ICES 1990 PAPER C.M.1990/L:92
Ref. Mariculture Committee
Ref. Pelagic Fish Committee

PRESENTATION OF A BIOLOGICAL DATA BASE FOR SPECIES OF CURRENT OR FUTURE IMPORTANCE TO AQUACULTURE OR FISHERIES

Rainer Froese

Institut fur Meereskunde
Düsternbrooker Weg 20
D-2300 Kiel 1

present address:

ICLARM
M.C. P.O.Box 1501
Makati, Metro Manila 1299
Philippines

ABSTRACT

A database (named FISHBASE) is being developed at the International Center for Living Aquatic Resources Management (ICLARM), which summarizes global information about living aquatic resources (fish, crustaceans and molluscs) in a standardized form and which will be made available to institutions in developing countries. For all species, stocks, and strains relevant to fisheries or aquaculture, FISHBASE contains information comparable in scope to that normally provided in species synopses such as those published by FAO. FISHBASE provides not only fast access to information on a given species but also allows for comparative studies between species groups or geographical areas. Although the project aims at developing countries and at tropical species, it might as well be of interest to scientists and to educators, planners and policy makers in the ICES area. The paper describes FISHBASE and suggest that ICES should recommend to its members: A) to support projects which aim to include the species within the ICES area into FISHBASE, B) to support cooperation agreements between institutions in ICES member countries and ICLARM on specific topics of FISHBASE such as diseases, genetics, or population dynamics, and C) to coordinate similar data base projects in member countries with FISHBASE to avoid redundant work as well as the development of incompatible systems.

1 Introduction

Knowledge about fisheries and aquaculture is distributed in numerous textbooks and thousands of papers, lacking, however, standardization of terms and units. Bibliographic databases like Aquatic Sciences and Fisheries Abstracts (ASFA) provide access to keyterms, titles and abstracts, but lack structured data; hence, their users must have access to original literature to extract the information of interest. As institutions in developing countries cannot afford to maintain extensive libraries, scientists thus often lack such access, and even if possible, such data retrieval is very costly and time-consuming.

To help solve this problem, the International Center for Living Aquatic Resources Management (ICLARM) is developing a database (named FISHBASE) to summarize global information about living aquatic resources (fish, crustaceans and molluscs) in a standardized form to be made available to institutions in developing countries. Inputting of species is currently concentrated on Asian and African waters. This effort is supported by the Food and Agriculture Organization of the United Nations (FAO) and the Commission of the European Community (CEC). Although the project aims at developing countries and tropical species, it might as well be of interest to scientists (biologists, economists, environmentalists and sociologists) and to educators, planners and policy makers in the ICES area.

This paper presents FISHBASE and proposes several ways in which the project could be supported by ICES and institutions within ICES member states.

2 A description of FISHBASE

For all species, stocks, and strains relevant to fisheries or aquaculture, FISHBASE summarizes information comparable in scope to that normally provided in species synopses such as those published by FAO. FISHBASE provides not only fast access to information on a given species but also allows for comparative studies between species groups or geographical areas.

2.1 Taxonomy and distribution

FISHBASE contains valid scientific names and synonyms, valid FAO names, common names by country and species distribution by FAO areas and countries. This information is derived from FAO.

2.2 Morphology, identification, and museum collections

For eggs, larvae, and adults, FISHBASE contains morphometric measurements, meristics and detailed descriptions including pictures, which allow quick, easy and accurate identification without reference to costly taxonomic monographs or scarce expert advice (FROESE and PAPASISSI 1990). The data on museum collections include full descriptions of specimens, their location and reference numbers and stimulate the upkeep and expansion of such collections. This part of FISHBASE also draws upon and assists the conservation of archival material, including drawings and descriptions from publications dating back to the end of the last century.

2.3 Ecology

FISHBASE contains structured information about habitats, behavior, reproduction, food items, predators, competitors, and ecological parameters, which help environmental and related studies in the context of global change.

2.4 Population dynamics

The need for structured population dynamics data was a primary reason for establishing FISHBASE. All the important parameters related to catch, growth, mortality, and length-weight relationships are included by species and/or stock. In addition, time series data on catches are provided.

