

Beitr. Ent.	Keltern	ISSN 0005 - 805X
55 (2005) 2	S. 499 - 503	27.12.2005

ECatSym - Electronic World Catalog of Symphyta : history and progress

(Hymenoptera)

ANDREAS TAEGER, STEPHAN M. BLANK and ANDREW D. LISTON

Summary

The content and technical development of the Electronic World Catalog of Symphyta (ECatSym) is briefly described. At present (April 2005) approximately 8.200 valid species group names and 850 genus group names are included. This represents significantly more than 90 % of the known world fauna. As well as taxonomic and nomenclatural information, the database contains comprehensive data on foodplants, distribution and references.

Zusammenfassung

Die technische und inhaltliche Entwicklung des Elektronischen Weltkataloges der Pflanzenwespen (ECatSym) wird kurz dargestellt.

Derzeit (April 2005) sind ca. 8.200 valide Artgruppennamen und ca. 850 valide Gattungsruppennamen enthalten, was einer Vollständigkeit von deutlich über 90 % entspricht. Neben taxonomisch-nomenklatorischen Informationen umfasst die Datenbank umfangreiche Hinweise auf Futterpflanzen, die Verbreitung der Taxa sowie Literaturquellen.

The need for a World Catalog

Even for experienced researchers on the Hymenoptera Symphyta, the extremely fragmented and contradictory literature dealing with the group is a severe hindrance to further progress. The following considerations also make the need for an up-to-date catalog of Symphyta apparent. For many species our knowledge is limited to a single or very few observations, not infrequently more than a hundred years old. Without cataloguing these data sources, the information will be lost. "Old" knowledge must be made accessible by application of modern taxonomy and nomenclature. If this is not done, information in older literature may easily be attributed to the wrong taxon as a result of overlooked homonymy, synonymy or historical misidentification of taxa. Current research can often build on already available knowledge, so that unnecessary, duplicate research may be avoided. Catalogs reduce the need for much "routine" work. Literature searches can be reduced to a minimum where a reliable catalog is available. It should not be forgotten that much important entomological work is undertaken by amateurs in their spare time, and, who has the "Zoological Record" at home?

The path to a World Catalog

The original arbitrary geographical limit set for the database (West Palaearctic) soon proved unsatisfactory. Many taxonomic problems can only be solved using the complete, worldwide literature. Further, many publications treating taxa which occur in the West Palaearctic refer also to areas outside this region. The conspicuous lack of published works on Symphyta of a monographic nature has increasingly hampered research on the group. In discussion with other specialists, the need for a modern World Catalog was strongly voiced.

Project history

The original intention in setting up a database for the Symphyta in 1989-1990 was to cope with the highly chaotic situation in the literature dealing with this insect group. The first stage of cataloguing was undertaken using the VCH-Biblio program (DOS). In a labourious operation, individual items of literature were systematically worked through and cross-referenced to key words (mostly taxonomic units). This procedure, using the literature sources as the "condensation point" for the data, was very effective during the first stage of data capture. In order to use the data for research, it was originally attempted to apply current taxonomy to the evaluated literature (e. g. using the currently accepted names, corrected placement of previously misidentified taxa and so on).

However, several weak-points in this approach soon became apparent. Although the interpretation of literature is obviously indispensable, this approach leads to the mixing of existing published (objective) data with secondary interpretative (subjective) material. Sooner or later it is no longer possible to identify the origin of these data. Connections between biological data and the insect names, distributional data and so on can not be rationally managed with this system. "Current systematics" are quickly superseded. It is essential to organise the data so that the application of the current classification leaves the original data unaltered.

As a result of the above problems, it was decided in 1991 that the database must be completely restructured. Data would no longer be bundled using the literature sources. The taxon was instead selected as the pivotal point of the new system, with emphasis on species- and genus-group names. "VCH-Biblio" is unsuited to this task. The choice fell to Borland Paradox 4.0 (DOS), a system also in use in a number of partner institutions (Museum Alexander Koenig, Bonn; Museum für Naturkunde, Berlin; Staatliches Museum für Naturkunde, Dresden), superseded by more modern versions of Paradox later.

