

# Cryptic matters: overlooked species generate most butterfly beta-diversity

Raluca Vodă, Leonardo Dapporto, Vlad Dincă and Roger Vila

*R. Vodă, V. Dincă and R. Vila (roger.vila@csic.es), Inst. de Biologia Evolutiva (CSIC-Univ. Pompeu Fabra), Passeig Marítim de la Barceloneta 37, ES-08003, Barcelona, Spain. R. Vodă also at: Dept de Genètica i Microbiologia, Univ. Autònoma de Barcelona, ES-08193, Bellaterra, Spain. VD also at: Dept of Zoology, Stockholm Univ., SE-106 91, Stockholm, Sweden, and Biodiversity Inst. of Ontario, Univ. of Guelph, Guelph, ON N1G 2W1, Canada. – L. Dapporto, Dept of Biological and Medical Sciences, Oxford Brookes Univ., Headington, Oxford, OX3 0BP, UK.*

The cryptic fraction of biodiversity is composed of morphologically similar species that are or have been overlooked by scientists. Although current research is increasingly documenting new cases, cryptic species are frequently ignored in large-scale studies and monitoring programs, either because they have not yet been discovered, or because of the practical difficulties involved in differentiating them. However, it is unknown if this could represent a bias extending beyond the number of missed species. By analyzing the butterfly fauna of the west Mediterranean (335 species), we defined cryptic species based on the current consensus of the scientific community, compared their properties to other congeneric species and investigated the consequences of their inclusion/exclusion in beta-diversity analyses. We show that, as defined, the cryptic fraction of butterfly diversity represents about 25% of the west Mediterranean fauna and is overwhelmingly composed by groups of species that are not sympatric. Our results show that co-occurrence among cryptic species is significantly lower than among congeneric non-cryptic species. Accordingly, albeit the frequency of cryptic species is homogeneously distributed over the study area, their distribution pattern accounts for most beta-diversity turnover over sea (from 50 to 100%). Beta-diversity turnover, a direct measure of the frequency of species replacement from site to site, is recognized as a fundamental parameter in ecology and is widely used to detect biogeographic patterns. These findings represent a change of paradigm in showing that cryptic diversity comprises original qualitative aspects in addition to merely quantitative ones. This highlights the importance of differentiating cryptic species for various research fields and opens the door to the study of further potential particularities of cryptic diversity.

Recent molecular studies unveiled a substantial fraction of biodiversity by recognizing a plethora of cryptic species in numerous taxonomic groups and across biogeographical regions (Pfenninger and Schwenk 2007). Cryptic species (Steyskal 1972) are those that have been erroneously treated as conspecific because of high morphological resemblance, and sometimes also behavioral or ecological similarity (Bickford et al. 2007). Thus, the concept of cryptic species is anthropocentric by definition because it is based on human perception of species. Irrespective of the evolutionary significance of delimiting such a group of species, it is justified to ask about the practical consequences of their discovery or omission. To date, no specific studies directly compared the properties of cryptic and non-cryptic diversity, perhaps due to the difficulty of unambiguously identifying cryptic species and their precise distributions for a sufficiently large area and taxonomic group. Consequently, researchers frequently simplify their datasets by not distinguishing cryptic species (Dennis 1997, Isaac et al. 2004, Pfenninger and Schwenk 2007, Descimon and Mallet 2009, Devictor et al. 2012).

Due to their similar natural history and appearance, cryptic species are predicted to compete for resources and

to reproductively interfere between them (Pigot and Tobias 2013). Results for a series of cryptic butterfly species revealed only restricted contact zones on mainland and chequered distributions over narrow (3–10 km) sea straits (Dincă et al. 2011, Dapporto et al. 2012, Sañudo-Restrepo et al. 2013). If chequered distributions are overrepresented among cryptic species, they could encompass a disproportionately high fraction of beta-diversity turnover, a direct measure of the frequency of species replacement from site to site. Spatial turnover is recognized as a fundamental parameter in ecology and is increasingly used to detect biogeographic patterns (Holt et al. 2013).