2.5 Aquaculture

FISHBASE is the first database to provide an organized and easily searchable structure to the highly heterogeneous data emerging from the rapidly evolving world of aquaculture. In addition to general data on the performance and tolerance of farmed species or strains, FISHBASE contains genetic data, time series data on production by country, information on breeding, hatchery and nursery systems, and on the farming systems used for growout.

2.6 Disease

Diseases are of increasing concern in aquaculture and fisheries. FISHBASE records the diseases reported for a species, stock, or strain, including their prevalence, symptoms, effects, treatment and prophylaxis. Symptoms are classified to allow diagnosis through FISHBASE (ACHENBACH and FROESE 1990).

2.7 Graphics

FISHBASE has a strong graphical component. It contains color images of eggs, larvae and adults for all species. It also contains distribution maps and drawings, explaining terms and definitions.

3 Mode of operation

FISHBASE is designed to run on low cost IBM-compatible microcomputers such as already exist in many institutions in developing countries. Special emphasis is given to user-friendliness.

Inputting is done mainly at ICLARM headquarters in Manila, supervised by ICLARM scientists, with the assistance of FAO and national institutions: e.g., Zoologisches Museum, Hamburg for museum collections and national Universities such as the University of the Philippines for national or regional information. Additional help is provided by the members of the

Network of Tropical Fisheries Scientists (NTFS) and the Network of Tropical Aquaculture Scientists (NTAS).

The entries for every species in FISHBASE will be examined by appropriate experts before its first release.

4 Progress to date

As of August 1990, substantial information on about one hundred species of finfish and nomenclature for about 800 other have been entered in FISHBASE. Entries for the crustaceans are beginning.

5 FISHBASE products

The main FISHBASE product will be a complete database stored on a CD-ROM laser disk for use with IBM-compatible microcomputers. Release of the first version is planned for the end of 1991. Updates will be available on an annual basis. Packages consisting of a microcomputer with FISHBASE and CD-ROM equipment will be available. Regional output will be distributed on standard diskettes and where appropriate as hard copy.

Training courses on the use of FISHBASE will be held in developing countries and at ICLARM headquarters in Manila.

FISHBASE is a research tool and a resource with which research can be done: comparative and interrelated studies. Such studies have already commenced using the data already entered (ACHENBACH and FROESE 1990; FROESE and PAPASISSI 1990).

6 Financial requirements

The core funding for the FISHBASE office, part-time inputting staff costs, and supervision by ICLARM scientists is provided by ICLARM. Additional funding for personnel and equipment is provided by FAO and the CEC. ICLARM is seeking additional support for coverage of the developing regions and special fisheries and aquaculture topics according to needs and donor interests.

7 Suggestion to ICES for support of FISHBASE

There are many possible forms in which ICES could support the further development of FISHBASE. In the first place, ICES should encourage its members to support projects which aim to include the species within the ICES area into FISHBASE. In addition, ICLARM is looking for scientists who would examine and complete the content of FISHBASE on specific topics such as diseases, genetics, or population dynamics. This could be done on the basis of cooperation agreements between institutions in ICES member countries and ICLARM.

Several institutions in ICES member countries are planning or developing similar data bases, e.g. on the local fish fauna. To avoid redundant work as well as the development of several incompatible systems, such projects should be coordinated with FISHBASE.

The author suggests on the basis of the outlined aims that the Biological Oceanography Committee recommends to ICES:

- To support projects which aim to include the species within the ICES area into FISHBASE.
- To support cooperation agreements between institutions in ICES member countries and ICLARM on specific topics of FISHBASE such as diseases, genetics, or population dynamics.
- To coordinate similar data base projects in member countries with FISHBASE to avoid redundant work as well as the development of incompatible systems.

Justification: FISHBASE provides instant access to information comparable in scope to that normally provided in species synopses. It allows for comparative studies between species groups or geographical areas. Such information is of interest to scientists and to educators, planners and policy makers in the ICES area. It is therefore, desirable to extend FISHBASE to the species within the ICES area.

8 References

ACHENBACH, 1. and R. FROESE 1990. Presentation of a database system for information on and diagnosis of fish diseases. ICES C.M.1990/F:72 13 pp.

