Trial test of external morphology-based identification of *Leptidea sinapis*, *L. reali* and *L. juvernica* (Lepidoptera: Pieridae) provides opportunity for an online identification platform

Sylvain Cuvelier & Dave Maertens

Abstract. The results of a trial to test identification criteria of external morphology of the *Leptidea* triplet (*Leptidea sinapis, L. reali, L. juvernica*) with blind readers using an online application are presented. The original application has been slightly modified and can now be used as a training module (http://butterfly.lifetrail.be) for the identification of these *Leptidea* species.

Samenvatting. De resultaten worden voorgesteld van een studie met geblindeerde lezers, voor de determinatie van het *Leptidea* triplet (*Leptidea sinapis, L. reali, L. juvernica*) aan de hand van uitwendige morfologische kenmerken. Hierbij is gebruik gemaakt van een online applicatie. De oorspronkelijke applicatie is licht gewijzigd en kan nu gebruikt worden als een trainingsmodule (http://vlinders.lifetrail.be) voor de determinatie van deze *Leptidea*-soorten.

Résumé. Les résultats d'une étude d'identification avec des lecteurs travaillant en aveugle pour la détermination du triplet *Leptidea* (*Leptidea sinapis, L. reali, L. juvernica*) se basant sur des critères morphologiques externes sont présentés. A cet effet une application web a été utilisée. L'application initiale a été légèrement modifiée et peut être utilisée comme un module de formation (http://butterfly.lifetrail.be) pour la détermination de ces espèces du genre *Leptidea*.

Key words: *Leptidea sinapis* – *L. reali* – *L. juvernica* – external criteria – online identification platform.

Cuvelier S.: Diamantstraat 4, B-8900 Ieper, Belgium. sylvain.cuvelier@pandora.be

Maertens D.: Landegemstraat 74, B-9850 Vosselare, Belgium. dave@lifetrail.be

These authors contributed equally to this work.

Introduction

For different goals (faunistics, nature management, conservation, monitoring, …) it is important to have precise information about the butterfly biodiversity of specific areas or regions. However, species identification based only on external criteria, often poses problems and is far from 100% reliable.

Examination of genitalia morphology, data from mitochondrial DNA (mtDNA) (e.g. DNA barcoding) or more detailed information from nuclear markers is often needed to resolve doubtful identifications.

Since the discovery of unexpected layers of cryptic diversity in Wood White butterflies (Dincă *et al*. 2011) the identification of *Leptidea* specimens based on their habitus, became very challenging. In subsequent studies, the existence of a triplet (*Leptidea sinapis* (Linnaeus, 1758), *Leptidea reali* (Reissinger, 1990) and *Leptidea juvernica* (Williams, 1946) was substantiated (Dincă *et al*. 2013; Šíchová *et al*. 2015). For reliable identification of specimens in the contact zones between *L. reali* and *L. juvernica*, DNA is now the recommended identification method.

Shortly after the original discovery of the *Leptidea* triplet, Mazel (2012) published external morphological criteria to separate the three species in France.

The first objective of the trial was to test, with blind readers, the reliability of these criteria on specimens that have been identified through DNA barcoding and/or genitalia and that originate from a much wider area within the Western Palaearctic.

The second objective was to build an online platform for the identification, by blind users, of many *Leptidea* specimens and for real time analysis of the results. The validation of such an application in a trial could lead to the development of a training module for the identification of this challenging *Leptidea* group and to new identification platforms for other difficult groups (e.g. *Pyrgus, Melitaea*,…).

Methods

Wing vouchers (Fig. 1a–b) of 85 specimens, including the three taxa and from different generations were kindly provided by Vlad Dincă and Roger Vila (Institut de Biologia Evolutiva CSIC-UPF, Barcelona, Spain). Specimens were included from a wide area in the Western Palaearctic (Supplementary Table 1).

Wing vouchers (upper and underside) of every specimen were photographed in standardized studio conditions (Tripod, Nikon D90, Sigma AF 180mm f/3.5 EX DG HSM APO).

The external characters described by Mazel (2012) were used to define 10 diagnostic features (6 on the upperside of the wings; 4 on the underside of the wings). Each diagnostic feature has a limited list of possible values, resulting in a total of 40 variables per specimen (Fig. 2a–b, Supplementary Table 2).

16 readers were included and had to identify the whole set of specimens at two different stages of the study. For the first identification (first stage = 1S) no information concerning the origin (date and locality) of the specimens was given. After all readers finished 1S, the origin of each specimen was released and a second identification (final determination = FD) was performed.

