera taxonomists and phylogeneticists (Nieukerken et al. 2011). A website has been established for 'Barcoding Fauna Germanica' (Balke et al. 2011) that continuously updates project progress, such as lists of species that lack barcode coverage.

At present, the Barcode of Life Data System (BOLD) includes barcodes for 1264 of the 1338 species of German Rhopalocera and Macroheterocera. Although 74 species (5.5%) are missing, 28 of these (2.1%) are under analysis raising coverage of the German fauna to 96.6%. The species which lack coverage are noted in Appendix S1 and a full list can be downloaded from the BFB website www.faunabavarica.de/taxa1-1.

**DNA analysis**

PCR amplification and DNA sequencing was performed at the CCDB following standard high-throughput protocols (Ivanova et al. 2006, deWaard et al. 2008), that can be accessed under http://www.dnabarcoding.ca/pa/ge/research/protocols. PCR amplification with a single pair of primers consistently recovered a 658 bp region near the 5' terminus of the mitochondrial cytochrome c oxidase I (COI) gene that included the standard 648 bp barcode region for the animal kingdom (Hebert et al. 2003). All barcoded voucher specimens are listed in Appendix S1. A DNA extract from each specimen is stored at the CCDB following standard high-throughput protocols. All sequences were deposited in GenBank according to the iBOL data release policy, and accession numbers are provided in Appendix S2. Complete specimen data including images, voucher deposition, GenBank accession numbers, GPS coordinates, sequence and trace files can easily be accessed in the Barcode of Life Data System.
two sections on species. These species are discussed in detail in the (1.2 %) shared barcode sequences with another classic entomological approaches. Only 15 species differences from all other species recognized through maximum pairwise distances. While intraspecific variation is reported as maximum pairwise distances.

Data analysis

Sequence divergences for the barcode region were calculated using the Kimura 2 Parameter model, employing the analytical tools on BOLD. Genetic distances between species are reported as minimum pairwise distances, while intraspecific variation is reported as maximum pairwise distances.

Results

Sequence records were obtained from 3467 specimens representing 957 species. Most sequences (3261) were longer than 500 bp, meeting the length requirement for barcode status (Ratnasingham & Hebert 2007). A 658 bp record was obtained from 2574 specimens (907 species). Supplemental data were included for 307 species (see Appendix S1) which are part of the German fauna, but which are currently only represented by specimens from other localities in Europe.

Genetic distances between species

Our data indicate that COI barcode sequences are diagnostic for nearly 99 % of the species of Rhopalocera and Macroheterocera known from Germany. ‘Diagnostic’ barcode clusters show constant differences from all other species recognized through classic entomological approaches. Only 15 species (1.2 %) shared barcode sequences with another species. These species are discussed in detail in the two sections on ‘barcode sharing’, and ‘overlapping barcodes’. Another 32 species showed deep intraspecific divergences (see below), but none of these cases can lead to misidentifications because the component sequence clusters in each species are distinct from those of any other taxa and are always nearest neighbours.

Forty-nine species pairs (Table 1) showed low divergence values from 1 % to 3 %. Among these cases, twelve species showed low divergence from two other species and one showed low divergence to four other taxa. Thus, Table 1 contains 83 species that possess low divergence of 1 % to 3 % from one or more other taxa. The three sections on ‘barcode sharing’, ‘overlapping barcodes’ and ‘low divergences under 1 %’ involve a further 30 species with similar or identical barcodes to their nearest neighbours (32 species minus two listed also in Table 1). Thus, 112 (8.9 %) of currently recognized German Rhopalocera and Macroheterocera species showed less than 3 % sequence divergence from their nearest neighbour, 66 (5.2 %) less than 2 %, and 32 (2.5 %) less than 1 % (these values include extraterritorial data and barcodes from 500–658 bp). If the analysis is restricted to 658 bp barcodes from the study area (Figure 1), these numbers change to 101 (11.2 %), 55 (6.1 %), and 23 (2.6 %), for the 3 %, 2 % and 1 % thresholds respectively.

German Rhopalocera and Macroheterocera species show a mean genetic distance of 11.9 per cent among species within the same family (SE = 0.003; n = 830,382 comparisons in the analysis of full-length barcodes), while congeneric species average 9.0 per cent divergence (SE = 0.016; n = 30,778 comparisons in the analysis of full-length barcodes).

Discrimination of species pairs with very similar morphology

DNA barcodes allow the unambiguous identification of seven species pairs whose discrimination by morphology is challenging:

- Leptidea sinapis (Linnaeus, 1758) – L. reali Reissinger, 1989 (Pieridae): Minimum Pairwise Distance 3.0 %; These species cannot be identified with certainty based on characters of wing pattern. The genitalia of both sexes are usually morphometrically distinct, but some individuals (especially males) cannot be discriminated unambiguously (Neumayr & Segerer 1995, Segerer 2001).

