BIODIVERSITY INVENTORY: REFLECTIONS ON PREPAREDNESS AND EFFICIENCY

Michael J. Donoghue[1] and Marilyn Smith[2]
1. Department of Ecology and Evolutionary Biology
and Peabody Museum of Natural History,
P.O. Box 208118, Yale University, New Haven,
CT 06520-8118 USA
2. DIVERSITAS, 51 bd de Montmorency,
75016, Paris, France

Abstract

Why should we even try to compile a comprehensive list of the species and clades on Earth? Here we develop an insurance analogy to highlight the value of an inventory—without it we might not recognize changes when they occur or be prepared to respond appropriately when our assets are threatened. If an inventory is necessary, could it ever really be accomplished? Considering current constraints on time and money, "business-as–usual" won't work. We need to develop and deploy new technological solutions to increase the efficiency of fieldwork and museum informatics.

Introduction

As demonstrated so clearly at the International Conference on Biodiversity: Science and Governance (Paris, January 2005), scientists, government leaders, and the business community agree that Earth's biological diversity is undergoing rapid change. They also recognize that these changes will have some major, and possibly quite dangerous, consequences for humanity. At the same time, the meeting highlighted how little we still know about the diversity of life. New and, in many cases, dramatically different species and more inclusive branches (clades) of the Tree of Life are being discovered at an astonishing rate, apparently limited only by the effort we put into exploring the biosphere (Donoghue and Alverson 2000).

Just as we know enough about global climate change right now to attempt to minimize its impacts, we surely have enough information on the loss of biodiversity to know that we must act to prevent it. Yet, there is a genuine concern—which frequently arises even in discussions amongst committed biodiversity scientists and policy-makers—about how to best spend our limited time and money. Given the daunting nature of the task, and with so many other pressing biodiversity issues in need of attention, it is perfectly legitimate to ask whether we really need a comprehensive list—an inventory—of the many millions of species and more inclusive clades that exist on Earth (e.g., see Renner and Ricklefs 1994a, b, and responses). If the reason is only "because it's there", or even "because we just might find something useful", the skeptics might be right. This question deserves our careful attention and a well-reasoned answer.

We think that the right response centers on the notion of preparedness. Obtaining the base-line knowledge of what biodiversity exists is fundamental to recognizing changes when they occur, to understanding what we're responding to, and to knowing how best to respond. In the battle to combat biodiversity loss, being able to identify what is changing within the system (e.g., extinction, introduced species) is key to tackling questions of when, where, how, and why. To better appreciate the connection between an inventory and this notion of preparedness, we think it is useful to consider analogous situations, and the first aim of our essay is to briefly develop one such analogy, namely creating a household inventory for insurance purposes.

Our second goal is to briefly reflect on a closely related question: in view of the rather pitiful state of our knowledge of biodiversity, how can we even hope to generate a comprehensive list? We develop the view that it is, indeed, rather hopeless to take on such a monumental task if the taxonomic community carries on with business as usual. Some dramatic changes are needed in how we operate. Specifically, we need to develop and to fully utilize new technological solutions, both in conducting fieldwork and in advancing museum informatics.

Preparedness: An Insurance Analogy

The biodiversity inventory problem compares quite closely, we suggest, to making a list of one's household possessions for insurance purposes. Many of us have filled our houses with so many things that it's impossible to remember them all. Some of these objects are of great value, either because they cost a lot in the first place or have a high replacement value, or because we cherish them for sentimental reasons. In the United States, at least, it is common to purchase insurance to cover one's belongings to avoid being financially crippled by a catastrophic event (e.g. fire, theft, or natural hazard). To determine appropriate premiums for a homeowner's policy, the insurance agent typically asks the client to prepare a comprehensive list of possessions and to assign an approximate replacement value to each item. Ideally, this list comprises short descriptions, receipts to demonstrate ownership, and serial numbers. It might also involve photographs or videotapes. The goal is to document what the client owns and how these possessions can be distinguished from anyone else's.

The objective of the inventory in this case is entirely obvious, as is the need. The list is the base-line against which the client and the insurance agent will measure change in the event that something happens to the home; it provides the proof of what was present before the incident occurred. With the list in hand, it is possible to identify any losses or additions, damage or movement. Without it, the homeowner may not know if, what, or when changes occurred, and certainly couldn't prove such changes to the satisfaction of the insurance agent.