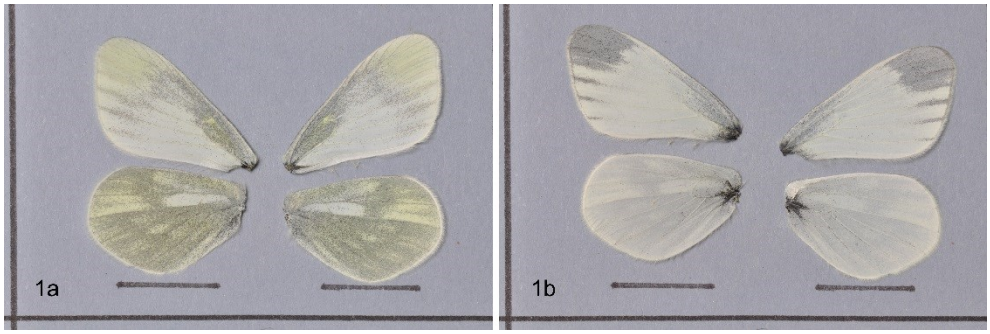

Fig. 1a–b. Wing vouchers of underside (a) and upperside (b) of ♂ *Leptidea juvernica*, Cork (Ireland), 26.v.2012 (leg. V. Dincă & C. Wiklund; © S. Cuvelier).

To facilitate the trial, an online tool was developed. This custom application was based on the technology of Oracle Application Express (Oracle Apex). The main features of this application are:

User self-registration, identification, authentication and authorisation.

Usage of different states per specimen. For all specimens, every blind reader had to go through the different, colour-based states: from red (to start) over orange (started, at least one diagnostic feature completed) to yellow (completed = 1S) and after releasing the specimen's origin to a green state (final = FD). At the end the administrators released the full dataset to the readers bringing all the specimens to a blue state.

Reporting personal and overall progress in order to interactively follow the individual progress of the reader.

Reporting personal identification scores.

Activity logging.

Extensive reporting and analysing possibilities for the administrators.

The analysis of the identifications was planned at different levels:

FD, global.

FD, comparison of the individual results.

1S and FD, global comparison of each taxon.

1S and FD, comparison per generation for each taxon.

1S and FD, comparison per gender for each taxon.

1S and FD, comparison of combination generation and gender for each taxon.

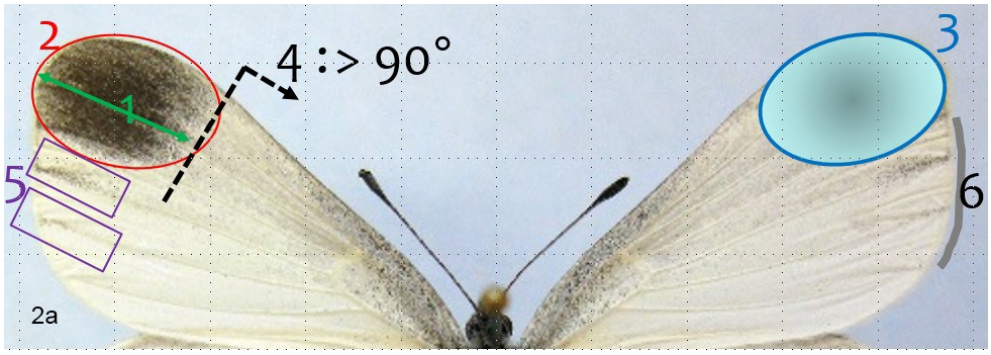

Fig. 2a. Diagnostic features on the upperside of the wings.

1: apical spot, dimension;

2: apical spot, shape;

3: apical spot, pattern;

4: corner, costal margin;

5: dusted veins, outer margin;

6: dusted scales, outer margin.

Fig. 2b. Diagnostic features on the underside of the wings.

7: forewing colour;

8: hindwing colour;

9: markings, hindwing;

10: stripes between veins, hindwing. (© S. Cuvelier).

A screenshot from the application (Fig. 3), shows a completed screen after stage 1S with the ten diagnostic features for a given specimen. The blue bar on the left is the menu, which remains visible throughout the entire application. For optimal visual analysis of every specimen the two screens with the wing vouchers can be maximised in a pop-up window by clicking on “Detail upperside” or “Detail underside”.