- Pontia edusa (Fabricius, 1777) – P. daplidice (Linnaeus, 1758) (Pieridae): Minimum Pairwise Distance 8.1 %. The sequence of P. daplidice was based on specimens from Spain (Roger Vila pers. comm.) and from the Middle East and requires confirmation for central European populations. These species cannot be distinguished basing on wing coloration or pattern, but allozyme markers and slight differences in male genitalia led to their taxonomical discrimination. Genitalic differences, however, are not constant, and introgression has been reported from hybrid zones in Italy. As a consequence, their taxonomic status as well as the occurrence of P. daplidice in Germany remains controversial (Geiger & Scholl 1982, Geiger et al. 1988, Ebert & Rennwald 1991a, Reinhardt 1992, 1995, Adam et al. 1997, Gaedike & Heinicke 1999).

- Colias hyale (Linnaeus, 1758) – C. alfacariensis Ribbe, 1905 (Pieridae): Minimum Pairwise Distance 1.9 %, inferred from extralimital data for C. alfacariensis. Adults of these species cannot be morphologically separated with certainty even by specialists and there are no significant genitalic differences. However, the larvae are very distinct, and the species, though
Table 1. Forty-nine species pairs of German Rhopalocera and Macroheterocera species with a minimum pairwise distance (K2P) in the range from 1.0 to 3.0%. Fifteen other species pairs/triplets with less than 1% sequence divergence are discussed in the text. Forty-seven of divergence values are based on the comparison of 658 bp sequences. The two exceptions (*) involve comparisons of 619 and 633 bp sequences. In one case (**) the analysis was performed with extraterritorial data. min p.d. = minimum pairwise distance.