Take the case of a robbery. Most people will immediately notice the absence of some items—particularly bigger things, such as the TV or the stereo. The absence of other, smaller items (e.g., a camera or a seldom-worn ring) might go unnoticed for some time, until the need to use them arises again. With a list in hand, the homeowner could carry out a systematic monitoring as soon as the incident occurred and establish precisely what was taken. The more comprehensive the list, the more complete the client's knowledge of what was stolen and the greater the opportunity to reap the maximum benefit from the insurance coverage. Items not listed on the inventory might be overlooked when preparing a claim. Some of these might turn out to have been quite valuable, but without proper documentation there is little chance of being compensated.

In this analogy the insurance list obviously corresponds to a base-line biodiversity inventory and the loss of biodiversity (owing to habitat loss, invasive species, etc.) is represented by the burglary. There are some important differences, such as the fact that we don't "own" biodiversity in the conventional sense of the word. Of course, the fundamental disanalogy is that when

it comes to species loss there is no insurance agent, no policy, and zero possibility of buying a replacement. If the species exists elsewhere on the planet, a re-introduction might be possible (probably at very high cost); if not, it's gone forever. This absence of a compensation mechanism should only strengthen the need to be prepared for any eventuality and to institute measures to safeguard our biological resources. Taking steps to maintain healthy ecosystems is obviously a form of insurance in itself (Perrings 1995), as these will surely improve the chances of recovery from perturbations of all sorts.

Before we extend this analogy, it's important to appreciate that whereas the list of possessions helps in cases of loss, and this is what insurance agents are primarily concerned with, a good list can identify other changes as well, including additions and movements. An inventory might be used as proof that a particular item was added to the home, or that an item was replaced by something similar (e.g., a forged painting for an original work of art). Again, reference to the inventory, depending on the detail of the documentation, makes it possible to identify such changes.

In this analogy it is also useful to consider strategies for creating the list and how this might depend on the circumstances. Insurance agents will agree that even a partial list is better than no list at all, and that the task might need to be approached in manageable bites. External factors might dramatically increase the need for a list while also imposing a time frame for getting it done. The distant threat of rare events—theft or lightening damage that might happen "some-day"—tends to make homeowners somewhat lackadaisical about completing an inventory. But events that are anticipated or highly likely to occur within a particular time frame (e.g., a lava flow approaching the neighborhood) provide excellent incentive for setting pen to paper.

The perceived urgency of a situation might also influence the way in which an inventory is conducted. One approach might be to concentrate first on documenting high-priority items—listing possessions with the greatest monetary or sentimental value, or perhaps the items viewed as most critical to ensuring the family's survival or re-establishing a household in case of catastrophe. Another tactic may be to make a comprehensive list for a few key rooms in the house and leave the others for later (e.g., one may need to quickly restock the kitchen to survive, but could get by without the things in the attic). If the situation is urgent, one might forego a comprehensive list but instead quickly take photographs of each room, which would at least show the basic contents (oriental rug in the living room, large clock in the corner) even though details about what's inside the closets would be missing. In each case, the homeowner faces the question of how much information is necessary and how best to gather it. Is it best to gather a little information about a lot of items or more detailed information about what seems most valuable? How many people, and with what knowledge, are needed to carry out the task given the time and the resources available?

Again, the biodiversity connections are obvious. A sudden disaster such as the December 2004 tsunami is a stark reminder of events that can suddenly alter entire ecosystems about which little was known. The anticipated impacts of human-induced climate change provide an incentive to inventory as much as possible, as quickly as possible. The need to respond to ongoing habitat destruction (e.g., deforestation in the tropics), or the sudden possibility of establishing a refuge area, call for even more immediate action. Given a range in time-sensitivity, a variety of approaches can be considered. Maybe we should put together a quick list of as many species as possible, but with relatively little depth of coverage (e.g., a "rapid biological inventory," such as those carried out by the Field Museum; www.fieldmuseum.org/rbi), as opposed to conducting a more protracted "all taxon biotic inventory" (e.g., Discover Life in America; www.dlia.org). Perhaps we should focus on the species that we think play the most important roles in maintaining ecosystem function, or provide the most important goods and services. Or maybe we should concentrate on what we think are key environments or biodiversity hotspots. Should we, for example, focus first on tropical forests and set aside tundra and the deep oceans for the time being?

