

## **Biodiversity and Bioinformatics**

In these days of burgeoning human population and concomitant habitat destruction, knowledge of centers of high biodiversity is critical if rational conservation decisions are to be made.

The problem is that this information is largely unavailable to the decision makers. The reason for this is twofold.

First, although a few groups of organisms such as birds and mammals and a few geographic areas, such as western Europe, are well studied and well characterized most groups of organisms are only partly known and the tropical parts of the world and the deep ocean have only just begun to be studied in detail.

These problems are compounded by the fact that there is little incentive for biologists these days to go into the classical fields of taxonomy and systematics. The glamour (and money) is in molecular and cell biology.

Secondly, although there is an enormous amount of biodiversity information in the worlds museums and universities it is not readily accessible.

It is ironic that most of these data are in the great museums which are located in the cool temperate parts of the world whereas, most of the organisms are in the warm humid parts of the world.

The data that exist are paper based. Descriptions by collectors and curators, herbarium sheets, diagrams and photographs, and of course, pickled and preserved specimens with their labels.

If a researcher wishes to consult these data he/she has to travel to the museum in question and do the work there. For the people who need a breadth of information to make decisions, this is obviously not an option.

There are moves afoot to remedy this situation.

There are two areas in biology where enormous amounts of information are generated. One is in molecular biology which deals with base sequences in DNA and amino acid sequences in proteins, and the other is the biodiversity information crisis.

Mathematics and computers are being used to tackle these problems with procedures which come under the label of Bioinformatics.

Attempts are being made at places like the University of Kansas, the Natural History Museum in London, the American Museum of Natural History in New York etc., etc. to computerize all of their data. This is an enormous and complex task, but this is only part of it. To make the data efficiently available world-wide it has to be accessible via the internet.

This is being done by using a relatively new computer language called XML (Extensible Markup Language).

Superficially XML looks like the more familiar HTML but whereas HTML is concerned with the appearance of a document, XML defines the internal structure, the organization of the content of a document and when it is fully implemented will revolutionize the speed and efficiency of access to all types of information over the net.

When, and if, we reach a point where most of the information has been accumulated and furthermore, is readily accessible, it will be much easier for concerned individuals and groups to make the case for the preservation of areas that are particularly rich in diversity. They will be able to overwhelm the politicians with actual data that hopefully even they will be able to understand and find difficult to ignore.

by Brian Pasby

'Science' - 29 September 2000 -vol 289, contains articles discussing the application of bioinformatics to biodiversity.

Hernando Chapter of the Florida Native Plant Society, P.O. Box 921, Brooksville, FL 34605. © Jan 2001  
Web site address: <http://members.tripod.com/~hcfnp>