FROESE, R. and C. PAPASISSI 1990. The use of modern relational data bases for identification of fish larvae. J. Appl. Ichthyol. 6: 37-45

A NEW PROCEDURE TO EVALUATE FISH COLLECTION DATABASES

by

FROESE, Rainer(l), BAILLY, Nicolas(2), CORONADO, Grace(l), PRUVOST, Patrick(2), REYES, Rodolfo, JR.(l) & HUREAU, Jean-Claude

(1) ICLARM, MC P.O. Box 2631, 0718 Makati City, PHILIPPINES
(2) Museum national d'histoire naturelle, Ichtyologie générale et appliquée, 43 rue Cuvier, 75231 Paris Cedex 05, FRANCE

ABSTRACT. - A procedure is presented that compares some of the information typically contained in fish collection databases (scientific name, identifier, locality, coordinates . . .) with reference information compiled in FishBase, a large biological database on finfish available on CD-ROM. The procedure detects possible errors in spelling, locality, and identification. It can be used to assign a reliability indicator to collection records, detailing the confidence in the given scientific name, identifier, area, country, and coordinates. For records with country assignment or with coordinates, various maps can be produced to visually detect possible errors in locality or identification, and to check for marine species being recorded from land, and vice-versa. Preliminary experiences in applying the procedure to subsets of several collection databases are presented.

RÉSUMÉ. - Les auteurs ont mis au point une procedure qui permet de comparer certaines informations habituellement presentes dans les bases de donnees de collections de poissons (nom scientifique, déterminateur, localité, coordonnées géographiques…) avec les informations presentes dans FishBase, une grande base de donnees biologique sur les poissons disponible sur CD-ROM. La procedure détecte les possibles erreurs d'orthographe, d'origine et d'identification. Cette procedure peut être utilisée pour attribuer un indice de fiabilite aux enregistrements de collections, en détaillant la confiance que l'on peut avoir dans le nom scientifique, le déterminateur, la zone géographique, le pays et les coordonnees. Pour les enregistrements qui font reference a un pays ou qui possèdent des coordonnees géographiques, des cartes peuvent ëtre produites afin de détecter visuellement des erreurs de localisation ou d'identification, et pour verifier que des especes marines ne sont pas enregistrées en eau douce et vice-versa. Des tests préliminaires d'application de la procedure a des sous-ensembles de plusieurs collections sont présentés.

Key-words. - Fish Collections, Databases, Reliability indicator.

The importance of museum collections for the study of biodiversity is now widely acknowledged. It is estimated that at least 3.8 million lots of fishes exist in North American museums (Poss and Collette, 1995). About 47% of the lots were computerized in North America in 1990, with a strongly increasing trend (Poss and Collete, 1995). In Europe the number of fish specimens held in museums was estimated at 7-8 million individuals in 1990 (Kottelat *et al.*, 1993), which may result in about one million lots if an average number of 8 specimens per lot is assumed. Additional large collections exist in Australia, Japan, South Africa and South America (Hureau, 1996). The total number of lots in the world may reach 10 million. About 60 collection databases were accessible through the Internet in October 1997.

The quality of the available information about the fish samples is highly variable. with some having been used in a recent revision by world experts, and others not having been looked at for more than 200 years. Most museums are understaffed and curators have a hard time just keeping up with the cataloging of recent acquisitions. Systematically re-identifying each specimen other than in the context of revisionary work seems not to be an option.

Thus, we propose here to compare the information available in typical collection databases with corresponding reference tables in FishBase to assign automatically a reliability indicator that can be used to identify misspellings and doubtful scientific names, likely misidentifications and range extensions, as well as records where the identification and locality are most likely correct.

MATERIAL AND METHODS

FishBase is a large biological database with key information on 17,600 species of finfish, as of October 1997 (Froese and Pauly 1997). FishBase 98 was released after this study and contained 54,000 names for 20,000 species (Froese and Pauly 1998). For the purpose of this study, FishBase 97 contained an OCCURRENCE table with 53,563 museum collection records from several museums in Europe and USA. It also contained extensive synonymies, distribution records by FAO area and by country, and a preliminary list of fish taxonomists, with indication of their family of expertise. This information was used to derive a combined reliability indicator for the stated scientific Name, Identifier, FAO Area, Country, and Coordinates (NIACC). The NIACC reliability indicator was assigned automatically to existing records, based on criteria explained below. For each category a number between 1 and 5 was assigned, with the following meaning:

Scientific Name

1 : unambiguous (in the synonymy of one species only);

2 : unambiguous, but has been misapplied (in more than one synonymy, species may be difficult to identify);

3 : ambiguous (matches no name in FishBase or points to more than one species, especially if no author was given);

4 : not available (e.g., for 'sp.');

5 : not yet evaluated.