<table>
<thead>
<tr>
<th>species 1</th>
<th>species 2</th>
<th>min p.d. [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thera variata (Denis &amp; Schiffermüller, 1775)</td>
<td>Thera britannica (Turner, 1925)</td>
<td>1.1</td>
</tr>
<tr>
<td>Polyommatus bellargus (Rottemburg, 1775)</td>
<td>Polyommatus coridon (Poda, 1761)</td>
<td>1.2</td>
</tr>
<tr>
<td>Autographa buraetica (Staudinger, 1892)</td>
<td>Autographa pulchella (Haworth, 1802)</td>
<td>1.2</td>
</tr>
<tr>
<td>Rhyacia latens (Hübner, 1809)</td>
<td>Rhyacia griseescens (Hübner, 1794)</td>
<td>1.2</td>
</tr>
<tr>
<td>Eupithecia absinhtiata (Clerck, 1759)</td>
<td>Eupithecia goossensitata Mabille, 1869</td>
<td>1.4</td>
</tr>
<tr>
<td>Chloroclysta sitarta (Hufnagel, 1767)</td>
<td>Chloroclysta miata (Linnaeus, 1758)</td>
<td>1.5</td>
</tr>
<tr>
<td>Oligia dubia (Heydemann, 1942)</td>
<td>Oligia versicolor (Borkhausen, 1792)</td>
<td>1.6</td>
</tr>
<tr>
<td>Oligia strigilis (Linnaeus, 1758)</td>
<td>Oligia versicolor (Borkhausen, 1792)</td>
<td>1.6</td>
</tr>
<tr>
<td>Euxoa aquilina (Denis &amp; Schiffermüller, 1775)</td>
<td>Euxoa obelisca (Denis &amp; Schiffermüller, 1775)</td>
<td>1.6</td>
</tr>
<tr>
<td>Eriogaster arbuscula (Freyer, 1849) **</td>
<td>Eriogaster lanestris (Linnaeus, 1758)</td>
<td>1.6</td>
</tr>
<tr>
<td>Pyrgus cataculae (Rambur, 1839) *</td>
<td>Pyrgus andromedae (Wallengren, 1853)</td>
<td>1.7</td>
</tr>
<tr>
<td>Phengaris arion (Linnaeus, 1758)</td>
<td>Phengaris teleius (Bergsträsser, 1779)</td>
<td>1.7</td>
</tr>
<tr>
<td>Boloria aquilonaris (Stichel, 1908)</td>
<td>Boloria napaea (Hofmannsegg, 1804)</td>
<td>1.7</td>
</tr>
<tr>
<td>Orthosia gracilis (Denis &amp; Schiffermüller, 1775)</td>
<td>Orthosia opima (Hübner, 1809)</td>
<td>1.7</td>
</tr>
<tr>
<td>Apamea zeta (Treitschke, 1825)</td>
<td>Apamea maillardi (Geyer, 1834)</td>
<td>1.9</td>
</tr>
<tr>
<td>Euxoa decora (Denis &amp; Schiffermüller, 1775)</td>
<td>Euxoa obelisca (Denis &amp; Schiffermüller, 1775)</td>
<td>1.9</td>
</tr>
<tr>
<td>Euxoa nigricans (Linnaeus, 1761)</td>
<td>Euxoa obelisca (Denis &amp; Schiffermüller, 1775)</td>
<td>1.9</td>
</tr>
<tr>
<td>Nola cristata (Hübner, 1793)</td>
<td>Nola subchlamydula Staudinger, 1870</td>
<td>1.9</td>
</tr>
<tr>
<td>Eupithecia semigrappatha Bruand, 1850</td>
<td>Eupithecia imprivata (Hübner, 1813)</td>
<td>1.9</td>
</tr>
<tr>
<td>Arcia aegistis (Denis &amp; Schiffermüller, 1775)</td>
<td>Arcia artaxerxes (Fabricius, 1793)</td>
<td>1.9</td>
</tr>
<tr>
<td>Apamea rubricera (Treitschke, 1825)</td>
<td>Apamea maillardi (Geyer, 1834)</td>
<td>2.0</td>
</tr>
<tr>
<td>Lacanobia suasa (Denis &amp; Schiffermüller, 1775)</td>
<td>Lacanobia thalassina (Hufnagel, 1766)</td>
<td>2.0</td>
</tr>
<tr>
<td>Thera britannica (Turner, 1925)</td>
<td>Thera vetustata (Denis &amp; Schiffermüller, 1775)</td>
<td>2.2</td>
</tr>
<tr>
<td>Xestia c-nigrum (Linnaeus, 1758)</td>
<td>Xestia ashworthii (Doubleday, 1855)</td>
<td>2.3</td>
</tr>
<tr>
<td>Eupithecia subumbbrata (Denis &amp; Schiffermüller, 1775)</td>
<td>Eupithecia orphnata Petersen, 1909</td>
<td>2.3</td>
</tr>
<tr>
<td>Pieris napi (Linnaeus, 1758)</td>
<td>Pieris bryoniae (Hübner, 1806)</td>
<td>2.3</td>
</tr>
<tr>
<td>Pyrgus armoricanus (Oberthür, 1810)</td>
<td>Pyrgus alceae (Hübner, 1803)</td>
<td>2.3</td>
</tr>
<tr>
<td>Thera variata (Denis &amp; Schiffermüller, 1775)</td>
<td>Thera vetustata (Denis &amp; Schiffermüller, 1775)</td>
<td>2.4</td>
</tr>
<tr>
<td>Ennomos quercinaria (Hufnagel, 1817)</td>
<td>Ennomos autumnaria (Werneburg, 1859)</td>
<td>2.5</td>
</tr>
<tr>
<td>Colias palaeno (Linnaeus, 1761)</td>
<td>Colias philomone (Esper, 1780)</td>
<td>2.5</td>
</tr>
<tr>
<td>Phengaris alcon (Denis &amp; Schiffermüller, 1775)</td>
<td>Phengaris arion (Linnaeus, 1758)</td>
<td>2.5</td>
</tr>
<tr>
<td>Pyrgus andromedae (Wallengren, 1853)</td>
<td>Pyrgus carthami (Hübner, 1813)</td>
<td>2.5</td>
</tr>
<tr>
<td>Euxoa recussa (Hübner, 1817)</td>
<td>Euxoa obelisca (Denis &amp; Schiffermüller, 1775)</td>
<td>2.5</td>
</tr>
<tr>
<td>Entephria nobilisaria (Herrich-Schäffer, 1852)</td>
<td>Entephria flavata (Osthelder, 1869)</td>
<td>2.7</td>
</tr>
<tr>
<td>Eupithecia linaria (Denis &amp; Schiffermüller, 1775)</td>
<td>Eupithecia pulchellata Stephens, 1831</td>
<td>2.7</td>
</tr>
<tr>
<td>Melitaea athalia (Rottemburg, 1775)</td>
<td>Melitaea britomartis Assmann, 1847</td>
<td>2.7</td>
</tr>
<tr>
<td>Coenonympha arcania (Linnaeus, 1761)</td>
<td>Coenonympha hero (Linnaeus, 1761)</td>
<td>2.7</td>
</tr>
<tr>
<td>Eilema lutarella (Linnaeus, 1761)</td>
<td>Setema cereola (Hübner, 1803)</td>
<td>2.7</td>
</tr>
<tr>
<td>Furcula bifida (Brahm, 1877)</td>
<td>Furcula bicuspis (Borkhausen, 1790)</td>
<td>2.7</td>
</tr>
<tr>
<td>Noctua janthe (Borkhausen, 1794)</td>
<td>Noctua janthina Denis &amp; Schiffermüller, 1775</td>
<td>2.7</td>
</tr>
<tr>
<td>Furcula bifida (Brahm, 1877)</td>
<td>Furcula furcula (Clerck, 1759)</td>
<td>2.8</td>
</tr>
<tr>
<td>Apamea monoglypha (Hufnagel, 1766)</td>
<td>Apamea sublustris (Esper, 1878)</td>
<td>2.8</td>
</tr>
<tr>
<td>Caradrina clavipalpis (Scopoli, 1763)</td>
<td>Caradrina selini (Boisdouval, 1840)</td>
<td>2.8</td>
</tr>
<tr>
<td>Cucullia absinthii (Linnaeus, 1761)</td>
<td>Cucullia artemisiae (Hufnagel, 1766)</td>
<td>2.8</td>
</tr>
<tr>
<td>Xestia triangular (Hufnagel, 1766)</td>
<td>Xestia c-nigrum (Linnaeus, 1758)</td>
<td>2.97</td>
</tr>
<tr>
<td>Furcula bicuspis (Borkhausen, 1790)</td>
<td>Furcula furcula (Clerck, 1759)</td>
<td>2.97</td>
</tr>
<tr>
<td>Leptidea sinapis (Linnaeus, 1758)</td>
<td>Leptidea reali Reissinger, 1990</td>
<td>2.97</td>
</tr>
<tr>
<td>Lasiosmata megera (Linnaeus, 1758)</td>
<td>Lasiosmata petropolitana (Fabricius, 1787)</td>
<td>2.97</td>
</tr>
<tr>
<td>Caradrina morpheus (Hufnagel, 1766)</td>
<td>Caradrina selini (Boisdouval, 1840)</td>
<td>2.97</td>
</tr>
<tr>
<td>All other species</td>
<td>All other species</td>
<td>&gt;3</td>
</tr>
</tbody>
</table>
frequently occurring syntopically, have different foodplants and ecological requirements (Weidemann 1986, Ebert & Rennwald 1991a).

