

Vanden Berghe E., M. Brown, M.J. Costello, C. Heip, S. Levitus and P. Pissierssens (Eds). 2004. p. 71-76  
Proceedings of 'The Colour of Ocean Data' Symposium, Brussels, 25-27 November 2002  
IOC Workshop Report 188 (UNESCO, Paris). x + 308 pp.  
– also published as VLIZ Special Publication 16

## **Development of an Atlantic Canada marine species information system based on a museum collection: a case study**

Angela Martin, Lou Van Guelpen, Gerhard Pohle and Mark J. Costello

Atlantic Reference Centre, The Huntsman Marine Science Centre  
1 Lower Campus Road, St. Andrews, N.B., Canada, E5B 2L7  
E-mail: arc@dfo-mpo.gc.ca

### **Abstract**

The Atlantic Reference Centre (ARC), established in 1984 as an amalgamation of regional collections, is a research museum of Canadian Atlantic marine biota and a provider of regional marine biodiversity information. Computerization of all catalogued lots has been accomplished through museum catalogue database restructuring and various funding programs. Because of its value as an archive of marine biodiversity data, concurrent ARC goals have been to place the database on-line and to use it for developing a species information system (SIS) for Canadian Atlantic marine biota. The SIS is an evolving entity of products, often collaborative, and funded by various biodiversity information initiatives. The ARC SIS is currently comprised of specimens and catalogues; on-line sites with querying, mapping and analysis capability for each museum taxon or displaying images, information, and a distribution map for each Canadian Atlantic fish species; a comprehensive list of all species inhabiting the Bay of Fundy, supplemented with substantial species information; a list and classification of all fish species in Canadian Atlantic waters; an on-line site integrating the biodiversity databases of the ARC and other providers with DFO physical and biological databases, into a biodiversity atlas service; and funding to 1) place ARC fish records on Species Analyst Canada, 2) develop a comprehensive Canadian Atlantic marine biodiversity register, 3) develop a web site for the Bay of Fundy species information system, and 4) perform quality assurance and quality control of the ARC museum database, thus improving data quality in the SIS. Future plans include collaboration with LarvalBase and FishBase in 2003-04, and expansion of the Canadian Atlantic marine biodiversity register. The ARC SIS will aid students, educators, researchers and managers in studying, protecting, and promoting the sustainable use of Atlantic Canada's natural resources.

Keywords: Biodiversity; Information; Marine; Atlantic Canada.

### **Introduction**

The Atlantic Reference Centre (ARC) is a research museum of Canadian Atlantic marine biota containing an archive of biodiversity information in the form of preserved specimens, and paper and computerized museum catalogues. This paper describes how, through a wide variety of short-term projects, the ARC is developing these resources into a Species Information System (SIS) for waters of Atlantic Canada, comprised chronologically of:

- ARC museum specimens, and paper and electronic catalogues,
- on-line sites with querying, mapping and analysis capability for each museum taxon,

- an on-line site displaying images and information for each Canadian Atlantic fish species,
- a comprehensive list of plant, invertebrate, and vertebrate species inhabiting the Bay of Fundy, supplemented with substantial species information,
- a list and classification of all fish species in Canadian Atlantic waters,
- an on-line site displaying a distribution map for each Canadian Atlantic fish species,
- an on-line site integrating the biodiversity databases of the ARC and other providers with DFO (Department of Fisheries and Oceans) physical and biological databases, into a biodiversity atlas service,
- 2003 funding to 1) place ARC fish records on Species Analyst Canada, 2) develop a comprehensive Canadian Atlantic marine biodiversity register, 3) develop a web site for the Bay of Fundy species information system, and 4) perform quality assurance and quality control of the ARC museum database, thus improving data quality in the SIS.

