POR-LINNAEUS : The application of interactive multimedia software for species data storage and computer assisted identification of Porifera

by Rob W.M. VAN SOEST, Peter H. SCHALK, Kathleen SMITH, Bernard E. PICTON, Marc BRUGMAN, Maria Cristina DIAZ, Miranda L. SANDERS, Wallie H. DE WEERDT, & Klaus RÜTZLER

Abstract

Porifera species data storage and identification using ETI's multimedia software package Linnaeus II were tested in several regional projects (Western Europe, Central West Atlantic and Papua New Guinea). Multimedia species data storage (texts, pictures, films) and retrieval with this software program appear to be easy to use, through pull down menus, button and text clicking, and uniformly applicable through standardized formats. Identification using multiple keys and a diversity of multi-media characters may be easily customized for different user groups. Regional or taxonomic monographs are assembled on CD-ROM for cheap and easily accessible dissemination. The stored data will also be made accessible on line.

Keywords : species-database, multimedia, IdentifyIt™, Linnaeus I1™.

Résumé

L'efficacité de stockage de données relatives à des espèces d'éponges et l'efficacité de leur identification ont été testées en utilisant le programme multimédia "Linnaeus II", produit par ETI, pour plusieurs projets régionaux : Europe de l'Ouest, Atlantique Central Ouest et Papouasie Nouvelle-Guinée. Avec "Linnaeus II", le stockage et l'accès de textes, d'images et de films sont facilités par l'usage de menus déroulants, de boutons et d'hypertextes, ainsi que par l'usage de formats standardisés. Le système d'identification par des clés aux entrées multiples et une variété de caractères multimédia peuvent être adaptés aisément pour différents groupes d'utilisateurs. La publication de monographies multimédias relatives aux éponges, organisées sur bases géographiques ou taxonomiques, est réalisée sur CD-ROM pour une diffusion aisée et économique. Les données seront aussi accessibles "on-line".

Mots-clés : banques de données, multimedia, IdentifyIt™, Linnaeus II™.

Introduction

For the study of sponge biodiversity and for monitoring changes in sponge occurrence which may be indicative of changes in our biological environment, high quality taxonomic and distributional data are imperative. Modern multimedia computer techniques, allowing nearly unlimited storage of graphic information in addition to text, will facilitate distribution of easy accessible sponge data. There is an increasing need for accessible and reliable information on all known species on the basis of user-friendly biodiversity databases (BRAATZ et al., 1993). A universal tool for biodiversity documentation is needed to further science and make existing knowledge faster accessible. A powerful answer to this growing demand for directly accessible biological data is a universal standardized computer based biodiversity information system (LOS, 1994). This system must be flexible enough to cope with, and incorporate, the many small scale databases in existence and must have the potential to produce (regional) master databases (BISBY, 1993). For determining species composition, fast, standardized and universal identification tools are needed based on expert knowledge to ensure high quality of data. For monitoring studies it is evident that good and reliable identifications are imperative to be able to compare datasets between different working groups and geographical areas, especially when one aims at comparing biodiversity values (RAVEN & WILSON, 1992). A new computer approach for biodiversity surveys is the answer.

The Linnaeus II software package for biodiversity documentation developed by ETI is an answer to these needs as it presents a universal tool for biologists to document biodiversity, create interactive identification keys and to exchange standardized data. Regularly updated electronic monographs on CD-ROM will form the basis for modern, fast accessible and compact libraries.

By using computer tools and modern interactive multimedia technology it is possible to create easy to use and cheaply updatable identification and information systems (ESTEP et al., 1989; SCHALK, 1992; ESTEP et al., 1993) that can be distributed to biologists worldwide. By including simple Geographic Information Systems (GIS) that run on PC's, distribution data can be stored and compared in an
interactive way as well. On line service through WWW is being developed and this will give easy access of biodiversity information to users worldwide. In this article we discuss the possibilities and application of such a system, as a basic tool for biodiversity documentation in the marine Porifera. The Porifera are one of the remaining phyla in which biodiversity data have hardly been accumulated. At the recent Brussels workshop several initiatives have been announced and steps to arrive at a unified names-, records- and specimen database have been taken. Next to these data, at present access to Porifera species information and identification keys is severely limited by the fact that literature is scattered over a vast amount of sources. Exchange of sponge biodiversity data between various disciplines and scientists is hampered by the lack of universal documentation tools and the lack of continuous active communication among taxonomists working in the same biogeographic areas. ETI's Linnaeus II software package Linnaeus II (including IdentifyIt) is here presented as a tool for a Porifera species documentation and identification. ETI's electronic accessibility (through FTP and WWW) is available for communication between taxonomists of various projects and networks.