_Aricia agestis_ (Denis & Schiffermüller, 1775) – _A. artaxerxes_ (Fabricius, 1793) (Lycaenidae): Minimum Pairwise Distance 1.9%. These two species were shown to be genetically and ecologically distinct, but due to overlapping morphological traits, previous records are not always reliable (Høegh-Guldberg 1966, Kames 1976, Ebert & Rennwald 1991b, Aagaard et al. 2002). DNA barcoding seems to allow their reliable discrimination, but detailed studies of populations at the contact zones for these species are in progress.

_Spilosoma lubricipeda_ (Linnaeus, 1758) – _S. urticae_ (Esper, 1789) (Erebidae/Arctiinae): Minimum Pairwise Distance 5.4 %. Specimens of _S. lubricipeda_ with few dark markings are easily confused with _S. urticae_, especially when the specimens are not fresh.

_Autographa buraetica_ (Staudinger, 1892) – _A. pulchrina_ (Haworth, 1802): Minimum Pairwise Distance 1.2 %. Individuals of _A. buraetica_ can be recognized by their lack of a reddish tinge in the wing coloration, but worn specimens may need genitalic dissection for identification.

_Noctua janthina_ (Denis & Schiffermüller, 1775) – _N. janthe_ (Borkhausen, 1792) (Noctuidae): Minimum Pairwise Distance 2.7 %. Although these species can usually be separated by minor differences in their hindwing pattern, some individuals with intermediate patterns are difficult to identify.

Other taxa:

Aside from these seven species pairs, there are other cases where barcode results aid species identification. For example, most species in the taxonomically difficult genus _Euxoa_ (Noctuidae) show a minimum pairwise difference from their nearest neighbour of 1.5–2.5 %. Since identifications based on morphology and genitalia are not always reliable, DNA barcoding improves the situation considerably although certain problematic cases are discussed in the sections on ‘low divergences below 1 %’ and ‘introgression’. Hausmann et al. (2011a) discuss several additional cases where DNA barcodes allow the separation of geometrid species with little morphological divergence. For instance, worn specimens of most _Eupithecia_ species (65 species in Germany) cannot be identified by their external appearance, but members of different species show substantial barcode divergence.
Barcode results also provide insights into cases where the close affinity of species has not been recognized. For example, the arctines *Eilema lutarella* (Linnaeus, 1758) and *Setema cereola* (Hübner, 1803) show a sequence divergence of just 2.7% despite their current placement in different genera (Fauna Europaea 2011). Work is in progress to further probe their affinities.

**Barcode sharing**

Very few species of Lepidoptera in Germany were found to regularly share the same barcode sequence with another morphologically distinct species. These cases involved four species pairs:

- *Colias croceus* (Fourcroy, 1785) – *C. erate* (Esper, 1805) (Pieridae): The morphological discrimination of these species is sometimes challenging, especially in orange male and whitish female forms of *C. erate*. Bavarian specimens of *C. croceus* are genetically identical with *C. erate* from other countries, suggesting that these taxa share barcodes. Although no Bavarian specimen of *C. erate* was barcoded, we expect barcode sharing in the study area.