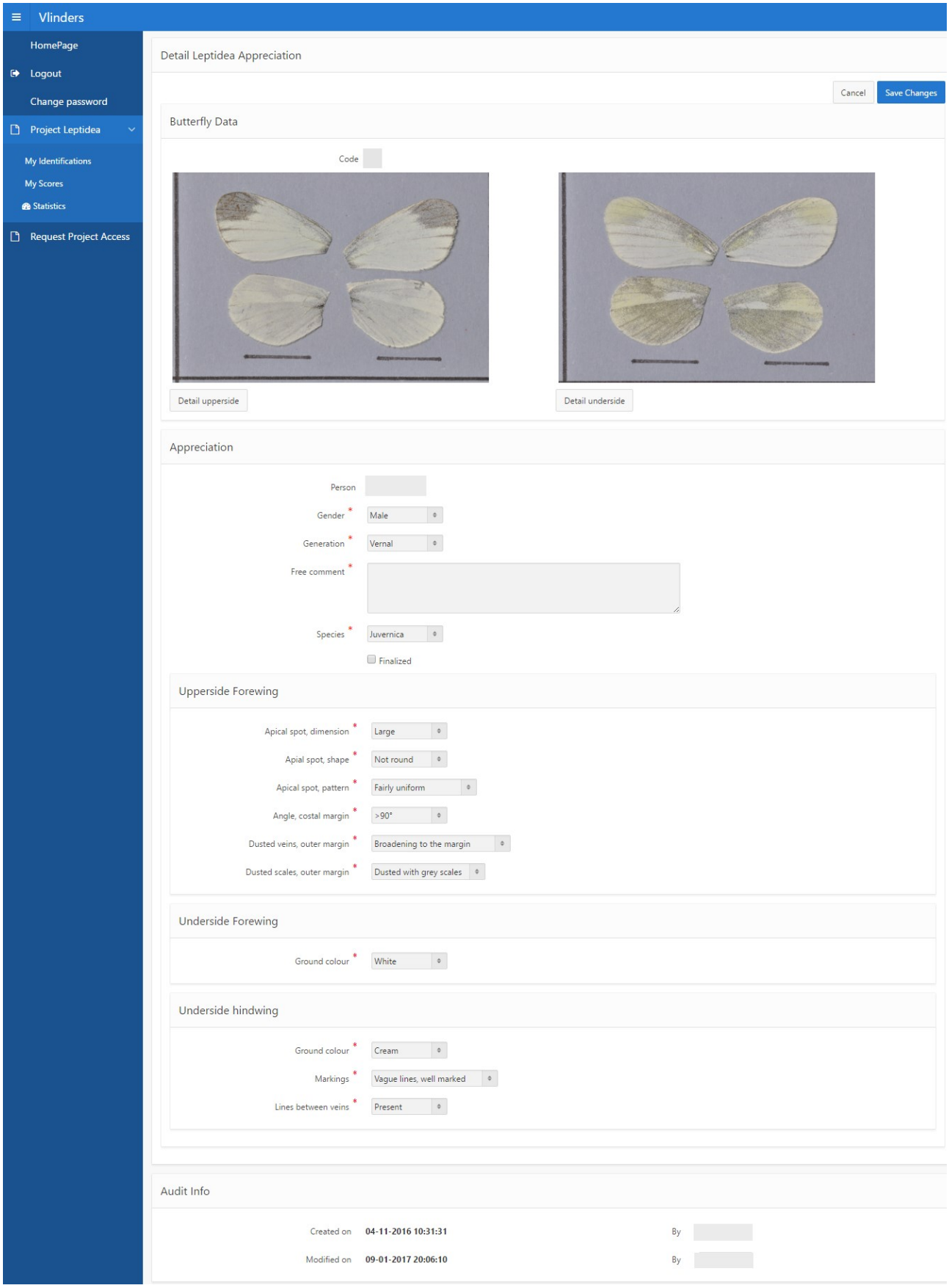

Fig. 3. Screenshot of the application with a completed input after stage 1S for one specimen (© D. Maertens).

Results and discussion

85 specimens were included in this trial. One reader did not complete the entire dataset and was withdrawn from the analysis of the study.

For all the *Leptidea* taxa the correct FD was 66% (Supplementary Table 1).