The Expert center for Taxonomic Identification

The Expert center for Taxonomic Identification (ETI) is a non-profit UNESCO associated organization at the University of Amsterdam "for and by taxonomists". The center was set up in 1990 with support from the Dutch government, UNESCO and the University of Amsterdam, with the aim to develop novel interactive multimedia software for biodiversity documentation and species identification. ETI has developed a user friendly, easy to handle, universal biodiversity software shell (named Linnaeus II) that allows specialists to store information on species in text, pictures, videos and sounds in a standard and exchangeable format using their own PC (SCHALK, 1992; ESTEP et al., 1993; SCHALK & LOS, 1994). More importantly, this shell allows for the construction of "smart" (expert) identification systems giving both specialists and laymen fast access to the multimedia species information. The Linnaeus II system has various help functions on-line, describing the character/states in text and graphics as well as instructions how to use them. This will open up taxonomic keys also to relative laymen or students and these programs may also be used as a modern teaching tool. The Linnaeus II system was designed to be used on relatively simple Macintosh and Windows PC's. It is an excellent and user friendly tool for biodiversity documentation. The Linnaeus II program contains six major sections (Fig. 1): databases for species information and for higher taxa, programs to build identification tools (IdentifyIt and Linnaeus Key), a references database (with all cited literature), a glossary section (with all terms defined, explained and illustrated) and an introduction section (with general information on the group and methodology). A seventh major section, MapIt (for distribution data), is being developed. All parts are fully multimedia and may include texts, photo's, sounds and videos. The simple graphic user interface and hyperlink system guarantees easy point-and-click navigation through the various parts of the program.

Until now (1995), ETI released five interactive multimedia titles on marine biology: Fishes of the Northeastern Atlantic and Mediterranean (±1300 species), Sea Cucumbers of Northern Australia (±90 species), Marine Planarians of the World (±80 species), Marine Lobsters of the World (±150 species), Sea Mammals of the World (±130 species). Also other fields are being covered, for example: fungi, insects, plants, and protoctista. The CD-ROMs can also contain regional data sets: e.g. higher plants of Indonesia or bats of India. Educational (student) titles are published as well e.g. "Five Kingdoms, a key to the phyla of life", and the "Protoctista glossary". All these information- and identification programs are disseminated on compact disk (CD-ROM) for both Macintosh and Windows platforms. As they all have the same "Linnaeus II" format, their user interfaces and program structures are similar, warranting the same easy access to every program for users. Moreover, as the data are stored in a standard format, combination of the datasets is possible and users can construct an electronic library that fits their needs. CD-ROMs can contain huge sets of data (up to 650 Mb), including thousands of full color pictures and drawings, but also instructive videos. CD-ROMs are cheaper to produce than printed works, especially when full color images are needed, and can therefore be regularly up-dated when data sets are extended. Easy distribution (mailing costs are minimal) and a low price (productions are subsidized) guarantees that all scientists can have access to the most up-to-date tool for their research.

Por-Linnaeus : application of Linnaeus II software to marine Porifera

The Linnaeus II software has been used by a group of sponge taxonomists to explore its aptness for Porifera biodiversity data. Four projects were started in the course of 1994: "Western European Sponges" (WES) with contributors R.W.M. VAN SOEST, B. PICTON & W.H. DE WEERDT, "Central West Atlantic Sponges" (CWAS), with contributors K. SMITH AND K. RÜTZLER, and "Papua New Guinean Sponges" (PNGS) with contributor M.C. DIAZ (cf. DIAZ, this volume). A separate initiative combining taxonomic and natural products information was taken by M. SANDERS. These groups of contributors will increase
as recently (1995) several other specialists have volunteered to participate. The feasibility of projects on sponges of other parts of the world is being studied.