- *Phyllodesma ilicifolia* (Linnaeus, 1758) – *P. tremulifolia* (Hübner, 1810) (Lasiocampidae): Most specimens of this species pair can be discriminated unambiguously by slight differential features in wing pattern and genitalia, but many specimens are misidentified in collections (Pro Natura 2000). There are further differences in morphology of larval stages and ecological niche. The proper identification of specimens examined in this study was confirmed by a specialist (V. Zolotuhin). Barcode sharing between these two species was also observed in several other European countries.

- Two additional cases of apparent barcode sharing in the genera *Chlorissa*, *Lycia* (Geometridae) were discussed in a previous paper (Hausmann et al. 2011a).

**Synonymy**

Populations of *‘Mythimna scirpi* (Duponchel, 1836)* from the Danubian valley are traditionally identified as this species (Schmid 1892: 38), but Hacker et al. (2002) proposed that this species is a synonym of *Mythimna sicula* (Treitschke, 1835). This conclusion is supported by the present study because the Danubian specimens share the same barcode as populations from the rest of Bavaria. One additional case of potential synonymy involving species in the genus *Thera* (Geometridae) was discussed in a previous paper (Hausmann et al. 2011a).

**Overlapping barcodes**

Overlapping barcodes were detected in three species groups:

The three German species in the genus *Setina* (Erebidae/Arctiinae) seem to have partly overlapping barcodes that may reflect introgression as noted in other high Alpine species complexes (Hausmann et al. 2011a). Introgression has been suggested for this group based on the presence of morphologically intermediate individuals (Trawöger 1991). Further study is required to investigate this case which is based on limited data from Germany and additional records from the ‘Lepidoptera of the Alps’ project (P. Huemer pers. comm.).

Two cases of supposed barcode overlap in the Geometridae (genera *Perizoma*, *Sciadia*) were discussed in a previous paper (Hausmann et al. 2011a).

**Different species with minimum pairwise distances under 1%**

Cases of low sequence divergence (< 1%) were detected in eight species pairs or triads, but there was no evidence for sequence sharing:

- *Erebia euryale* (Esper, 1805) – *E. ligea* (Linnaeus, 1761): This species pair possesses quite clear-cut differences in external appearance coupled with a constant barcode divergence of 0.8%.

- *Plebejus argyrognomon* (Bergsträsser, 1779) – *P. idas* (Linnaeus, 1761): Discrimination of this species pair based on external appearance is challenging, but both the male and female genitalia are diagnostic (Segerer 2001). The species also show a constant barcode divergence of 0.5% corresponding to three diagnostic substitutions.

- *Pyrgus warrenensis* (Verity, 1928) – *P. alveus* (Hübner, 1803): Discrimination of these species based on external appearance is challenging, and the genitalia are also similar (Pro Natura 1997, Segerer 2001). DNA barcoding provides an additional character set for resolving these taxa. For instance, we found a constant minimum pairwise distance of 0.5% between *P. warrenensis* and lowland *P. alveus*, corresponding to three diagnostic substitutions in the barcode fragment. Detailed analysis of these taxa and their relationships to additional allied taxa from other geographic regions will be addressed in a subsequent publication.

- *Diachrysia stenochrysis* (Warren, 1913) – *D. chrysitis* (Linnaeus, 1758) (Noctuidae): These two species show a minimum pairwise divergence of 0.93% in the study area. Even when specimens are examined from a larger geographical area in Europe, the typical diagnostic features (medial area uninterrupted versus interrupted) remain correlated with this COI haplotype divergence. Although several papers have
tried to resolve the taxonomy of this species pair with integrative approaches (e.g. Hille et al. 2005), further research is needed at a broad geographical scale. Since identifications based on morphology and genitalia are not reliable, detailed studies involving rearing and hybridization studies coupled with DNA barcoding of specimens should be undertaken. There is also a need to verify the application of the name *D. stenochrysis* to specimens from Europe, because the type series derives from the Eastern Palearctic (Fauna Europaea 2011). It is possible that the name ‘*D. tutti* (Kostrowicki, 1961)’ should be re-validated at species rank for the European populations.

*Conistra vaccinii* (Linnaeus, 1761) – *C. ligula* (Esper, 1791): Discrimination of these species based on external morphology is challenging (*C. ligula* with forewing apex more tapering and wing coloration usually darker), and the genitalic structures are also similar. However, the two species show a minimum pairwise divergence of 0.77 % in the barcode region.

*Photetes captiuncula* (Treitschke, 1825) – *P. minima* (Haworth, 1809): These two species are easily discriminated based on external morphology, but show a minimum pairwise distance of just 0.64 % in the study area.