100 % identification was reached for 6 specimens (4 *L. sinapis*, 1 *L. juvernica* and 1 *L. reali*) of the 85 *Leptidea* (7.1%) and 1 specimen of *L. juvernica* was not at all identified by the 15 readers. This shows that the criteria described by Mazel (2012) can help to get closer to a correct identification of the *Leptidea* group in the Western Palaearctic. But as we need fully reliable identifications for different goals, it is not sufficient. The result can probably improve when considering the analysis of higher numbers of *Leptidea* in a single locality even though the three *Leptidea* taxa might be sympatric in certain contact zones (e.g. Dincă *et al*. 2013).

The identification results for the FD (Fig. 4) by reader, show a broad range: 47.1%–77.6%. This can be explained partly by the different level of experience in the identification of the *Leptidea* triplet inherent within the readers. However, the difference of 30.5% clearly indicates the need for a training tool to improve the identification ability of readers. It is clear that it can be difficult or even impossible to correctly identify some taxa even when providing a standardised set of specimens and criteria. The interpretation remains subjective and we need to improve the quality of the identification keys in standard entomological literature. This also shows the extent of the difficulty in the identification of butterflies photographed in nature. The interpretation of such photographic material often needs a lot of common sense. In all taxa where certain identification is not possible, we advocate the sampling of specimens for further study.

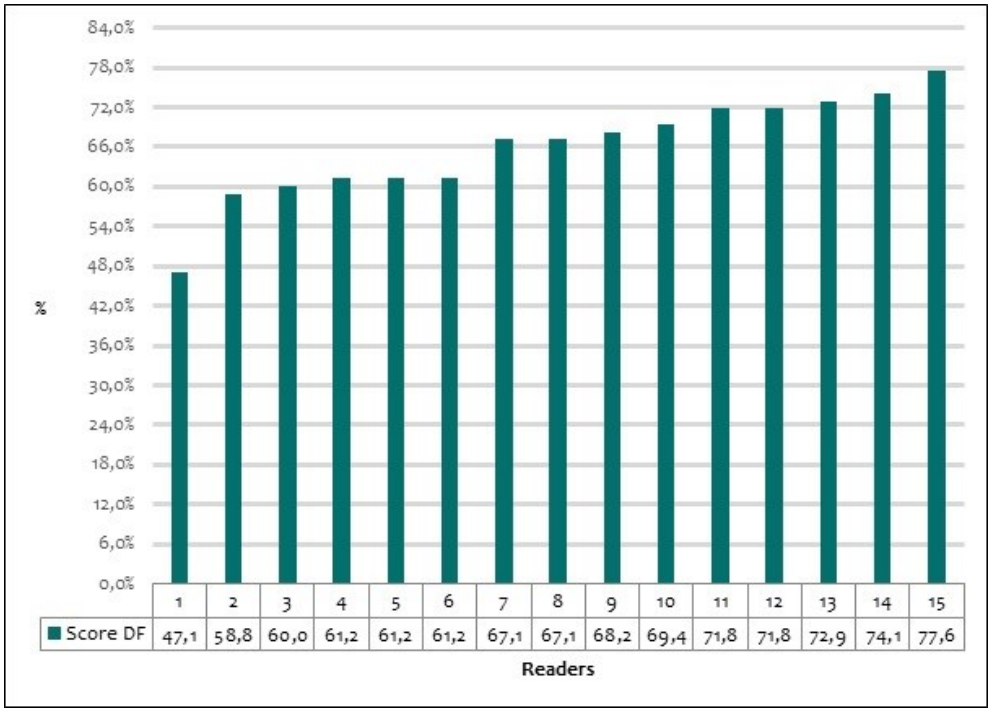

Fig. 4. Comparison of the FD results by reader. Mean FD: 66.0% (47.1–77.6%).

Without the origin of a specimen being given the three taxa are identified correctly in approximately half of the specimens (Supplementary Table 3a). Releasing the origin of each specimen the FD identification clearly increased for *L. sinapis* (+ 9.8%) and *L. juvernica* (+19.7%) but had little influence on the correct identification of *L. reali* (+3.7%).

Comparing the results by generation (Supplementary Table 3b) for the three taxa, seems to provide some tools to improve the FD identifications at a given locality. Aestival *L. sinapis* were correctly identified in 79.6%. In springtime the reliability is higher for *L. juvernica* (77.9%) and *L. reali* (74.3%).

Comparing the results by gender (Supplementary Table 3c) gives better FD identifications for males *L. sinapis* (74.5%) and *L. reali* (65.0%) and for females *L. juvernica* (75.6%). However, a higher number of specimens is needed to confirm these findings.