To assess the effects of ignoring cryptic diversity, we studied the biogeographical properties of cryptic species uncovered since 1970, after the cornerstone work of Higgins and Riley (1970) that have been generally accepted by the scientific community. In particular, we explored the prevalence of chequered distribution patterns among groups of cryptic taxa in west Mediterranean butterflies, which represent one of the best-studied invertebrate groups, especially regarding their taxonomy and distribution. We also assessed if cryptic taxa account for an unbalanced fraction of the beta-diversity

turnover and concluded that, although they are frequently excluded from large-scale surveys and studies, their contribution may be far more important than the mere percentage of species they represent.

## Methods

### Presence data and co-occurrence analyses

We analyzed butterfly presence data from 82 areas comprising 27 small islands and 55 sites of 25 km radius located on mainland (43) and on large islands (12) (Supplementary material Dataset 1). An objective identification of species is a crucial step for comparative studies at continental level, especially on a diverse group such as butterflies. Although European butterflies have been among the most studied invertebrates in the last centuries, recognition of species is not always straightforward and it has been based on different concepts for various taxa. In some cases species are readily recognizable based on their external morphology, while in other cases, in depth morphological and molecular analyses, as well as the assessment of reproductive barriers, have been needed (reviewed by Descimon and Mallet 2009). For these reasons, we used a taxonomic species concept (reviewed by Mayden 1997) recognizing species as ‘whatever a competent taxonomist chooses to call a species’ (Wilkins 2009). We applied this criterion and compared past taxonomy with the current consensus of the scientific community in the following way.

1) We produced a list of species based on the widely accepted checklist of the Fauna Europaea Project (Karsholt and Nieuwerkerken 2013, <www.faunaeur.org>), funded by the European Commission (EVR1-1999-2001). We complemented this list with the north African species recognized in the recent synthesis works that include this region (Tennent 1996, TARRIER and Delacré 2008, Tshikolovets 2011). We also added four species (*Zerynthia cassandra*, *Polyommatus celina*, *Coenonympha lylus*, *Coenonympha darwiniana*) and eliminated one (*Polyommatus gallois*) based on recent molecular studies, generally combined with morphometrics, mate choice experiments and/or ecological data (see Supplementary material Appendix 1 for references). These changes have not yet been introduced in Fauna Europaea but will most likely be in the next version.

2) Within the list of species obtained, we considered as cryptic those pairs or groups of species that have been considered as a single species in at least one of the most widely used synthesis works regarding Europe and/or north Africa published since the reference work of Higgins and Riley (1970) (see Supplementary material Appendix 1 for discussions on each case). The references provided should be regarded as examples and usually cover general works that likely represent widely accepted trends in European butterfly taxonomy. In specific cases where relevant studies were available, these have been cited as well.

Co-occurrence between pairs of species has been computed as the fraction of shared areas by using the Sørensen index  $QS = 2C/(A + B)$ , where C is the number of areas where the two species co-occur and A and B are the number of areas where the first and second species are

present, respectively. We computed the QS between pairs of cryptic species and between pairs of non-cryptic species belonging to the same genus. To test for significant differences in QS values we performed the Mann–Whitney U test with 1000 Monte Carlo permutations.

### Beta-diversity turnover

To evaluate the Contribution of Cryptic Species to the beta-diversity Turnover (CCST), we calculated the Simpson index among the 82 study areas. Subsequently, for each presence submatrix composed of each cryptic group, we computed a matrix (sm) for areas by using the Simpson dissimilarity denominator. Finally, we added all sm and divided this sum matrix by the total Simpson matrix (also containing the contribution of non-cryptic species). Following the previous description, CCST can be computed by the formula:

$$CCST = \frac{\sum_{sm=1}^{36} (A_{sm} + \min(B_{sm}, C_{sm}))}{A - \min(B, C)}$$

where  $A_{sm}$  is the number of shared species and  $B_{sm}$  and  $C_{sm}$  are the number of non-shared species among areas for each of the 36 cryptic groups considered separately (sm). A is the overall number of shared species and B and C the overall number of non-shared species among the same pairs of the 82 areas. With the ‘deldir’ R package we calculated a Delaunay triangulation among areas using their center GPS coordinates and plotted the CCST index on a map. The CCST among pairs of areas were attributed to the midpoints connecting those areas and were interpolated through Inverse-Distance-Weighting using Spatial Analyst (<www.esri.com>).