*Mesapamea secalis* (Linnaeus, 1758) – *Mesapamea didyma* (Esper, 1788) (= *secalella* Remm, 1983) – *Mesapamea remmi* (Rezbayani-Reser, 1985): *M. secalis* and *M. didyma* cannot be discriminated by external appearance, but both male and female genitalia show clear differences (Rezbayani-Reser 1989). The two species also show two diagnostic substitutions in the barcode region, producing a minimum pairwise distance of 0.31 %. One Bavarian male of the strongly controversial taxon *M. remmi* was barcoded. Although its identification was ascertained by dissection, its barcode was identical to that of *M. didyma*, supporting the hypothesis that *M. remmi* is a F1 hybrid between female *M. didyma* with male *M. secalis*. Nomenclaturally we retain the lectotype designation of *Mesapamea didyma* in Lempke (1988) as valid. After years of general acceptance of the name *M. didyma*, Zilli et al. (2005) suggested that the lectotype of *didyma* would not belong to the type series because it was not figured on Esper’s plates. Therefore, Zilli et al. (2005) and Fauna Europaea (2011) proposed to use *secalella* instead of the putatively doubtful name *didyma*. However, in many similar cases, the entire syntype series has been accepted even if not all specimens were figured in the original description. The text accompanying the original illustration, though published later, makes it clear that Esper did not base his description on a singleton but on a series. The Code (ICZN 1999: §72.4.1., §72.5.6., §73.1.4., 73.2.) provides an opportunity for a wider interpretation and its effort to promote nomenclatorial stability supports this view. Therefore we propose to re-validate *Mesapamea didyma* (Esper, 1788). A 130 bp fragment of the barcode region was recovered from the 223 year old lectotype of *M. didyma* (Esper collection of the ZSM) and it showed 100 % similarity to modern specimens of the species.

*Euxoa nigrofusca* (Esper, 1788) – *E. obelisca* (Denis & Schiffermüller, 1775): Minimum Pairwise Distance 0.77 %. The latter species can usually be recognized by its larger size, its reddish coloration and the diffuse postmedial line without crossing arrow streaks. Data from sites outside Germany suggest that *E. tritici* (Linnaeus, 1761) and *E. nigrofusca* show a minimum pairwise divergence of less than 1 %, but this result awaits confirmation for Germany.

Possible cases of introgression

With an average of just four specimens per species, sample sizes for most species were too low to reveal cases of rare introgression. Nevertheless we detected a few cases where a certain specimen appeared to have ‘the wrong’ mitochondrial genotype. The results of this section are tentative because all cases involve species pairs whose identification through traditional means is challenging:

*Pieris napi* (Linnaeus, 1758) – *P. bryoniae* (Hubner, 1806) (Pieridae): Minimum Pairwise Distance 2.3 % (n = 1/2) in Germany, 1.9 % (n = 25/17) when considering all European data. Although these species are generally easily distinguished by barcodes, one male from southern Germany (Starnberg) had a DNA barcode typical of *P. bryoniae*, while its morphology and capture at low elevation (585 m) suggested that it was *P. napi*. However, the morphological discrimination of male *P. napi* and *P. bryoniae* is often very difficult, and their genitalia are identical. The specimen in question may have been misidentified or it may be a F1 hybrid or advanced generation introgressant. Three cases of possible introgression in the Geometridae (one case in *Isturgia* and two cases in the Eupithecia) were discussed in a previous paper (Hausmann et al. 2011a).

Cases of deep intraspecific divergence

Most German Lepidoptera show very limited intraspecific sequence variation at COI (Figure 2), but about 3 % of the species included two lineages with more than 2 % sequence divergence. These cases included 23 ‘traditionally recognized species’ among the 907 species with full-length barcodes (2574 individuals). When analysis was expanded to the complete data set of 3261 specimens with barcodes (>500 bp; 944 species), nine more haplotype-pairs were detected. 13 of these 32 cases involved
a single specimen outside the main cluster, but 19 taxa included multiple individuals in each barcode cluster. Ten of these splits involved more than 4% divergence, and the deepest divergence was 9.2%. Genitalic dissections and sequencing of nuclear genes are underway to test how many of these cases represent overlooked species pairs. Closely related, young species pairs may be overlooked by a 2 per cent screening threshold, but they can still show constant barcode differentiation as emphasized by the examples of 19 species pairs with divergences under 2% as shown in Table 1 and of 7 species pairs/triplets with constant divergences under 1% discussed above under the section ‘low divergences under 1%’.

Correlations with biogeography

Most cases of intraspecific divergence involved the sympatric occurrence of both barcode clusters, but sample size must be increased and regional coverage extended to properly examine intraspecific geographical patterns within Germany.

In a larger number of German species deep genetic divergences from foreign populations were detected. For instance, Italian specimens of *Spiris striata* (Erebidae/Arctiinae) differed by 4.75% from their German counterparts. This and many other similar cases await careful integrative research.

**Discussion**

**Identification success**

Patterns of DNA barcode variation in 1264 species of German Rhopalocera and Macroheterocera were examined in this study. The results highlight the efficiency of DNA barcoding in species identification as it enables unambiguous re-identifications for nearly 99 per cent of the fauna.