## The museum collections

The ARC, a joint operation of the Huntsman Marine Science Centre (HMSC) and the DFO, is located in St. Andrews, New Brunswick, Canada on the Bay of Fundy shore. Established in 1984, the ARC united two organizations: the HMSC Ichthyoplankton Laboratory, established in 1977 and archiving a valuable collection of Canadian Atlantic ichthyoplankton and adult fishes, and the DFO Identification Centre, initiated in the 1960s and with excellent holdings of local invertebrates and fishes, some dating back to the early 1900s. Collection growth has continued to the present such that holdings now represent all Canadian Atlantic waters.

A major benefit of the collecting history is that, for selected DFO research surveys, especially ichthyoplankton surveys, all material collected was archived at the ARC and is of great value for ecological, population, and biogeographic, as well as systematic studies. ARC holdings by December 2002 comprised 126,403 catalogued lots (a lot contains all specimens of a species from one collecting event, such as a bottom trawl) (Table I). The ichthyoplankton collection is the largest in the United States and Canada (Poss and Collette 1995). The ARC also houses type specimens and osteological preparations.

Table I. ARC collection holdings, December 2002

Collection	Catalogued lots	Representing
Invertebrates	13,475	24 phyla
Fishes	24,318	204 families
<u>Larval fishes</u>	<u>88,610</u>	<u>105 cruises</u>

In-house taxonomic expertise has contributed to major collaborative research undertakings, such as the US GLOBEC (Global Ocean Ecosystems Dynamics) program on Georges Bank, and DFO's Scotian Shelf Ichthyoplankton Program on Nova Scotia's offshore banks. This expertise also includes marine zooplankton, nektonic and benthic invertebrates, and freshwater insects.

## **The computerized catalogue**

Since its foundation the ARC has had a policy of cataloging its holdings in computerized databases. However, changing technological developments resulted in a long and complex process of catalogue database restructuring over time involving dBase, MINISIS, Revelation, Advanced Revelation, MUSE, and finally Specify. This database restructuring was a learning process that wasted resources. From this experience we learned that when in-house database expertise is lacking, selection of a dedicated museum database management system is advisable and should be undertaken with care.

In 2002 the ARC adopted the museum database Specify (<http://usobi.org/specify>), for which technical support is freely available. Specify is the successor to MUSE (see MUSE at <http://www.usobi.org/specify/DevaSearch/search.html>). In recent years the manual addition of records into MUSE and Specify has enabled computerization of all ARC catalogued lots, completed in 2002. Also completed in 2002 was a long-term project to electronically catalogue the ARC's extensive ichthyoplankton holdings, utilizing DFO databases. These tasks were accomplished through data management and biodiversity funding programs in Canada and the USA, together with various government internships and summer youth employment programs (see Acknowledgements).

The culmination of ARC database development is a functional museum database into which newly catalogued specimens are regularly added. Since this database is an archive of marine biodiversity data, concurrent ARC goals are to make the database available on-line, and to use it as a basis for developing a SIS for Canadian Atlantic marine biota.

## **Data on-line and the Species Information System**

The ARC SIS is an evolving entity comprising products funded by various, usually unrelated, biodiversity information initiatives, and sometimes done in collaboration with other institutions (see Acknowledgements). In 2000 the ARC museum database was an integral component of a program to develop a system for integration, visualization, and analysis of distributed biogeographic and oceanographic information for marine populations using the internet and built-in tools for analysis (<http://netviewer.usc.edu/web/index2.html>). Another major outcome of this project was a web site for interactive mapping and listing, on a taxon basis, of the ARC's georeferenced holdings (<http://gmbis.marinebiodiversity.ca/aconw95/aconscripts/ARCSpecimenMap4.html>). This site was last updated in November 2002, and contains all ARC georeferenced records (approximately 91% of catalogued lots).

In 2001, production of the 'Distribution of rare, endangered and keystone marine vertebrate species in Bay of Fundy seascapes' (Bredin *et al.*, 2001) required the development of a preliminary list of fish species inhabiting Canadian Atlantic waters.