To evaluate the importance of cryptic species in determining the overall turnover pattern, we calculated the mean value of the ratio between the Simpson and the Sørensen dissimilarity matrices ( $\beta_{simp}/\beta_{sor}$ ) among areas. This ratio is considered to be the best measure describing the contribution of turnover to the overall beta-diversity pattern (Dobrovolski et al. 2012). We then removed all cryptic taxa and recalculated the same mean value. Finally, we created 10 000 matrices by removing a random number of species equal to the number of cryptic ones and calculated their mean  $\beta_{simp}/\beta_{sor}$ . We tested if the matrix lacking all cryptic taxa shows a significantly lower ratio than the random ones (if at least 95% of the random matrices showed a  $\beta_{simp}/\beta_{sor}$  higher than the matrix without all cryptic species).

## Results

### Co-occurrence analyses for cryptic and non-cryptic diversity fractions

In total, 86 species out of 335 (25.67%) are involved in 36 cryptic groups. In most groups (20), cryptic species have been separated only on the basis of differences in morphology (mostly male genitalia), while in six groups species were

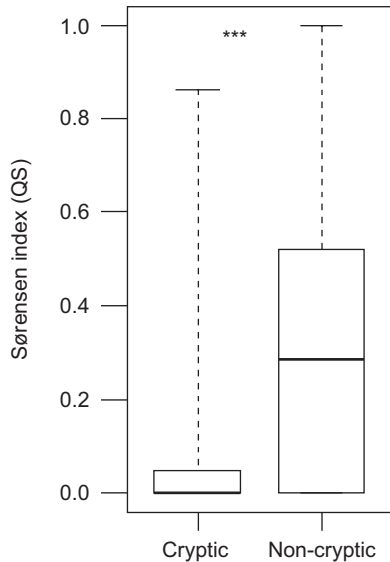


Figure 1. Boxplot indicating median values of the Sørensen index (QS) for pairs or groups of cryptic species (Cryptic) and congeneric pairs of non-cryptic species (Non-cryptic). Cryptic species share a significantly smaller fraction of their distributions compared with the non-cryptic ones.

recognized based exclusively on molecular evidence (DNA, allozyme or number of chromosomes) because no relevant differences in morphology have been detected. In nine groups morphological and molecular evidence have been coupled, and in one group (*Leptidea*) morphological and molecular analysis have been combined with mating experiments (Supplementary material Appendix 1). The Mann–Whitney test ( $U = 15319.5$ , Monte Carlo  $p < 0.001$ ) revealed that cryptic species pairs have significantly lower QS than non-cryptic pairs of species (Fig. 1), which means that they share a smaller fraction of their distributions.

### Contribution of non-sympatric cryptic species to the overall beta-diversity pattern

Although the fraction of cryptic diversity was relatively constant among the 82 study areas ( $23.02 \pm 0.04$ , Fig. 2a), the fraction of beta-diversity turnover due to replacements between cryptic species showed pronounced variation over the west Mediterranean (Fig. 2b). As a general rule, the CCST was higher than 50% only over sea areas, with especially high values of more than 80% between Sardinia-Corsica-Italian mainland, Sicily-Italian mainland, Sicily-north Africa and Balearics-Catalonia-France. By contrast, the CCST was lower (about 10%) between sites separated by land (Fig. 2b). The complete presence matrix revealed a mean  $\beta_{\text{simp}}/\beta_{\text{sor}}$  value of 0.450 (Fig. 2c). When all the 86 cryptic species have been removed the mean value decreased to 0.346, while the mean value for cryptic species alone was 0.666. None of the 10 000 matrices generated by removing 86 random species showed a  $\beta_{\text{simp}}/\beta_{\text{sor}}$  mean value lower than the matrix lacking the 86 cryptic species ( $p < 0.001$ , Fig. 2c).

