

Taxonomy as Information Science

H. Charles J. Godfray

*NERC Centre for Population Biology, Imperial College London, Silwood Park Campus,
Ascot, Berkshire SL5 7PY, Email: c.godfray@imperial.ac.uk*

The taxonomy of a group of organisms today consists of the sum of papers on their classification in a myriad of different journals. This distributed taxonomy, however, is not the only possible model. For a particular group, species descriptions, keys, consensus and alternative classifications, and references to type material could all be mounted on the web to provide what might be called a unitary taxonomy. Once established, future work on the group would refer only to the species and other taxa hypotheses on the unitary web site and not to the preceding paper literature. The paper will explore the advantages and disadvantages of unitary taxonomies. The advantages include greater accessibility and visibility, preservation of the useful but not the hindering legacy of the past, the ability to employ greater numbers of images, and the ease of linking morphological and molecular approaches. The disadvantages include the costs of maintaining unitary taxonomy web sites, the requirement of web access, and the risk of authoritarianism. I argue that the advantages outweigh the costs and that unitary taxonomies may be the means of attracting the new funds that systematics so urgently needs and deserves.

Why do modern developed societies pour so much money into scientific research? In some cases the answer is straightforward, there is a clear economic benefit in the short to medium term. This is the reason for some government and most industrial funding. There may also be non-economic short to medium-term benefits that can be thought of as improvements to our quality of life. How can we disperse oil slicks, or improve air quality by better scrubbing power station emissions, or develop better forms of medical care (medical research has both economic and non-economic short-term aims)? All of these can be justified and accounted for by rational-choice economics using straightforward and easily communicated concepts of utility.

This, however, accounts for only a fraction of the activity in our great research institutions, universities, museums, and other research centers. Here, much of the research has only long-term potential benefits; or is capacity building, paving the way for more utilitarian studies; or is “blue-skies”, with no identified concrete benefits to mankind but a hope that something unexpected and useful might arise; or is just curiosity driven, motivated by a belief that accumulating knowledge is somehow a noble human endeavor. There is a curious social contract between the citizen and the scientist that allows the latter considerable latitude in what he or she does with taxpayers’, charity-givers’, or philanthropists’ money, yet nevertheless demands certain some ill-defined returns for whatever type of science is funded.

Where does taxonomy fit into this picture? Some taxonomists can point to direct short- to medium-term benefits of their work. The value of fossil stratigraphy to the oil and mining industries is an obvious example, as is the importance of entomological taxonomy to biological control for pest management and — although often exaggerated — plant taxonomy to bio-prospecting. But these are a minority; few businesses employ taxonomists as part of their core enterprises, and pri-

vatising taxonomy would be a disaster. The short- and medium-term goals of taxonomy relate to quality of life than rather wealth creation. Where taxonomists are employed or used as consultants by industry, it is more typically in their environmental units, which require expertise in assessing negative effects of their activities on the environment. Governments and NGOs also often use taxonomists in this capacity.

But all of this accounts for only a small fraction of taxonomy — the justification for the remainder lies in the realm of capacity building, “blue skies” and research driven by curiosity. Of these, “blue skies” is perhaps the least persuasive, at least as I define it here as research that may trigger unexpected benefits to society. Though it is rather nonsensical to say one cannot think of any unexpected benefits, the history of the subject is in delivering research along predicted rather than unpredicted lines. The argument for the importance of taxonomy in enabling other science is unimpeachable. Much of whole-organism biology requires the accurate identification of species and, increasingly, an accurate understanding of their phylogeny. This is perhaps most critically true in community ecology, the science of the interaction of assemblages of different organisms. The lack of good taxonomies for many groups of especially tropical organisms is a real impediment to research. It is not only the supra-organismic sciences that benefit from taxonomy. One must be careful not to exaggerate but an increasing number of interesting questions in molecular, genetic and developmental biology are benefiting from a comparison across species, especially when the comparison is based around an *Arabidopsis*, *Caenorhabditis*, *Drosophila* or other model organisms.