**Low divergences and putatively young species**

Nearly all of the 49 species pairs with low sequence divergence (1–3%; Table 1) are known to be taxonomically problematic. Adopting a standard calibration, COI divergences of 1–2.5 per cent correspond to divergence times of roughly ½ to 1 million years,
suggesting diversification of these species during the Pleistocene (Hausmann et al. 2011a). The fragmentation of distributions and the restriction of populations to glacial refugia were undoubtedly a major driving force for younger speciation events in the European fauna (Dapporto 2009).

Hybridization and introgression

Similar to the results from barcoding Bavarian geometrids (Hausmann et al. 2011a), barcode-sharing was rare in German Rhopalocera and Macroheterocera as it was detected in just 15 species (1.2 %). Several of these cases were inferred from the analysis of specimens from other parts of Europe, but will likely be discovered as sample sizes grow for Germany.

Though awaiting corroboration with additional data, there are, apparently, four species pairs where sharing of mitochondrial sequences reflects either as a result of introgression or F1 hybrids. The case of apparent ‘horizontal exchange’ of mitochondria between species in the high-alpine genera Sciadia and Elophos (Hausmann et al. 2011a) represents a particularly interesting case for subsequent studies.

Cryptic diversity

Despite more than 250 years-long history of research on Lepidoptera of Europe a surprising number of German Rhopalocera and Macroheterocera species (32) was found to show deep barcode divergences, although sample sizes were small and most specimens derived from southern Germany. Although it is premature to reach a final conclusion, we anticipate that some of these cases represent cryptic diversity. DNA barcoding studies have similarly revealed overlooked Lepidoptera species in North America (Hebert et al. 2010, Handfield & Handfield 2006) and in many other parts of the world (e.g. Hausmann et al. 2009a,b, Hausmann & Hebert 2009). At present, the 32 deep divergences are being addressed by a multivariate analysis including large-scale dissecting, enlarging sample-size with data connected with reliable ecological traits (e.g. natural host-plants), and sequencing additional markers.

Interestingly, in butterflies (Rhopalocera) five of the six cases of deep splits range between 2 % and 3 % divergence, and only one shows a deeper split of close to 5 %. In geometrid moths the range is greater, with divergences of up to 9.3 %, and seven (of twenty) splits exceeding 4 %. Surprisingly few cases of deep divergence were detected in the Noc-tuoiidea and Bombycoidea (six of 392 species with ≥2 sequences from Germany), but this may reflect the limited sampling effort on these groups (average = 3.1 specimens per species). Twenty species with deep divergences were found in Geometridae with a sample size of approximately 4.6 specimens per species and 281 species with two or more sequences from the study area (Hausmann et al. 2011a).

DNA barcoding as an efficient tool for biodiversity research

After just two years, the Barcoding Fauna Bavaria (BFB) project has assembled a DNA library for 1264 species of Rhopalocera and Macroheterocera from Germany. Moreover all sequences, georeferenced specimen data and images are freely accessible online.

We believe that this study has established that DNA barcoding provides a reliable, quick and very economical method for monitoring the faunistic and taxonomic aspects of the biodiversity of a whole country. Serious monitoring projects usually require support from many experts for the identification of taxonomically difficult species. Because of the need for careful morphometrical analysis, it often requires 10 to 200 hours per 100 specimens, depending upon the number of dissections required. Although DNA barcoding is not yet cost-effective for routine identifications of random samples (if enough experts are available), it is already a cheaper option for difficult samples. Our current costs for DNA barcoding are about 10 € per specimen for all steps. In groups such as Microlepidoptera, Hymenoptera, or many Coleoptera where morphological identifications are more challenging, DNA barcoding is still more cost-efficient. Moreover, the few specialists can be relieved from routine identifications, allowing them to focus efforts on alpha-taxonomic or phylogenetic work. DNA barcoding also provides an extremely effective approach for the detection of cryptic diversity. Traditional methods are always more time consuming and expensive than DNA barcoding because they require large numbers of dissections and morphometrical analyses for each case. By contrast, when DNA barcoding reveals cryptic lineages, specimens can be carefully targeted for analysis. The taxonomist is not replaced by the molecular technique, but is able to work more efficiently.

Because of its low cost and effectiveness in species identification, DNA barcoding will soon play an important role in biomonitoring programs linked to industrial development, soil and water protection, pest control in forestry and agriculture, food and seed control, environmental studies, nature conservation, performance of monitoring and biodiversity assessment, such as subsequent biological research addressing host plant specificity, phylogeographic patterns, genetic distances correlated with pheno-
logical traits a.s.o. Several studies have already employed DNA barcode data of German Lepidoptera, such as an ecological study monitoring herbivores on a neophytic plant (Gossner & Hausmann 2009), the revision of a geometrid genus (Huemer & Hausmann 2009) and the detection of three overlooked Microlepidoptera species for the fauna of Bavaria (Segerer et al. 2011a,b).