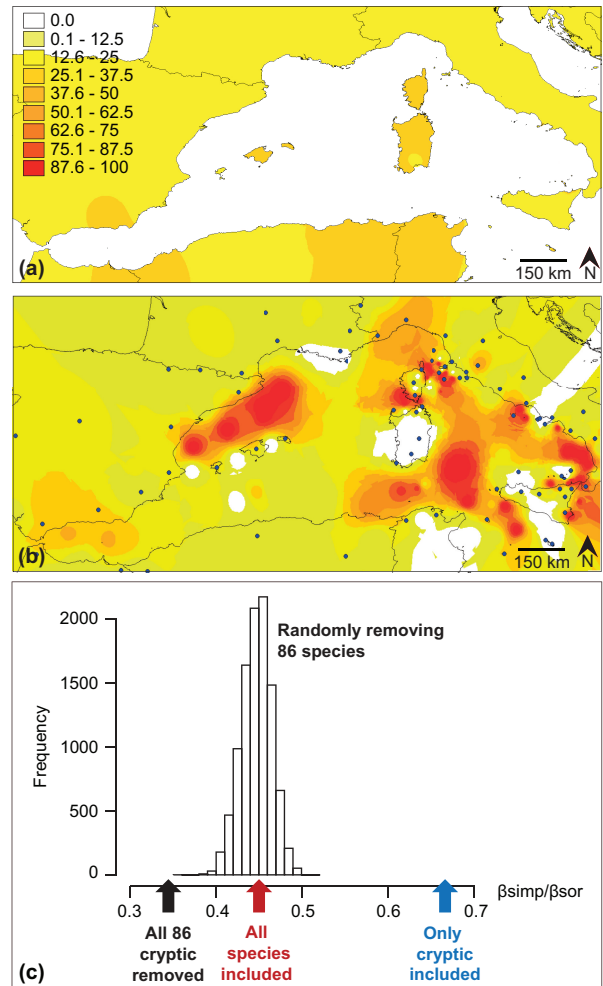


Figure 2. (a) Distribution map (equirectangular projection) of the percentage that cryptic butterfly species represent of the total diversity. (b) Landscape analysis for the Contribution of Cryptic Species to the beta-diversity Turnover (CCST). (c) Mean  $\beta_{\text{simp}}/\beta_{\text{sor}}$  value of the complete matrix (red arrow) compared with the value obtained by removing all 86 cryptic species (black arrow) and by using only cryptic species (cyan arrow). The histogram represents frequencies of  $\beta_{\text{simp}}/\beta_{\text{sor}}$  mean values for 10 000 matrices generated by removing 86 random species.

### Discussion

Due to their morphological similarity and complex taxonomy, cryptic groups are frequently ignored or treated as a unit in large-scale studies and monitoring schemes. In this study we objectively defined cryptic species and showed that co-occurrence within cryptic groups over the west Mediterranean is much lower than among congeneric non-cryptic species. This result could be explained if we assume that cryptic groups are composed by entities that recently evolved in allopatry and maintained their distribution pattern simply because they did not have enough time to disperse. However, many cryptic taxa in our study are not sister species and show substantial genetic divergence, while other morphologically differentiated species are their sister taxa; in most such cases a chequered distribution still occurs (Dincă et al. 2011, Sañudo-Restrepo et al. 2013).

The mechanisms maintaining chequered distributions across specific narrow sea straits are still unknown, but a convincing hypothesis can be formulated. Recent studies have shown that butterflies can experience range expansions/contractions even over relatively short time periods (Dapporto and Bruschini 2012, Dapporto et al. 2012). It is thus likely that, in addition to temporal and physical constraints, the dispersal of most cryptic species has been hampered by the presence of other members of the same group in the recipient areas (Waters et al. 2013), a phenomenon recently demonstrated in birds (Pigot and Tobias 2013). Presumably, no strong precopulatory barriers exist between many cryptic butterfly species (Wickman 2009) and hybrids between related taxa often display reduced fitness (Descimon and Mallet 2009). In the absence of specific mate recognition mechanisms, contact areas may act as population sinks unlikely to enlarge, although they are predicted to shift until areas allowing low dispersal and low population densities are reached (Dasmahapatra et al. 2010). A concentration of boundaries between cryptic species over sea straits perfectly matches this hypothesis. Noteworthy, the fraction of cryptic diversity is homogeneous across the study region and almost all cryptic taxa that we analyzed occur on both mainland and islands. For this reason, the striking contribution to island divergence is not due to an unbalanced occurrence of island endemics among cryptic taxa. These mechanisms are generally applicable to most organisms, and thus it is probable that the phenomenon here discovered for butterflies holds for other taxa as well, albeit the exact impact on the overall patterns may vary.