Here, already a brief survey of the first results is presented with details on the structure of the entered data and the construction of the identification tools. Currently, about 170 species have been entered and are in various stages of completion. The WES project is based for a major part on the Marine Conservation Society's "Sponge IV" and "Sponge V" guides (ACKERS et al., 1985, 1992). The CWAS project is rooted in the Smithsonian Carrie Bow Cay Coral Reef and Mangrove Project (cf. RÜTZLER & MACINTYRE, 1982). The PNGS project is inspired by Californian natural products chemists working together with taxonomists (e.g. CREWS et al., 1994), and the same group also initiated the natural products project (cf. below).

Like all Linnaeus II organism databases Por-Linnaeus contains the above mentioned six major sections (or "stacks"). For the three projects (WES, CWAS and PNGS) three of these sections are identical and shared (Introduction, Glossary and Higher Taxa). The other three sections (Species Cards, References and IdentifyIt) are custom-built for the respective projects, although the IdentifyIt files dealing with classes and orders will be built to serve all sponge projects, regardless of area.

Species Cards and Higher Taxa Cards

The Species Card database can be accessed by a mouse click on its button in the Navigator card of the program. In its "Overview" mode the species card looks like a file card with a default picture of the species on the right and an overview description on the left (Fig. 2A). The overview in general contains a summary of the characteristic features of a species with emphasis on macro morphological characters. In this overview mode it is easy to browse through the species cards stack for quick comparisons of species. The species cards are arranged in order of entry, but may be sorted in alphabetical or taxonomical sequence for finding a certain species card. Next to the default picture and an overview description the
Halichondria panicea (Pallas, 1758) is an intertidal or shallow-subtidal, thickly encrusting, massive or occasionally branching sponge, with typical volcano-shaped oscular chimneys. Colour green or yellow. It emits a characteristic disagreeable smell, said to be similar to that of "carbide". Surface smooth. Consistency firm, texture crumb-of-bread (hence its popular name "crumb-of-bread sponge" and its Latin name "panicea"). Ectosome detachable. Spicules oxeads only. One of the commonest intertidal sponges of North West Europe, but occurring all over the Holarctic region.

Fig. 2A. - The species database of the program consists of various "cards", here illustrated by the common Halichondria panicea. The overview card is for quickly comparing species by browsing through the stacks.

species card also contains a "Details" section. A click on the "Detail" button in the lower left of the overview leads to the "Detail" mode of the species card, from which the following information fields can be accessed (Fig. 2B):

A. Taxonomic description (the eye icon)
B. Classification (the cladogram icon)
C. Synonyms (the mask icon)
D. References (the book icon)
E. Multimedia, i.e. pictures, sounds, films (film camera icon)

A. The field "Taxonomic Description" has the following default paragraphs to ensure uniformity of species descriptions:
- Color
- Shape, size, surface and consistency
- Spicules
- Skeleton
- Reproduction
- Additional information such as cytology, chemistry, etc.
- Ecology
- Distribution
- Type specimen information
- Remarks (including comparisons with similar or related species)

The text in the Overview and Taxonomic Description is "hot" through hyperlinking: clicking on an unknown word starts a search in the glossary section of the program (cf. below). Any word can also be "found" by prompting the menu and typing the desired word in a dialog box.

B. The second field is "Classification", where all higher taxa names associated with the species are given. These names are hyperlinked to give access to the higher taxa database (cf. below): clicking on a name leads to the higher taxa card of that name.

C. The next field is "Synonyms" where all synonyms, common names and citations of the species are listed. Synonyms can also be searched for throughout the whole Species Card stack by prompting the menu and typing the desired name in a dialog box.