Of course, all such choices entail risks. For example, it's not always entirely clear which species are crucial to ecosystem function, or which areas truly harbor the most biodiversity. In addition, our fragmented knowledge of biodiversity increases the chance of not paying adequate attention to—or even completely overlooking—something that turns out to be critically important. We might focus on forest trees at the expense of understanding fungal biodiversity in the soil, only to find out that the forest system crashes without the mushrooms. In view of such uncertainties, it might be best to pursue a mixed strategy, one team setting out to compile a quick list of species names and characteristics, while other teams focus on more detailed accounts for what are deemed to be key taxa or key areas.

This last point brings up the thorny issue of assigning value, which is a key element of making a list for insurance purposes. This is much easier, of course, for household possessions than it is for species and clades, or for ecosystem services (e.g., Pearce and Moran 1994; Perlman and Adelson 1997; OECD 2002). Some elements of biodiversity can be measured in economic terms because they produce (or contribute to the production of) commodities that are bought and sold in real markets. But this doesn't pertain in the case of most species or, indeed, for most biodiversity goods and services. Fortunately, more attention is now being paid to the aesthetic and cultural values of biodiversity, which in some ways are akin to the sentimental value of personal possessions. Just as a homeowner might seek the expertise of an antique dealer or art agent in assessing the value of a table or painting that has been in the family for years, natural scientists clearly need to draw on other expertise, engaging local people and social scientists in the dialogue.

One of the most important aspects of a household inventory is its capacity to act as a tool for monitoring change over time. Once a list is created, it is relatively easy to add new purchases or delete items that are lost, stolen, or no longer functional, and to track factors that influence the value of individual items. Just as the value of an original painting might skyrocket as the fame of the artist increases, the economic value of a particular plant species might increase when it is found to contain the key ingredient for a drug in high demand. The tiny cyanobacterium *Prochlorococcus marinus*, which was not discovered until the 1980's, is now known to be the dominant photosynthetic organism in the open oceans, therefore playing a major role in the global carbon cycle (Chisholm *et al.* 1988). Our perception of its value has risen accordingly.

Our hope is that this analogy adds weight to the argument for biodiversity list-making. We need a list of the species and the more inclusive lineages on Earth for a very good reason: to be able to properly monitor changes in the system--what's been lost, what's been added, and what's moved around. In a nutshell, lists prepare us for change—to recognize it when it occurs and to respond wisely. In the case of biodiversity, this might sometimes seem unnecessary. Won't we know when a bird has gone extinct or when a plant species has been introduced? Maybe so, but even this depends on authoritative lists having been made in the past. And what about organisms that have not attracted so much attention? We don't have decent lists of the bacteria that live in our soils or in the oceans, ergo we really have no idea if species are coming or going from these systems. Changes may already be occurring in these communities that could profoundly influence the functions that underpin life on our planet. Lacking the proper base-line knowledge, we are dangerously under-prepared to recognize change when it happens and to implement appropriate responses.

The insurance analogy intuitively links list-making to preparedness. When disaster strikes, some will reap the benefits of having prepared a list and can get on with rebuilding their lives; others will deeply regret not having invested the time and energy to take stock of their assets. However, the insurance analogy has its limitations, and there may be others that are better suited for some purposes. For example, where would doctors be without an accurate "inventory" of body parts (human anatomy)? Or, in these times of tightened homeland security, is it any wonder that governments want a complete list of who lives where, who's entering and who's leaving?

Our aim is not to push the insurance analogy as the sole or even the best one, but to use it as a reminder of the vital role that inventories play in our daily lives and to help clarify the need for biodiversity inventory in particular.

Technological Efficiency

It may be necessary (or, at least, highly advantageous) to inventory biodiversity, but what can we realistically hope to accomplish? In view of the growing threats associated with biodiversity change and loss, what strategies should we adopt? As we noted above, this choice depends on a number of factors, especially on our perception of the impending danger and on the resources that we can devote to the project. If we lived in a relatively static world, with many centuries before we might experience any major changes, then the status quo might suffice. But it is abundantly clear that we have a biodiversity crisis on our hands: biodiversity loss is occurring rapidly and is bound to accelerate.

We have also begun to appreciate the depth of our ignorance of biological diversity, and just how much more remains to be discovered. Some 1.7 million species have been named and described (usually only very briefly) over the last three centuries, but there may be ten or more times this many awaiting discovery. Likewise, we have discovered and named only a tiny fraction of the more inclusive branches of the Tree of Life (Donoghue 2004), which will provide the ultimate basis for organizing and navigating all biological information (Cracraft and Donoghue 2004). Unfortunately, on what would appear to be the relevant time scale—say, over the next 30 to 50 years—we're not likely to experience a major increase in the number of taxonomic experts, upon whom all of this work depends. After all, it takes time to develop this expertise.