We conclude that unveiling the cryptic fraction of biodiversity is of primary importance for a comprehensive description of biogeographical patterns. The finding that cryptic groups of species present characteristic geographical distributions highlights the need to study further potential particularities they may display, as well as to assess their implications in multiple research fields. We showed that ignoring cryptic species results in losing a significant fraction of spatially structured taxonomic variation, which hampers the opportunity to reveal phylogeographic suture zones (Moritz et al. 2009) and to investigate historical and ecological processes determining distribution range dynamics (Marske et al. 2013). Such phenomena underlie fundamental processes like dispersal, extinction, environmental filtering, competition, sexual interference, hybrid depression, and ‘founder takes all’ dynamics (Marske et al. 2013, Waters et al. 2013). Moreover, it is likely that cryptic species respond differently to climatic conditions (Pearman et al. 2010) and geographical shifts in their contact zones correlated with climate change have been documented (Mercader and Scriber 2008, Mallet et al. 2011). This suggests that studies quantifying the effects of global climatic changes on organisms may be missing part of the signal stemming from cryptic biodiversity. The same could apply to life history and other ecological preferences that may be differentially expressed in morphologically similar species. Morphological differentiation may well be the result of diverging selective pressures, which could generally lack in these groups of species. Different evolutionary and speciation factors and processes may result in varying degrees of morphological differentiation (*Heliconius* genome consortium 2012), and

thus phylogenetic reconstruction and macroecological inference could be affected due to underrepresented cryptic diversity.

Case studies typically rely on well-assessed taxonomy, but macroecology and comparative studies are usually based on data that encompass a great number of species and are more susceptible of missing cryptic taxa. The costs and benefits of taking into consideration cryptic species have been rarely discussed due to a reductionist view of their importance limited to alpha-taxonomy, i.e. species numbers (Isaac et al. 2004). From this perspective, our study not only provides novel insights into biogeography, but it also serves as a heads-up about the importance of including the cryptic diversity in ecological, biodiversity and biogeographic studies to avoid biased results.

*Acknowledgements* – Funding was provided by the Spanish Ministerio de Ciencia e Innovación (Project CGL2010-21226/BOS and a predoctoral fellowship to R. Vodá, ref. AP-2010-5409), by the project ‘Barcoding Italian butterflies’ and by a postdoctoral fellowship from the Wenner-Gren Foundation Sweden and a EU Marie Curie International Outgoing Fellowship for Career Development (project no. 625997) to VD.

## References

- Bickford, D. et al. 2007. Cryptic species as a window on diversity and conservation. – *Trends Ecol. Evol.* 22: 148–155.
- Dapporto, L. and Bruschini, C. 2012. Invading a refugium: post glacial replacement of the ancestral lineage of a Nymphalid butterfly in the west Mediterranean. – *Org. Divers. Evol.* 12: 39–49.
- Dapporto, L. et al. 2012. Identifying zones of phenetic compression in west Mediterranean butterflies (Satyriinae): refugia, invasion and hybridization. – *Divers. Distrib.* 18: 1066–1076.
- Dasmahapatra, K. K. et al. 2010. The anatomy of a ‘suture zone’ in Amazonian butterflies: a coalescent-based test for vicariant geographic divergence and speciation. – *Mol. Ecol.* 19: 4283–4301.
- Dennis, R. L. H. 1997. An inflated conservation load for European butterflies: increases in rarity and endemism accompany increases in species richness. – *J. Insect Conserv.* 1: 43–62.
- Descimon, H. and Mallet, J. 2009. Bad species. – In: Settele, J. et al. (eds), *Ecology of butterflies in Europe*. Cambridge Univ. Press, pp. 219–249.
- Devictor, V. et al. 2012. Differences in the climatic debts of birds and butterflies at a continental scale. – *Nat. Clim. Change* 2: 121–124.
- Dincă, V. et al. 2011. A combined genetic-morphometric analysis unravels the complex biogeographical history of *Polyommatus icarus* and *Polyommatus celina* common blue butterflies. – *Mol. Ecol.* 20: 3921–3935.
- Dobrovolski, R. et al. 2012. Climatic history and dispersal ability explain the relative importance of turnover and nestedness components of beta diversity. – *Global Ecol. Biogeogr.* 21: 191–197.
- Heliconius* Genome Consortium 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptation among species. – *Nature* 487: 94–98.
- Higgins, L. G. and Riley, N. D. 1970. *A field guide to the butterflies of Britain and Europe*. – Houghton Mifflin, 1st American ed.
- Holt, B. et al. 2013. An update of Wallace’s zoogeographic regions of the World. – *Science* 339: 74–78.
- Isaac, N. J. B. et al. 2004. Taxonomic inflation: its influence on macroecology and conservation. – *Trends Ecol. Evol.* 19: 464–469.