D. The "References" field holds all references of the species cited in the various fields. References are linked to the "References" stack (cf. below).
Colour: basically light orange-yellow or pale yellowish green. However, intertidal specimens exposed to the light may be dark greyish green, presumably due to microsymbionts. Shape, size, surface and consistency: quite variable in shape, related to exposure to water movement. Specimens growing in intertidal localities exposed to the full oceanic surf may be entirely smooth with barely visible oscular chimneys. More intermediate environments show the typical volcano-shaped chimneys up to 4 or 5 cm high. Oscules are relatively large, conspicuous, 2-4 mm in diameter. Sponge body may be up to 25 cm thick and 60 cm across, but much smaller specimens are the rule. In localities with strong currents specimens may grow out to form longer oscular chimneys closely adhering and verging toward palmate forms. In deeper or stagnant waters, with absence of wave surge, specimens may form masses of anastomosing branches, with oscular chimneys here and there on the branches. Consistency is firm, compressible, easily torn.

Spicules: oxacete only; size in specimens from Western Europe: 124-482 by 2-15 μm. Average sizes are: 300 by 7 μm (Holland), 280 by 8 μm (France), 360 by 9 μm (Ireland). Several

Fig. 2B. - The "detail card" can be reached by a button push (on the "Detail" button bottom left) and gives access to different database fields (button icons on the upper right): taxonomic description (eye icon, with chapters on taxonomy, distribution, reproduction, ecology, comparative remarks), classification (cladogram icon, giving systematics and access to higher taxa information), synonyms & common names (mask icon), species specific literature (book icon), and media clips (camera icon, listing pictures, sounds and video files that can be accessed by clicking).
discorhabd

Didiscus antitdidiscus
Didiscus placospingioides
Didiscus aceratus

Fig. 3. - Card from the Glossary stack illustrating the term "discorhabd". Each card contains a term and its synonyms, its definition and almost always one or more pictorial illustrations. Items in the Glossary can be reached either by clicking on the term where it occurs in descriptions in Species and Higher Taxa Cards or through clicking on the Glossary button in the Navigator and asking the pull-down menu to "find" a term.

All text in Por-Linnaeus is "hot" through hyperlinking: clicking on an unknown word starts a search in the glossary section of the program. Currently about 275 terms, including names of spicules, skeletal structures, anatomical features, etc. are defined and most are illustrated by one or more pictures taken from handbooks and authoritative texts. Prominent sources are the Thesaurus of Sponge Morphology [BOURY-ESNAULT & RÜTZLER (Eds), in prep.], Atlas of Sponge Morphology (DE VOS et al., 1992) and WIEDENMAYER's (1977) glossary. Most species so far included in Por-Linnaeus are Demosponges, and accordingly the glossary at present is biased towards that class. Eventually, it will be extended to encompass terminology for all Porifera. It may then serve as a general tool for all future sponge biodiversity projects using Linnaeus II software and can also be released as a separate sponge term glossary tool on CD-ROM.

The References stack is a literature database with fields for title, author, journal etc., and fields for keywords. The references are linked to species and higher taxa cards, and these links are listed in the keywords field.

The Introduction stack contains general information on sponges, taken from authoritative sources (e.g. HARTMAN, 1982). Cards in the stack include "General info", "Anatomy", "Cells", "Feeding", "Reproduction", "Predators", "Skeleton", "Spicules", "Taxonomy", "Thick Sections" and "Spicule Preparation". A commented video containing underwater images of Indonesian sponges is added to show sponges in their natural habitats.

A PC based GIS system "MapIt" is currently being developed by ETI to store distribution data. MapIt will enable the user to overlay distributions of several species, study species diversity in a certain area, or to compare distribution patterns. MapIt will have various levels of resolution to enable storage of distribution patterns in detail as well as on a worldwide basis.
Identification of the species stored in Por-Linnaeus can be achieved by clicking the Identifylt button in the Navigator. Identifylt (Fig. 4) is an identification program using the matrix approach, i.e. it scores hits in a matrix of taxa and their character states. Instead of a rigid binary choice system that one has to follow step by step, in Identifylt the user has a free choice to select any character. Every character has a number of states, one or more states may be chosen, or even "unknown" may be selected. If a character state is uncertain, it may be entered as "unknown" or as "multiple state", or it may be ignored. Characters and states are linked to descriptive texts (both short descriptors indicated as "A" and long text explanations indicated as "T"), but also where appropriate to pictures (indicated with the camera icon, cf. Fig. 4A), which can be called up by a button click (Fig. 4B). Once a choice of character states has been completed, the program sorts the species list in order of probability, shown as hit percentages, i.e. the number of matching character states as a percentage of the total number of character states entered. A click on the species name leads to the species card of the species for a quick check, after which one may return to the identification program. It is also possible to compare species in the "hit list" for both common and exclusive characters (by clicking the "Compare" button). This way of identifying is superior to the conventional, printed, and rigid dichotomous keys, as it allows the user to start with any known or available character. Thus, even damaged or incomplete specimens may be identified as the system also works with incomplete character/state sets and answers with a list of probable species.