Identifier

1 : expert of respective family (has published—as first author—a revision of the family or a genus within the family);

2 : expert (has published a taxonomic revision of another family);

3 : other identifier (has not published a revision as first author);

4 : not available (no identifier stated);

5 : not yet evaluated.

FAO area

1 : species known from stated FAO area;

2 : (category not used);

3 : species not known from stated FAO area;

4 : not available (no FAO area stated);

5 : not yet evaluated.

Country

1 : species known from stated country;

2 : species not reported from stated country, but list of countries for this species is incomplete in FishBase;

3 : species not known from stated country (and list of countries for this species in FishBase is complete);

4 : not available (no country stated);

5 : not yet evaluated.

Coordinates

1 : coordinates of locality are within (fresh) or adjacent to (marine) the geographic range of the stated country (the range is a rectangle of the nearest latitudinal and longitudinal degrees that include the country; for marine species, the range is extended by 4 degrees in each direction);

2 : (category not used);

3 : coordinates are not within or adjacent to the stated country;

4 : not available (no coordinates or no country stated);

5 : not yet evaluated.

The following reference tables in FishBase were used:

The **SYNONYMS** table with 46,318 names for 17,640 species, classified as: junior synonym, misidenfication, misspelling, original combination, new combination, questionable, other.

The EXPERTS table, which contains the names of currently 387 taxonomists for which the FishBase **REFRENS** table contains a publication that is classified as revision of a genus or a family, and of which they are the first author. The **EXPERTS** table also states the respective families of these revisions.

The FAOAREAS table with 35,678 records of FAO statistical areas from which a species has been reported, classified either as native, endemic, introduced, extirpated, reintroduced, or unclear.

The COUNTRY table, which states for each of the 17,640 species in FishBase the UN country names from which they have been reported. Note that currently such country lists are complete for only about 50% of the species in FishBase.

The COUNTREF table, which contains for each of the UN countries, as well as for a number of islands the nearest latitudinal and longitudinal full degrees that include the country.

RESULTS

Scientific name

About 83% of the scientific names were unambiguous synonyms (category 1). Eleven percent of the names referred to one valid species, but had also been misapplied to other species, thus having a higher chance for possible misidentifications (category 2). Four percent of the names could not be linked to any synonym in FishBase and thus may either be misspelled or not yet contained in FishBase, or point to more than one biological species, such as in the case of *Alectisindica*, which is a misspelling of *Alectis indicus* (Rüppell 1830), and also refers to Alectis indica (Cuvier 1833), which is a junior synonym of *Alectisciliaris* (Bloch 1787). Since authority names are often omitted in collection databases, the name could not be assigned to a valid species (category 3).

FishBase contains a routine to find misspellings in scientific names (Froese, 1997). Two percent of the names did not provide specific epithets (category 4).

Identifier

For the purpose of this study we created an EXPERTS database with 387 family experts. We standardized the names in 13,948 of the 16,025 records for which Identifier names were given in the FishBase OCCURRENCE table. Of these records, 11% were identified by family experts, 73% were identified by other experts, and 10% were identified by persons of unknown experience. Note that not all museums have a system in place that keeps track of subsequent identifications.

FAO area

About 61% of the lots stated FAO areas that were compatible with the established distributional range of the species (category 1). Eight percent gave areas outside the established range (category 3). A closer look at these records revealed three possible reasons: 1) a wrong FAO area had been assigned to the locality stated in the museum database; 2) the FAO area was compatible with the locality stated in the record, and thus was either a range extension or a misidentification; and 3) the area was correctly assigned and compatible with the established range, but the FAOAREAS table in FishBase erroneously did not contain a record. In this case, FishBase was corrected. About 28% of the records had no FAO area assigned to them (category 4), and three percent of the records could not yet be evaluated because the scientific name had no match.

Country

About 68% of the country names provided matched a country assigned to the species in the FishBase COUNTRY table (category 1). Twelve percent did not match a record in the COUNTRY table, but the list of countries for that particular species was known to be incomplete in FishBase (category 2). Four percent of the countries provided were not contained in lists that were complete for the species, and thus were either range extensions or misidentifications. In a few cases, countries were found to be missing from lists that were supposed to be complete, and this was corrected in FishBase. Fourteen percent of the records had no country name assigned (category 4), and two percent of the records could not yet be evaluated because the scientific name did not match a valid species (category 5).