The next ARC SIS building block, in 2001, was an internet product entitled 'Fishes of Atlantic Canada: a photographic compendium' (<http://collections.ic.gc.ca/compendium>). The web site provides images, taxonomy and names, and important biological and ecological information for each fish species living in Canadian Atlantic waters. The fish list was developed from that described above.

Also in 2001, the ARC began the 'Development of a species information system for the Bay of Fundy'. The result was a database detailing taxonomy, general distribution and abundance, basic biological and ecological traits, conservation, economic, or other importance, and pertinent literature for annelids, crustaceans, echinoderms, and fishes inhabiting the Bay of Fundy. In 2002 this project continued as the 'Completion of a marine species information system for the Bay of Fundy', with the same details for phytoplankton, seaweeds, molluscs, remaining arthropods, birds, and mammals. During these projects, to facilitate future biodiversity data interoperability, standardized taxonomy was adopted using Sears (1998) for plants, the Integrated Taxonomic Information System (ITIS) (<http://www.itis.usda.gov>) for invertebrates, and FishBase ([www.fishbase.org](http://www.fishbase.org)) for fishes.

Late in 2001, the ARC developed a framework for a registry of Canadian Atlantic marine biodiversity datasets (<http://www.marinebiodiversity.ca/mbw/index.html>). The DFO Maritimes Science Data Inventory, a compilation of regional departmental datasets, was modified to incorporate those from other federal and provincial government departments, universities, museums, nongovernmental organizations, industry, and the literature. This was a non-standards-based framework containing metadata for each electronic and non-electronic dataset describing ownership, contacts, content in detail, computerization, and availability. It is anticipated that this framework will contribute to the basis for a national plan for marine biodiversity (Zwanenburg *et al.* 2003), as well as to the ARC SIS.

In 2002 the ARC produced 'An atlas of distributions of Canadian Atlantic fishes' (<http://collections.ic.gc.ca/FishAtlas>), in which a modified distribution map was created for each ARC georeferenced species and placed in an electronic atlas. Each species is linked to the same species in the photographic compendium above. The species list for the atlas was much larger than that for the compendium, since the very large ARC ichthyoplankton collection, and thus many more species, had been added to the museum database. FishBase standardized fish taxonomy was adopted for all species. From this project the ARC SIS evolved, with the addition of the new fish species, standardized taxonomy, and inclusion of synonyms previously used by the ARC.

The most recent SIS development was a grant to integrate on-line the biodiversity databases of the ARC and other providers with DFO physical and biological databases, into a biodiversity atlas service. This atlas service will feature collection-based and integrated browsing and mapping of data, with links to services such as ITIS, FishBase, CephBase (<http://cephbase.utmb.edu/>), Hexacorals (<http://www.kgs.ukans.edu/Hexacoral/>), and internet images. Query results will be exportable in various formats, including input for on-line biodiversity analytical applications.

After presentation of this paper at the Colour of Ocean Data symposium, the ARC obtained funding in 2003 to 1) provide ARC museum database records for fishes in Darwin Core Version 2.0 format to Species Analyst Canada ([http://www.cbif.gc.ca/speciesanalyst/fish/speciesquery\\_e.php](http://www.cbif.gc.ca/speciesanalyst/fish/speciesquery_e.php)), enabling on-line searching of ARC fishes; 2) develop a Canadian Atlantic marine biodiversity register containing a list and classification (based on Sears (1998) for plants, ITIS for invertebrates, and FishBase for fishes) of marine species in all major phyla inhabiting the Canadian Atlantic, and the range and concentrations of each species. This will become an on-line resource; 3) develop a web site for the Bay of Fundy species information system described above; and 4) perform comprehensive quality assurance and quality control of the ARC museum database, thus improving data quality in all components of the SIS.

Future plans involve growth of the ARC SIS. In 2003, ARC larval fish museum database records will be placed on LarvalBase ([www.larvalbase.org](http://www.larvalbase.org)), and on FishBase in 2003 or 2004. The Canadian Atlantic marine biodiversity register will be expanded to include taxonomy, identification, images, distribution and abundance, conservation status, basic biological and ecological traits, and commercial, scientific, educational, and/or social importance.