Under these circumstances, it seems abundantly clear that business as usual won't work. In every respect, taxonomic effort is currently too limited and the taxonomic process is too slow. If we hope to make a serious dent in developing a list that will truly be useful in recognizing and combating biodiversity loss, some dramatic changes must be made. We need new approaches to conducting fieldwork—to more rapidly identify what's living where, and to uncover new species. We also desperately need to unlock the information currently hidden away in the world's natural history museums. There may be well over two billion specimens in these collections, almost all of them accompanied by some basic biodiversity information: what the organism is, and where and when it was collected. These data need to be rendered accessible as quickly as possible so that we can more efficiently establish the existence and the limits of species, chart their known geographic distributions, and make predictions about where they may occur in nature.

We need new electronic publishing mechanisms and, as unpleasant as it may seem, we need to take a hard look at whether our current nomenclatural systems are encouraging or impeding the rapid publication of new taxonomic discoveries (Hibbett and Donoghue 1998). And we need to develop new ways to quickly translate discoveries into the resources that are used by decision-makers.

Our aim here is to highlight just a few of the new technologies that would dramatically increase the efficiency of biodiversity inventory. In terms of the insurance analogy that we developed above, these all can be viewed as ways of speeding up the list-making process and improving its quality—getting the most and best information as quickly as possible.

Beginning first with fieldwork, we can and must supplement standard collecting techniques—which yield standard specimens—with new approaches that take advantage of digital cameras. For some groups of organisms (e.g., many land plants) proper digital images may often be sufficient for accurate identification, and could certainly be recorded as evidence of the presence of a species at a particular time and location (measured precisely, using a GPS device). The technical problem

for the taxonomic community is to develop the relevant standards and the proper means of archiving such digital data and integrating them with data obtained from traditional specimens. This simple use of digital cameras in the field would have a dramatic impact on the rate at which inventories could be carried out, as well as on our knowledge of species ranges. It may also be possible (using wireless communication) to instantly transmit images to experts situated in remote locations, thereby greatly accelerating routine identifications and the discovery of new species.

Another technological innovation is more far-fetched, but is now within the realm of possibility: the development of a hand-held, automated species-identifier, modeled after the fictional "tricorder" used to identify alien life-forms in the popular "Star-Trek" films (e.g., www.tricorder.cjb.net). The idea is to obtain a tiny sample of an organism, to quickly extract, amplify, and sequence a set of target DNA markers, and then to compare these to known sequences and situate the unknown specimen within the phylogeny of life. Such a device has not yet been developed, but in principle it can (and presumably will) be done. This is not to suggest it would be easy, however, or to overlook the challenge of creating the necessary underlying databases. The basic motivation behind such a hand-held identifier also underlies the DNA "barcoding" initiative (www.barcodinglife.org), although the barcoding project is currently much more narrowly focused on mitochondrial CO1 sequences and not on nanotechnological solutions per se. Of course, all such projects will need to deal with sequence variation within species (and therefore with sampling protocols) as well as the difference between genes trees and species trees (e.g., Maddison 1997).

Museum-related informatics tools are developing at a rapid rate and are already setting the stage for a global biodiversity cyber-infrastructure. Phylogenetic information is becoming more widely accessible (e.g., TreeBASE, www.treebase.org; The Tree of Life Web Project, www.tolweb. org) and work is underway to dramatically extend such resources (e.g., CIPRES, www.phylo.org). Museum specimen information is accessible through a wide array of national (e.g., CONABIO in Mexico, www.conabio.gob.mx; AVH in Australia, www.anbg.gov.au) and international (e.g., GBIF, www.gbif.org) database-centered organizations. In addition, schema-independent protocols (e.g. DiGIR, www.digir.net) have been developed to allow rapid access to specimen information on particular groups of organisms that are distributed across museum collections (e.g., MaNIS, elib.cs.berkeley.edu/manis; HerpNET, www.herpnet.org). And several taxon-oriented informatics efforts have focused special attention on providing excellent digital images (e.g., Automontage in AntWeb; www.antweb.org).