The following conclusions were drawn from experiences made in the restructuring of the database in 1992. Major adaptation of databases, whether involving alteration of the technology used or of the content, are exceedingly time-consuming, particularly when they already contain data. On no account should one succumb to pressure to update the system (use of a newer software version) without good reasons. The development of the database "in house" proved worthwhile. This allowed speedy adaptation in response to new requirements. The database has a modular construction, thus facilitating the incorporation of new elements as well as the retrieval of information on particular topics.

As the result of our work towards a World Catalog, lots of taxonomic and nomenclatural discrepancies have shown up, which we have published in several works (Blank 1996, Blank et al. 1998, Blank & Taeger 1998, Saini et al. 2005, Taeger & Blank 1996). The database, representing the hidden backbone of the catalog, has also become indispensable for collecting and sorting data for other publications (e.g. Blank et al. 2001, Taeger & Blank 2004).

Data included in the catalog and their organisation

The planned scope of the catalog includes the following types of information:

- valid taxa (names of the species-group, genus-group and family-group)
- distribution (zoogeographic regions)
- reference to further information (identification keys, phylogenetic analyses, taxonomic actions)
- biology (hostplants, parasitoids).

The database in figures (April 2005)

8,200 valid species-group names

850 valid genus-group names

400 family group names

23,000 species- and genus-group names (variant spellings, including approximately 13000 basionyms)

8,000 literature references

65,000 cross-references between names and literature

6,000 hostplant records derived from 13000 literature records

35,000 locality records, of which approximately 15000 from literature sources

20,000 sets of distributional data based on approximately 35000 mentions in the literature

Work currently in progress

The first version of the systematic world catalog as part of the remit of GBIF-International has been published in 2005 (Taeger & Blank 2005) in cooperation with E. K. Groll (Müncheberg), A. D. Liston (Müncheberg), A. Shinohara (Tokyo), D. R. Smith (Washington), and M. Wei (Changsha). The content will be completed and improved in steps by the incorporation of additional literature references. The taxonomic part has been available since September 2005 at http://www.zalf.de/home_zalf/institute/dei/php/ecatsym/ecatsym.php.

As part of the Fauna Europaea project 11,000 datasets for the occurrence of sawfly species in European countries (or subdivisions of these) have been published on <http://www.faunaeur.org> (Taeger & Blank 2004). An improved checklist providing access to the underlying references is being prepared (Blank & Taeger 2005). The GBIF-Germany project "GISHym: Global Information System on Hymenoptera" is working on a database of the primary sawfly type specimens held by German museums. All these datasets are being included in the ECatSym database. Photographs of types will be managed as an additional resource, but technical limitations inherent to Paradox at present preclude their direct inclusion.

Future possibilities

A principal goal for the near future is the digitalisation of literature on Symphyta. For this purpose publications are scanned, processed with the OCR software ABBYY FineReader and saved as PDF with "text under image". Currently we hold 1,800 publications as PDF documents, which equals roughly 20% of the ECatSym references. Search commands, e.g. of AdobeReader, provide an easy and direct access to primary sources.

The inclusion of larval hostplants, relationships to parasitoids and more detailed faunistic data shall increase in steps the amount of information available in the ECatSym online database. Finally we intend to use the constantly expanding database as an "information centre" on sawflies. New data are currently captured in a database version running on the

server of the DEI / ZALF, and a static version is displayed on the web. For practical reasons the revision of the database through the internet is desirable, since this will enable active access during visits to external institutions and external specialists can be immediately included in the improvement process.

The provisions of the ICZN include a ruling to propose the draft for a List of Available Names in Zoology (ICZN 1999, Art. 79). The ECatSym database appears to be the suitable tool for this process.

Acknowledgments

The realisation of the world catalog has been enabled by seedmoney from GBIF (Copenhagen) for the project "ECatSym: Electronic World Catalog of Symphyta (Insecta, Hymenoptera)" in 2004-2005. Additional funding has been provided by the Federal Ministry of Education and Research (BMBF, funding ID 001001LI02045) in course of the GBIF-D project "GISHym: Global Information System on Hymenoptera" (part of the German GBIF sub-node Evertebrata I), which has worked hand in hand with the ECatSym project.