NIACC

Five digits with 4 or 5 possible entries each allow for 2,000 possible combinations for the NIACC reliability indicator. Of these, only 301 were assigned by the algorithm in the current study. The most common combination was NIACC 14111, which was assigned to 37% of the collection records, indicating highly reliable records with so far unevaluated identifier. Similar good marks were NIACC 24111 for 3%, NIACC 12111 for 4%, NIACC 12121 for 2% and NIACC 11111 for 1.4% of the records.
Some combinations gave considerable insight in the evaluation of the collection record. For example, NIACC 11111 refers to a specimen with an unambiguous name, identified by a family expert, and reported from a locality, country and area within its established range. NIACC 11131 indicates a similar well identified specimen from a locality in a new country that is within its broader range, i.e., most likely a range extension (Table 1). NIACC 23331 indicates a species with a name that has been misapplied before, identified by an unknown identifier, and reported from a locality clearly outside the established range, i.e., most likely a misidentification. We think that these examples demonstrate why a '5 telling digits' indicator is more useful than a summarized single digit indicator. Several combinations related to the unfinished status of the evaluation process (see below), e.g., NIACC 14411 (11%), indicating that no FAO area were available, nor NIACC 14311 (2%), indicating that the name, country and coordinates were correct, but there was a probably erroneous mismatch of country and FAO area in the collection database.

DISCUSSION

One of the problems with museum collection databases is a lack of consistent standards. For example, identifier names are stated with one, two, or without initials, with leading or trailing initials, with spelled out prenames, with additional information such as years or collaborators, etc. For example, the unique list of 807 different entries in the Identifier field of the OCCURRENCE table contained 8 variations of what should have been 'Randall, J.E.'. Similarly, there are rarely any standards used for geographic areas and country names. In addition, these fields are usually filled manually be typing in the information, instead of selecting it from pick-lists, and consequently there are numerous typos. Cleaning up of these fields is the first task before the NIACC quality indicator can be assigned. The printout of an alphabetic list of unique entries in a given field is an efficient way to approach this work.

There are already some reliability indicators in use in collection databases. A system applied by Australian museums assigns 5 levels from 'Highly reliable identification' to 'Identification superficial', based on the 'taxonomic experience of the identifier, their knowledge of the group considered, and the amount of effort given to make the identification'. The quality level is assigned manually by data encoders or by the identifiers themselves (Williams *et al.* 1996). The MUSE collection database system which is used by several museums in North America has a Yes/No field for 'questionable' records, which refers to the confidence that the cataloger has in the identification. Also, locality records can be marked as 'proofed' if the cataloger is satisfied with the accuracy and quality of the data (Dave Catania, pers. comm.).

Our criterion for inclusion in the EXPERTS database (first author or a generic or family revision) was a pragmatic one, and we are waiting for feedback to improve on this. One could, for example, argue that also co-authors of such revisions should be considered experts of the respective family. However, we want to stress that the criterion for inclusion should be an objective one, allowing the computer to do the assignment whenever a new revision is entered in the FishBase REFRENS table.

An algorithm has been developed to plot coordinates using the WINMAP software that comes with FishBase (Coronado and Froese, 1997) and write back a file that states for every point whether it fell on 'land' or 'sea'. For example, of 282 occurrence points plotted for Chaetodontidae, 29 fell on land. Some of these were true mistakes, whereas others just reflected the fact that butterfly fishes occur close to the shore and that the often used accuracy of full geographical minutes is not sufficient to avoid such points being plotted on land. Point data by family can also be evaluated on screen, to find obvious outliers (see Fig. 1).

We realize that the reliability of the NIACC indicator itself depends on the quality and completeness of the respective reference tables in FishBase. However, these tables will continuously be improved the more they are used for this and other purposes, and we feel that even in their current status the frequency of errors was relatively low.

Table 1 lists all records that were classified as likely range extensions or new records for a country (NIACC 11131), with the possible new country indicated in the collection database and the distributional range given in FishBase. Countries in FishBase are assigned to a species if 1) they are mentioned in a distributional range, such as shown in Table 1, 2) they are included in a map showing the distribution of a species, or 3) there is a taxonomic reference explicitly naming the countries. The list of countries for a species is marked 'complete' in FishBase if the available sources allow such a statement. As for the records in Table 1, most countries are within or very close to the given range, making the addition of the new country to the established list highly likely. One record (Tahiti) is an omission in the FishBase country list, because Society Islands are stated in the range. Altogether the algorithm classified these records correctly.