Combining generation and gender (Supplementary Table 3d) for each taxon gives subgroups that have higher identification FD results: aestival females *L. sinapis* (86.7%), vernal females *L. juvernica* (81.3%) and vernal males *L. reali* (84.4%). Some of these subgroups are however based on very few specimens. It is necessary to increase the sampling size to verify if these trends are correct.

Without the use of the application, it would be more difficult to standardize the way of performing this trial. All data are stored in a standardized relational database (Oracle). The system has numerous built-in checks for completeness and accuracy, without which the manual

quality control and project follow-up would be extremely time consuming and labour intensive.

There is a noticeable encouraging effect on the participants when they receive feedback from the application regarding their work status.

The application proved to provide real-time and easy administrative follow-up of both the individual and overall progress.

Conclusion

Considering the 66.0% FD, the correct identification of only 6 specimens (7.1%) by all readers and the single specimen that was not identified by any of the readers, the identification of the *Leptidea* triplet based on the criteria published by Mazel is unsatisfactory for the Western Palaearctic. Results in some subgroups give the impression that the identification can be optimized (with potentially higher success at a more local scale) but a higher sample size is needed to confirm this.

Next to larger sample sizes including both sexes and different generations of the *Leptidea* triplet, we recommend the inclusion of other external anatomical features, e.g. wing scales and venation, antennae and legs.

Considering the wide range of individual FD identifications from the 15 readers and, by extrapolation, thinking about what can be expected for many other taxa with highly similar external morphology, improving the quality of the identification keys in future entomological literature is mandatory.

The original application has been slightly modified to be used as a generic platform for the identification of these *Leptidea* species.

On the website http://butterfly.lifetrail.be you can register and request access to the *Leptidea* project. Once approved, you can then perform the entire exercise. A user manual can be downloaded from the website.

The developed platform is online to train individual users improving their identifications of the *Leptidea* triplet whilst taking into consideration that in many cases, sampling for further study is needed to have reliable identifications.

The results of the new trainees will be included in the actual data and will serve for future developments of the platform.

Acknowledgements

We are grateful to Vlad Dincă and Roger Vila (IBE, Butterfly Diversity and Evolution Lab, Barcelona, Spain) for kindly providing the wing vouchers for this study. We express our gratitude to Peter Russell and Vlad Dincă for fruitful discussions about the final draft before submission.

We are indebted to the following blind readers for their patience, perseverance and identifications: J. Couckuyt, D. De Backer, R. De Mol, K. Jonckheere, B. Maes, M. Mølgaard, M. Rowlings, L. Parmentier, S. Spruytte, M. Taymans, Ph. Van de Velde, A. Van Grimberge, P. Vantieghem, L. Verhelst, J. Vervaeke and Raphaël Windey.

References

Dincă V., Lukhtanov V., Talavera G. & Vila R. 2011. Unexpected layers of cryptic diversity in wood white Leptidea butterflies. — *Nature Communications* **2**: 324. DOI: 10.1038/ncomms1329.

Mazel R. 2012. Critères morphologiques de séparation des *Leptidea sinapis* L., *L. reali* Reissinger et *L. juvernica* Williams (Pieridae, Dismorphiinae). — *Revue de l'Association Rousillonnaise d'Entomologie* **11**(1): 1–9. http://www.proserpine.org/forum/Leptidea_Mazel_2012.pdf (accessed 27.i.2017).

Dincă V., Wiklund C., Lukhtanov V., Kodandaramaiah U., Norén K., Dapporto L., Wahlberg N., Vila R. & Friberg, M. 2013. Reproductive isolation and patterns of genetic differentiation in a cryptic butterfly species complex. — *Journal of Evolutionary Biology* **26**(10): 2095–2106.

Šíchová J., Volenikova A., Dincă V., Nguyen P., Vila R., Sahara K. & Marec F. 2015. Dynamic karyotype evolution and unique sex determination systems in Leptidea wood white butterflies. — *BMC Evolutionary Biology* **15**: 89. DOI 10.1186/s12862-015-0375-4.

Supplementary information accompanies this paper:

– Supplementary Table 1: http://www.phegea.org/Phegea/Appendices/Phegea45-2_Table-S1.pdf

– Supplementary Table 2: http://www.phegea.org/Phegea/Appendices/Phegea45-2_Table-S2.pdf

– Supplementary Table 3a-d: http://www.phegea.org/Phegea/Appendices/Phegea45-2_Table-S3.pdf