References

- BLANK, S. M. 1996: Revision of the sawflies described by Lothar Zirngiebl. (Preliminary studies for a catalogue of Symphyta, part 2.) (Insecta, Hymenoptera, Symphyta). - *Spixiana*, München 19 (2): 195-219.
- BLANK, S. M.; DETERS, S.; DREES, M.; JÄNICKE, M.; JANSEN, E.; KRAUS, M.; LISTON, A. D.; RITZAU, C. & TAEGER, A. 2001: Symphyta. - In: DATHE, H. H.; TAEGER, A. & BLANK, S. M. (eds.): Verzeichnis der Hautflügler Deutschlands (Fauna Germanica 4). - *Entomologische Nachrichten und Berichte*, Dresden Beiheft 7: 8-27.
- BLANK, S. M.; SHINOHARA, A. & TAEGER, A. 1998: Revisionary Notes on Pamphiliid Sawflies (Hymenoptera, Symphyta: Pamphiliidae). Preliminary studies for a catalogue of Symphyta, part 3. - *Deutsche Entomologische Zeitschrift*, Berlin 45 (1): 17-31.
- BLANK, S. M. & TAEGER, A. 1998: Comments on the taxonomy of Symphyta (Hymenoptera) (Preliminary studies for a catalogue of Symphyta, part 4). Pp. 141-174. - In: TAEGER, A. & BLANK, S. M. (eds.): Pflanzenwespen Deutschlands (Hymenoptera, Symphyta). Kommentierte Bestandsaufnahme. - Goecke & Evers, Keltern.
- ICZN [International Commission on Zoological Nomenclature] 1999: International Code of Zoological Nomenclature. - 4th ed., London, 306 pp.
- SAINI, M. S.; BLANK, S. M. & SMITH, D. R. 2005: Checklist of the Sawflies (Hymenoptera, Symphyta) of India (Preliminary studies for a catalog of Symphyta, part 5). - In: BLANK, S. M.; SCHMIDT, S. & TAEGER, A. (eds.): Recent Sawfly Research - Synthesis and Prospects. - Goecke & Evers, Keltern. (in print)
- TAEGER, A. & BLANK, S. M. 1996: Kommentare zur Taxonomie der Symphyta (Hymenoptera) (Vorarbeiten zu einem Katalog der Pflanzenwespen, Teil 1). - *Beiträge zur Entomologie*, Berlin 46 (2): 251-275.
- TAEGER, A. & BLANK, S. M. 2004: Fauna Europaea - Hymenoptera, Suborder Symphyta. - <http://www.faunaeur.org>
- TAEGER, A. & BLANK, S. M. 2005: ECatSym - Electronic World Catalog of Symphyta (Insecta, Hymenoptera). - version 1.0 (July 1, 2005), ECatSym Online Service Müncheberg, [http://www.zalf.de/home_zalf/...](http://www.zalf.de/home_zalf/) URL

Author's addresses:

DR. ANDREAS TAEGER; DR. STEPHAN M. BLANK & ANDREW D. LISTON
 Deutsches Entomologisches Institut (DEI)
 Leibniz-Zentrum für Agrarlandschaftsforschung (ZALF)
 Eberswalder Strasse 84
 D-15374 Müncheberg, Germany
 e-mail: taeger@zalf.de, blank@zalf.de, liston@zalf.de

Deutsches Entomologisches Institut

Beiträge zur Entomologie

CONTRIBUTIONS
TO ENTOMOLOGY

Volume 55 - 2005

Chefredakteur/Editor in Chief: Prof. Dr. HOLGER H. DATHE
Verantwortlicher Redakteur/Assistant Editor: Dr. ECKHARD K. GROLL
Redaktion/Editorial Board: CORNELIA GRUNOW

Redaktionsbeirat/Advisory Board:

Dr. P. VAN DOESBURG, LEIDEN; Prof. Dr. W. DUNGER, GÖRLITZ;
Dr. R. GAEDIKE, Bonn; Prof. Dr. B. KLAUSNITZER, Dresden;
Prof. Dr. H. MALICKY, Lunz; Dr. W. MEY, Berlin;
Dr. G. MEDVEDEV, St. Petersburg; Prof. Dr. J. OEHLKE, Eberswalde;
Dr. G. PETERSEN, Görzdorf; Dr. K. SATTLER, London;
Prof. Dr. W. SAUTER, Illnau; Prof. Dr. U. SEDLAG, Eberswalde

Goecke & Evers, Keltern