Table 2 lists all records that were classified as probable misidentifications (NIACC 23331), again with the suggested country and the range given in FishBase 97. As can be seen from the text, these species were already known to be misidentified outside their area, or to be part of a species complex that needs further study and may be one species only. The algorithm thus classified these records correctly.

Acknowledgments. - We thank Alice Laborte for implementing the various procedures to assign the NIACC reliability indicator to the existing records. We thank the FishBase team for having assembled the reference tables used in this study. FishBase has been developed with support from the European Commission, DG VIII.

REFERENCES

CORONADO G. & R. FROESE, 1997. - The WINMAP software. *In* R. Froese and D. Pauly (eds.) 216-231. FishBase 97: concepts, design and data sources. ICLARM, Manila, Philippines, 256 p., with CD-ROMs.

FROESE R. 1997. - An algorithm for identifying misspellings and synonyms in lists of scientific names of fishes. *Cybium*, 21(3): 265-280.

FROESE R. & D. PAULY (Eds.), 1997. - FishBase 97: concepts, design and data sources. ICLARM, Manila, Philippines, 256 p., with 2 CD-ROMs.

FROESE, R. & D. PAULY (Eds.), 1998. - FishBase 98: concepts, design and data sources. ICLARM, Manila, Philippines, 294 p., with 2 CD-ROMs.

HUREAU J.-C., 1996. - Les bases de données de collections de poissons de France et dans le monde: revue historique. Biosystema 14 : Systématique & Informatique : 34-47.

KOTTELAT M, NIELSEN J.G. & H. NIJSSEN., 1993. - Survey of ichthyological resources in European museums and collections. Soc. Europ. Ichthyol., Paris, Frankfurt : 24 p.

POSS S.G. & B.B. COLLETTE., 1995 - Second survey of fish collections in the United States and Canada. Copeia, 1995(1) : 48-70.

WILLIAMS A., P.R. LAST, M.F. GOMON & J.R. PAXTON., 1996. - Species composition and checklist of the demersal ichthyofauna of the continental slope of Western Australia (20-35% S). Rec. West. Austral. Mus. 18 : 135-155.

Table I. - Collection records marked by the algorithm as probable range extensions. The suggested new countries are given in square brackets. The distributional range is stated s contained in FishBase 97. See text for discussion of the table.

Abudefduf bengalensis [Hong Kong]

Indo-West Pacific: eastern Indian Ocean, north to Japan, south to Australia.

Abudefduf sordidus [Palau]

Indo-Pacific: Red Sea and East Africa Hawaiian and Pitcairn Islands, north to Japan, south to Australia.

Apistogramma cacatuoides [Colombia]

South America: Suriname.

Bellator militaris [Cuba]

Western Atlantic: North Carolina to southern Florida and northern Gulf of Mexico in USA; south to Yucatan in Mexico.

Caelorinchus caribbaeus [Trinidad and Tobago]

Western tropical Atlantic from Cape Hatteras to northern Brazil. Absent in Straits of Florida, uncommon to the north and along Antillean chain.

Caelorinchus multispinulosus [China]

Southern Japan to East China Sea.

Caelorinchus occa [Antigua Barbados]

Central North Atlantic: from Florida Straits to northeastern South America (uncertain). One record from Bermuda. Atlantic: southern Africa (must be confirmed), from Faroe Channel to Cape Verde Is. (Ref. 3587).

Caelorinchus parallelus [New Caledonia (2 records)]

Indo-west Pacific: southern Japan, East China Sea, and the Philippines, but may extend into Indian Ocean, Australia and New Zealand.

Chromis weberi [China]

Indo-Pacific: Red Sea and South Africa (Ref. 4391) to line Is. and Samoa, north to southern Japan, south to New Caledonia; Palau in Micronesia.

Chrysiptera glauca [Cook Islands, Palau (2 records)]

Indo-Pacific: East Africa to the Line and Pitcairn Is.; Australia northwards to Japan; throughout Micronesia.