Por-Linnaeus contains several types of Identifylt trial files directed at different user groups. There is a key to the classes of Porifera fully pictorial or with lengthy text explanations for each character state. This is intended for the non-specialist who wishes nevertheless to achieve an accurate identification. Likewise, there are keys for the orders - so far only finished for the Demospongea (Fig. 4A) - which are also extensively supported by picture characters (Fig. 4B) and descriptive explanations. Once an order has been keyed out, identification to a particular species is made through simple text character state choices. As a first case the Western European Haplosclerida recently revised by De Weerdt (1985, 1986) were completed and tested out. With the key any of the 16 closely similar species may be identified reliably.

A different approach, intended for a quick identification using a minimum of macro morphological characters, is the picture key based on Identifylt, developed for Papua New Guinean sponges (cf. Diaz, this volume), but applicable with some adaptation to sponges from other areas. This is useful for identifying morphologically conspicuous species in a limited, well-known area. Users of such a key may be chemists or ecologists, studying sponges in such a limited area, who lack knowledge or technical means to study microscopical structures. By offering these scientists a limited choice of habit characters (shape, size, oscule morphology, color, surface characteristics, consistency), it is often possible to let them identify sponges with some reliability. First results indicate that this may be a valuable tool. Final species identifications, however, would require the addition of microscopical characters (skeletal elements, choanosomal characters).

Lastly, a trial key was developed in which the chemistry of sponges is used as a means for identification. Classes of compounds as well as actual chemical structures (available as pictures) are used in the key. When completed this is potentially a valuable tool for chemists to find which sponge taxa are known to contain certain chemistry and through their species cards to check on habits and other macro morphological characters.

The multimedia identification software development by ETI is still in progress and new versions are forthcoming. In the most recent version of Linnaeus II, a traditional dichotomous key builder will be available. Identifylt and the Linnaeus Key may both be used to get access to the species database. A new "shell" for an entirely picture based identification is under construction and is expected to be released before the end of 1995.

User friendly data entry, storage and retrieval

Entering data in the species cards and constructing the keys is controlled from pull down menus and keyboard commands. Importing text into species cards is a simple matter of cutting and pasting digitized text from disc or scanning in printed text with simple optical character reading software. All fields of the species cards are scrolling fields allowing unlimited data entry. Pictures of most existing formats can be imported in a default format (for Macintosh this is the PICT format and for Windows TIFF), preferably by way of a photo-imaging program which allows insertion of texts, arrows, etc. In Identifylt, special functions like "cloning" allow the user to duplicate a species record and a selected number of fields, which simplifies the making of a new record of an almost similar species. Retrieval of information as described above is mostly a matter of clicking buttons, but may also be controlled from pull down menus or keyboard commands.

Linnaeus II was built using extensive feedback from ETI's international taxonomy networks now including about 500 specialists. It has been released for Macintosh platforms, a similar Windows version will be available in 1995. ETI also aims for incorporation of existing databases. The whole system is structured in such a way that data can be transported to other formats when needed. This also guarantees sufficient flexibility for the future.
Fig. 4A and 4B - Identify mode of the IdentifyIt file "Demosponge orders". The IdentifyIt part of the Linnaeus II program simplifies the identification process. The user may select any character from the scrolling list on the right hand top side of the card, for instance microscleres (Fig. 4A). Clicking on the character leads to a card listing all its states (e.g. euaster, microxea, discorhabd, etc.). One or more states may be selected by clicking on the name. Picture characters (indicated by the camera icon) have their states illustrated (Fig. 4B). After the choices of character states have been made (these are listed in the right hand corner, Fig. 4A) the program updates the taxa list (left side of the card in Fig. 4A) in order of probability (percentages in front of taxon name). The taxon on top of the list (in the example the order Homosclerophorida) fits best with the scored character states. Further characters of the taxon can be inspected by clicking on its name, which leads in the example to the Higher Taxa card describing the Homosclerophorida. Further identification to species can be done by accessing the Homosclerophorida IdentifyIt file. Similar IdentifyIt files exist for all the species of all the orders.