## Discussion

Museum collections provide verifiable records of species distributions. Despite changes in knowledge about species identity over time, museum material can always be re-examined for confirmation or correction.

Collections with multiple samples in space and time are of even greater research value. For example, the material can be studied for population structure (size, growth rate, age, maturity), genetic variation, diet and parasites. The ARC collections are thus of value not only to taxonomists and systematists, but to researchers of genetics, population biology, fisheries, parasites, and ecology.

It is unfortunate that many marine laboratories and institutes have not curated their collections as a resource for future generations. Those with collections have difficulty keeping their catalogues up to date. A survey in Europe, as part of the European Register of Marine Species project (Costello 2000), found 20% of collections had no catalogues at all, and only 36% had all of their collections catalogued on paper. The rate of computerized catalogues was less, with only 10% with full coverage of their collection, and 36% with no computerized catalogues.

Little use can be made of museum collections if basic information on their content, in terms of species, numbers of individuals, and geographic origins, are not known. Ideally, this information is not just catalogued on paper, but in a computerized database searchable from the internet. Such on-line publication of the collection information will increase awareness and hence usage of the collections.

As a provider of biodiversity information the ARC goal is to complete the SIS and become a primary source of data for species inhabiting Canadian Atlantic waters. We will continue to develop biodiversity products like those detailed above, and others, such as electronic species atlases. This paper shows how it is possible to develop an integrated SIS using a series of short-term projects and variety of funding sources. This comprehensive and easily accessible

assemblage of marine biodiversity information will aid students, educators, researchers and managers in studying, protecting, and promoting the sustainable use of Atlantic Canada's natural resources.

## Acknowledgements

We would like to acknowledge the funding agencies that contributed to the ARC SIS: Fisheries and Oceans Canada, the Centre for Marine Biodiversity, the USA National Ocean Partnership Program, the Alfred P. Sloan Foundation, the Census of Marine Life, Environment Canada's Ecological Monitoring and Assessment Network, the Gulf of Maine Council on the Marine Environment, Industry Canada, the New Brunswick Environmental Trust Fund, Natural Resources Canada's GeoConnections program, Environment Canada's Canadian Information System for the Environment, the Global Biodiversity Information Facility, Human Resources and Development Canada, Canadian Heritage Young Canada Works, and the New Brunswick Student Employment and Experience Development Program. Collaborators in the SIS component projects were: Fisheries and Oceans Canada, the Specify Project, and the Atlantic Canada Conservation Data Centre. We especially would like to thank Fisheries and Oceans Canada for its support through the years for the ARC and for our biodiversity information collaborations and initiatives. The paper was reviewed and enhanced by an anonymous reviewer. Attendance at the Colour of Ocean Data conference was made possible by a grant to the Ocean Biogeographic Information System from the Sloan Foundation through Rutgers University to MJC.

## References

- Bredin K.A., S.H. Gerriets and L. Van Guelpen. 2001. Distribution of rare, endangered and keystone marine vertebrate species in Bay of Fundy seascapes. Prepared for the Gulf of Maine Council on the Marine Environment by the Atlantic Canada Conservation Data Centre. v + 73p.
- Costello M.J. 2000. Developing species information systems: the European Register of Marine Species. *Oceanography* 13(3):48-55.
- Poss S.G. and B.B. Collette. 1995. Second survey of fish collections in the United States and Canada. *Copeia* 1995(1):48-70.
- Sears J.R. (Ed.). 1998. NEAS keys to the benthic marine algae of the northeastern coast of North America from Long Island Sound to the Strait of Belle Isle. A publication of the Northeast Algal Society, xi + 163p.
- Zwanenburg K., K. Qerbach, E. Kenchington and K. Frank. 2003. Development of a science plan for marine biodiversity in Canada. Canadian Technical Report of Fisheries and Aquatic Sciences 2432, viii + 72p.