Chrysiptera rex [Hong Kong]

Indo-West Pacific: eastern Indian Ocean (Scott and Ashmore Reefs), Ryukyu Is., Taiwan, Philippines, Palau, Indonesia, New Guinea, New Britain, Solomon Is., Vanuatu, New Caledonia, and the Great Barrier Reef.

Cirripectes castaneus [Hong Kong, Kiribati]

Indo-West Pacific: Red Sea to Tonga, north to southern Japan; south to Lord Howe Is.; Palau, Ifaluk, and Kapingamarangi in Micronesia.

Cirripectes polyzona [Palau (3 records)]

Indo-Pacific: South Africa to Kiribati, north to Japan; south to Rowley Shoals and the southern Great Barrier Reef, throughout Micronesia.

Cirripectes quagga [Palau (2 records)]

Indo-Pacific: South Africa to Tanzania, east to Henderson Island, Pitcairn and the Hawaiian Is.; north to China; south to the Great Barrier Reef, Australia.

Cirripectes stigmaticus [Kiribati]

Indo-Pacific: from Mozambique to Kenya, throughout the Indian Ocean and western Central Pacific to the Marshall and Samoa Islands.

Dascyllus flavicaudus [Tahiti]

Southeastern Central Pacific: southeastern Oceania including Society Is., Tuamotu Is., Pitcairn Group, and Rapa.

Nezumia convergens [Ecuador (4 records)]

Eastern Pacific: ranges from the Gulf of California, Mexico south to Chile, including Cocos and the Galapagos Islands.

Pomacentrus emarginatus [Solomon Is.]

Indo-Australian: Waigiu (off west New Guinea) and Palau.

Stegastes nigricans [Viet Nam]

Indo-Pacific: Red Sea and the East Africa to the Tuamotu, Marquesas, and Line Is., north to the Ryukyu and Bonin Is., south to New Caledonia and Tonga; throughout Micronesia. Excluding the Hawaiian Islands (Ref. 7247).

Ventrifossa macropogon [Puerto Rico, St Kitts Nev. (2 records)]

Western tropical Atlantic from off Guyana into the Caribbean and the Gulf of Mexico, and in Atlantic off northeastern Florida. Common to the south of the Gulf of Mexico, but relatively rare in the Gulf and Gulf stream.

Ventrifossa mucocephalus [Haiti]

So far known only from the western Caribbean, the Straits of Florida off Cuba, and the Atlantic off northeastern Florida.

Ventrifossa nigrodorsalis [New Caledonia (5 records), Viet Nam]

Known from southern Japan, Taiwan Island, Philippines, and parts of Indonesia (Borneo, Halmahera). Slight morphological variation seen in specimens from Japan and Indonesia, but not sufficient to recognize additional taxa.

Ventrifossa petersoni [Myanmar]

Indo-Pacific Ocean.

Table II. - Collection records that were identified by the algorithm to be probable misidentifications (NIACC 23331). The countries stated in the collection records are given in square brackets. The descriptive text is the range given in FishBase 97.

Johnius dussumieri [Hong Kong]

Indian Ocean: from the southern coast of South Africa (not the Red Sea) eastward to the Andamans. Other records outside this area are doubtful.

Carcharhinus wheeleri [New Caledonia, Solomon Islands]

Western and Central Indian Ocean: Natal, South Africa to Somalia, the Gulf of Aden, and Red Sea. Termed as C. amblyrhynchus by Wheeler (1962) and C. spallanzani by Bass, D'Aubrey & Kistnasamy (1973). Very close to C. amblyrhynchos and may prove to be not distinct from that species.

Malacocephalus laevis [Hawaii]

Eastern Atlantic: Faroe Is. to South Africa. Reported from Iceland. Western Atlantic: Straits of Florida to Brazil, Gulf of Mexico and Caribbean Sea. Indian Ocean: Arabian Sea, Bay of Bengal, off Maldives and off East African coast. Western Pacific: Indonesian area and Australia. *M. hawaiiensis*, *M. luzonensis*, and *M. nipponensis* are closely related to *M. laevis* and may eventually prove to represent populations of this species, if a comprehensive comparison of material from the Atlantic, Indian Ocean and Pacific is done.

Fig. 1. - Example of a map showing point data for Labridae. Note that some points were erroneously plotted on land whereas other points were plotted in the open ocean where the occurrence of wrasses is highly unlikely.