International Sponge Taxonomy Networks

Creating this type of electronic identification and information tools is an experts job. To build a biodiversity database ETI works together with UNESCO to set-up and support both regional and international networks of specialists who use the Linnaeus II software shell for local database building. Working with these networks ensures standardization and quality control of the data. Specialists who become an ETI partner receive the Linnaeus II software package for Biodiversity Information free of charge. The specialists send back a copy of the entered species data for back up in ETI's World Biodiversity Database on a yearly basis. At ETI the various information sets are compiled and submitted to a review procedure similar to scientific journals. Completed sets (on a particular taxonomic group or geographic area) are disseminated on CD-ROMs. The CD-ROMs will be free for the authors who composed the programs, while others (institutions, libraries, amateur users, etc.) pay only the costs of production since the CD-ROM productions are subsidized. Important is that authors share their copyrights on the data with ETI. All authors retain the right to use their data for any other purpose (including publications of articles and books or digital versions thereof).
Sponge taxonomists working with Por-Linnaeus are at present organized in two centers of activity, viz. at the Institute of Systematics and Population Biology of the University of Amsterdam and at the National Museum of Natural History (Smithsonian Institution) at Washington. Contributing specialists can send in and retrieve Por-Linnaeus data from these centers by File Transfer Protocol (FTP), allowing unlimited transfer of both texts and pictures. To ensure uniformity of data entry in the various projects of Por-Linnaeus, this will be guided by R.W.M. VAN SOEST (Amsterdam) and K. SMITH (Washington). Future potential Por-Linnaeus projects can count on the technical support of ETI in porting biological data from existing species databases to the Linnaeus II shell, or general support with scanning and entering data in printed form. The datasets will be regularly updated by the international taxonomy networks. Approximately one thousand specialists all over the world have stated to work on this biodiversity database that is published literary "bit by bit" on CD-ROM. ETI stores all information in one big World Biodiversity Database in Amsterdam which will be available on-line to all scientists in a later stage.

Acknowledgments

Christine MORROW (BioMar Unit, Dublin) contributed to Por-Linnaeus by allowing the use of her spicule drawings. Gideon GIJSWIJT (ETI, Amsterdam) assisted in the final stages of preparing the Por-Linnaeus demo.

References


BOURY-ESNAULT, N. & RÖTZLER, K., (Eds), in prep. Thesaurus of sponge morphology. Smithsonian Institution Press, Washington DC.


Rob W.M. VAN SOEST
Institute for Systematics and Population Biology (Zoological Museum)
University of Amsterdam
PO Box 94766, 1090 GT Amsterdam, The Netherlands

Peter H. SCHALK
ETI - University of Amsterdam
Mauritskade 61, 1092 AD Amsterdam, The Netherlands

Kathleen SMITH
National Museum of Natural History
Smithsonian Institution, Washington, DC 20560, USA

Bernard E. PICTON
BioMar, E.S.U.
Trinity College, Dept. of Environmental Sciences
Dublin 2, Ireland.

Marc BRUGMAN
ETI - University of Amsterdam
Mauritskade 61, 1092 AD Amsterdam, The Netherlands

Maria Cristina DIAZ
Marine Sciences, University of California at Santa Cruz
273 Applied Sciences Building, Santa Cruz
CA 95064, USA

Miranda L. SANDERS
Marine Sciences, University of California at Santa Cruz
273 Applied Sciences Building, Santa Cruz
CA 95064, USA

Wallie H. DE WEEERDT
Institute for Systematics and Population Biology (Zoological Museum)
University of Amsterdam
PO Box 94766, 1090 GT Amsterdam, The Netherlands

Klaus RÜTZLER
National Museum of Natural History
Smithsonian Institution, Washington, DC 20560, USA