A critical bottleneck is the rate at which information is being captured from individual museum specimens. At the moment, this generally requires that label data be physically typed into local database systems, which is laborious, costly, and altogether too slow. This process badly needs to be automated, so that it involves little or no human intervention. This is precisely the focus of a US NSF-funded project that one of us (MJD) is currently involved in: the HERBIS project (www.herbis.org), centered in the Botany Division of the Peabody Museum (Yale University). The idea of this project is to produce a system of web-services to automatically upload into a collections database—with the click of a button—the label data associated with a herbarium specimen (Fig. 1). In our prototype system, a digital camera captures a high-resolution image of the entire specimen. The label image is automatically extracted and then text recognition (optical character recognition, natural handwriting recognition) and natural language processing software parse the label information into relevant database fields. Systems of this sort have the potential to accelerate enormously the rate at which museum data are brought online.

Ready access to museum specimen data makes it possible to generate distribution maps and geographic range predictions based on niche modeling (Peterson 2001; Scott et al. 2002). These processes rely critically on geospatial referencing (georeferencing), which also needs to be automated. Fortunately, significant progress is now being made in this area (Beaman et al. 2004),

Figure 1. HERBIS (www.herbis.org) schema for automatically capturing label data and populating a collections database from a single digital image of an entire herbarium specimen. This process entails a series of web-services for text recognition (OCR = optical character recognition; NHR = natural handwriting recognition) and natural language processing (NLP). Automated georeferencing (e.g., using BioGeoMancer; www.beogeomancer.org) can be incorporated in this process.
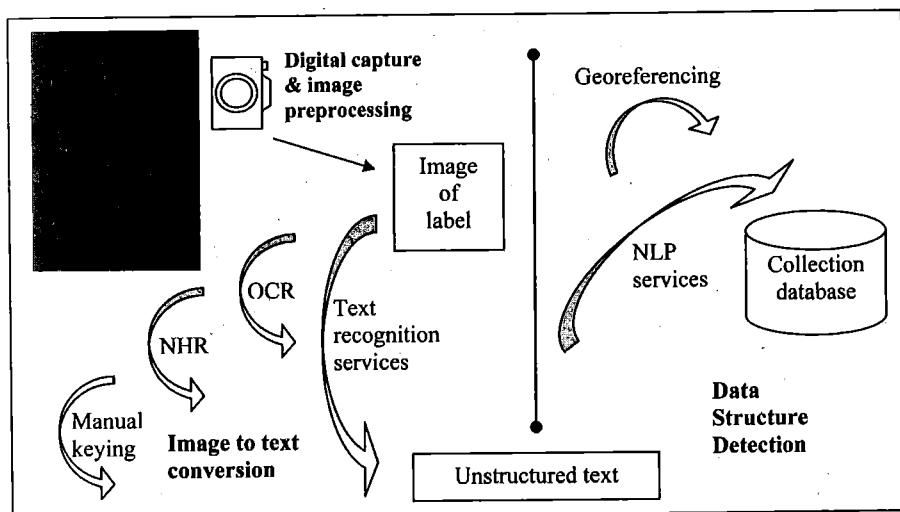
including the ability to translate English language placename descriptions into a set of latitude/longitude coordinates (BioGeoMancer, www.biogeomancer.org) and the development of uncertainty measures around a geospatial reference point (Georeferencing Calculator, http://elib.cs.berkeley.edu/manis/gc.html). These advances will be crucial in linking specimen data to predictions about the distribution of species and clades, and therefore to monitoring and managing biodiversity change.

We have focused attention on only a few of the technological advances that could (and will) make biodiversity inventory more efficient and effective. It remains clear, however, that there are no magic bullets, and that the "taxonomic impediment" recognized by the Convention on Biological Diversity in their establishment of the Global Taxonomy Initiative (www.biodiv.org/programmes/cross-cutting/taxonomy) is all too real. Support for taxonomic training and for building the infrastructure for taxonomic research must obviously be increased, and capacity-building in biodiversity-rich areas is especially critical. In the long-run we are quite optimistic about this, simply because the underlying science is so exciting and the opportunities for genuine discovery are so great. In the meantime, given the need for immediate action, we must make the best use of existing taxonomic expertise by developing the tools to make this critical work as efficient as possible.

Concluding Thoughts

It is important to ask the question: why should we invest in a biodiversity inventory of our planet? Our answer revolves around the notion of preparedness—we need proper base-line information against which we can measure change and respond to it appropriately. We hope that our insurance analogy will help get this point across. It is equally important to think carefully

about a cogent strategy, taking into account the obvious constraints on time and money. We are convinced that the process can be greatly accelerated through the invention and deployment of a variety of technological solutions. Of course, this will take serious commitment and investment, but we are certain that the increases in efficiency will be well worth the effort.

Acknowledgments