Acknowledgements

Many colleagues at the Biodiversity Institute of Ontario contributed to the success of this study. We are particularly grateful to Sujeewan Ratnasingham, Rodolphe Rougerie, Evgeny Zakharov and Nataly Ivanova. Grants to PDNH in aid of the International Barcode of Life project from Genome Canada through the Ontario Genomics Institute, from the Ontario Ministry of Research and Innovation, and from NSERC enabled both sequence and informatics analysis.

Ralf Bolz (Ullstadt), Ulf Buchsbaum (ZSM Munich), Theo Grünewald (Landschut), Alfred Haslberger (Teisendorf), Richard Heindel (Günzburg), Gerhard Karl (Öd/Mehring), Stefan Koch (Mindelheim), Bernd May (Munich), Rudolf Oswald (Munich), Walter Ruckdeschel (Westerbuchberg), Walter Sage (Niedergottau), Wolfgang Schacht (ZSM Munich), Michael Seizmair (Munich), Georg Stiegel (Augsburg), Ralf Sturm (Rain), Rudolf Tannert (Nürnberg), and Ludwig Weigert (München) submitted important specimens for analysis. Most of the tissue sampling and databasing was performed by Mei-Yu Chen, Johanna Dietrich, Lisa Gangkofner, Johanna Schlereth and Olga Schmidt (ZSM Munich). Stefan Schmidt, Michael Balke and Lars Hendrich play key roles in the BFB project, e.g. in the web presentation, public relations, coordination, administration. The collecting and processing of the specimens were funded as one component of the BFB project from the Bavarian ministry of Science, Research and Art.

We thank Jerome Barbut (Paris), Stoyan Beshkov (Sofia), Robert J. Borth (USA), Ernst Brockmann (Lich-Oberbessingen), Feza Can (Hatay), Sven Erlacher (Chemnitz), Peter Huemer (Innsbruck), Marco Infusino (Messina), Jean-Francois Landry (Ottawa), Vladimir Lukhtanov (St. Petersburg), Antoine Lévêque (Beaugenmcy), Marko Mutanen (Oulu), Wolfgang Nässig (Frankfurt), Stefan Naumann (Berlin), Vazrck Nazari (Guelph), Per Nedreberg (Oslo), Antonio Ortiz (Murcia), Gergely Petrány (Budapest), Norbert Pöll (Jischl), Franz Pühringer (Linz), Rodolphe Rougerie (Rouen), Stefano Scarlacio (Cosenza), Dejan Stojanovic (Novi Sad) and Jeremy deWaar (Vancouver) for granting access to their projects on BOLD. Some of their data were indirectly used for this publication (Appendix S1) and we are very grateful for this collegial, kind support.

Author contributions

Conceived and designed the experimental work: AH, GH, PDNH. Performed the experimental work: AH, AHS. Analyzed the data: AH, AHS, GB, WS. Contributed reagents/materials/analysis tools: AH, AHS, GH, PDNH. Wrote the paper: AH. Provided input into the manuscript: AHS, GB, WS, GH, PDNH.

References


**Supporting information**

Online available – see http://www.zsm.mwn.de/spixiana/toc.htm

**Appendix S1:** Species list and sequencing success

**Appendix S2:** GenBank accession numbers
DNA barcoding of aquatic macroinvertebrates holds much promise as a tool for taxonomic research and for providing the reliable identifications needed for water quality assessment programs. A prerequisite for identification using barcodes is a reliable reference library. We gathered 4165 sequences from the barcode region of the mitochondrial cytochrome c oxidase subunit I gene representing 264 nominal and 90 provisional species of mayflies (Insecta: Ephemeroptera) from Canada, Mexico, and the United States. No species shared barcode sequences and all can be identified with barcodes with the possible exception of some Caenis. (2011) Now DNA-barcoded: the butterflies and larger moths of Germany. Spixiana 34: 47â€“58. View Article. The study of global biodiversity will greatly benefit from access to comprehensive DNA barcode libraries at continental scale, but such datasets are still very rare. Here, we assemble the first high-resolution reference library for European butterflies that provides 97% taxon coverage (459 species) and 22,306 COI sequences. We estimate that we captured 62% of the total haplotype diversity and show that most species possess a few very common haplotypes and many rare ones. Specimens in the dataset have an average 95.3% probability of being correctly identified. Mitochondrial diversity displayed elevated haplotype richness in southern European refugia, establishing the generality of this key biogeographic pattern for an entire taxonomic group. A German sequence of the holarctic species Orgyia antiqua L. 1758 was used as outgroup. Two sequences of Spanish specimens of O. trigotephras from Murcia (by the courtesy of A. Ortiz), and one of a southern Italian specimen (A. Hausmann) were included into the analysis. stage egg-mass. (2003).