- Karsholt, O. and Nieuwerkerken, E. J. van. 2013. Lepidoptera, Papilionoidea. – Fauna Europaea ver. 1.0, <[www.faunaeur.org](http://www.faunaeur.org)>
- Mallet, J. et al. 2011. Hybridisation and climate change: brown argus butterflies in Britain (*Polyommatus* subgenus *Aricia*). – Insect Conserv. Divers. 4: 192–199.
- Marske, K. A. et al. 2013. Phylogeography: spanning the ecology–evolution continuum. – *Ecography* 36: 1169–1181.
- Mayden, R. L. 1997. A hierarchy of species concepts: the denouncement in the saga of the species problem – In: Claridge, M. F. et al. (eds), *Species: the units of biodiversity*. Chapman and Hall, pp. 381–424.
- Mercader, R. J. and Scriber, J. M. 2008. Asymmetrical thermal constraints on the parapatric species boundaries of two widespread generalist butterflies. – *Ecol. Entomol.* 33: 537–545.
- Moritz, C. et al. 2009. Identification and dynamics of a cryptic suture zone in tropical rainforest. – *Proc. R. Soc. B* 276: 1235–1244.
- Pearman, P. B. et al. 2010. Within-taxon niche structure: niche conservatism, divergence and predicted effects of climate change. – *Ecography* 33: 990–1003.
- Pfenninger, M. and Schwenk, K. 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. – *BMC Evol. Biol.* 7.
- Pigot, A. L. and Tobias, J. A. 2013. Species interactions constrain geographic range expansion over evolutionary time. – *Ecol. Lett.* 16: 330–338.
- Sañudo-Restrepo, C. P. et al. 2013. Biogeography and systematics of *Aricia* butterflies (Lepidoptera, Lycaenidae). – *Mol. Phylogenet. Evol.* 66: 369–379
- Steyskal, G. C. 1972. The meaning of the term “sibling species”. – *Syst. Zool.* 21: 239.
- Tarrier, M. R. and Delacre, J. 2008. *Les papillons de jour du Maroc*. – Biotope, Mèze.
- Tennent, J. 1996. *The butterflies of Morocco, Algeria and Tunisia*. – Gem Publishing, Walling for.
- Tshikolovets, V. V. 2011. *Butterflies of Europe and the Mediterranean area* – Tshikolovets Publications.
- Waters, J. M. et al. 2013 Founder takes all: density dependent processes structure biodiversity. – *Trends Ecol. Evol.* 28: 78–85.
- Wickman, P.-O. 2009. Mating behaviour in butterflies. – In: Settele, J. et al. (eds), *Ecology of butterflies in Europe*. Cambridge Univ. Press, pp. 17–28.
- Wilkins, J. S. 2009. *Species: a history of the idea*. – Univ. of California Press.

Supplementary material (Appendix ECOG-00762 at <[www.ecography.org/readers/appendix](http://www.ecography.org/readers/appendix)>). Appendix 1 and Dataset 1.