In my view, an unarguable case can also be made for curiosity-driven taxonomic research, of a type that will command public support. One line of reasoning, the one I find the least convincing, is that the human species has a duty to document the biodiversity with which it shares the planet. I think “duty” and similar expressions are philosophically suspect and really just a cover for a powerful emotional need to understand and conserve plants and animals. E.O. Wilson would call this emotion part of our species’ “biophilia” and I think that it is this type of notion, shorn of the highfaluting rhetoric of duty, and genuinely democratized so that it is much more than the special pleading of taxonomists and fellow travelers, that is the strongest justification for continued and increased support for taxonomy. That this ground swell of support for the subject is real is evidenced by the legion of amateur taxonomists (in which category I include keen bird watchers and wild flower enthusiasts, and perhaps largest of all, serious amateur gardeners) and the popularity of programs and books with significant taxonomic content (I’m reminded here of a six-year old girl I know who has an encyclopedic knowledge of dinosaur nomenclature).

Thus, is there a problem in taxonomy and, if so, what is it? I think there is a widespread belief amongst many taxonomists and scientists who use their output that species-level taxonomy is under-funded and failing to produce the systematic work needed by other fields. I am most familiar with the situation in the United Kingdom where the evidence for this is clear. Descriptive taxonomy has largely disappeared from the university staff room (and curriculum) while funding for major museums and herbaria has fallen in real terms. The Linnean Society, the Systematics Association, and similar societies produce manifestos lamenting this decline, and government responds by setting up serial enquiries (the latest at <http://www.parliament.the-stationery-office.co.uk/pa/ld200102/ldselect/ldsctech/118/11801.htm>) which do some good (one led to a five-year taxonomy initiative by a research council) but lead to nothing permanent and structural to reverse the situation. I think similar things have happened in other countries, a steady decline with some encouraging exceptions (for example, the couple of recent National Science Foundation taxonomic initiatives in the United States). Most pernicious of all, the standing of the field has become undermined, with taxonomy thought of as “old-fashioned” as and less sexy than the many

other disciplines that make up modern biology. I have concentrated here on species-level taxonomy; the standing of phylogenetics is somewhat higher, though this too suffers from the lack of funding for the whole field.

Here is an argument for funding less taxonomy. Scientific research is essentially a market and, give or take a little inertia and the minor distortions of fads and fashions, funding follows the questions that are either most useful or most interesting. In an environment where for quite healthy reasons there is intense competition for the research dollar, the lack of resources going to the subject is a logical and helpful response to changing economic and intellectual priorities. Taxonomists feel under-funded only because they remember a bygone age when they were at the cutting edge of science.

Clearly this argument is simplistic, and no one would seriously argue that an invisible hand is creating a perfect market to guide research funding. But there are some features of the modern science-funding scene that do resemble a market, and in this article I want to argue that thinking about how taxonomy operates in this milieu may help identify what changes the subject should make to increase its support, and what changes it should shun because they would do the reverse. Specifically, I want to concentrate on more long-term research and suggest that funds do tend to follow the most interesting questions but in a more sophisticated manner than that of the argument of the last paragraph. What matters is not only how interesting the question is, the potential extra science that the research may leverage, but how capable the subject is of delivering useful output. I think taxonomy is suffering not because it is any less interesting or important than it was fifty years, but because it is largely failing to deliver.

ARGUMENT BY ANECDOTE

The Alloxystinae are a group of tiny and fascinating wasps that are hyperparasitoids of aphids (and a few other Homoptera). They belong to the insect order Hymenoptera and are traditionally included in the paraphyletic Parasitica, the vast majority of whose members are parasitoids of other insects (Gauld and Bolton 1988). They are now normally placed in the family Charipidae of the superfamily Cynipoidea, whose most famous members are the gall wasps (Cynipidae) which have secondarily lost the parasitoid life style. Taxonomists consider there are six or so genera of which only two, *Alloxysta* and *Phaenoglyphis*, have more than a few species.

The biology of the aphid-attacking species is relatively uniform. Aphids are prey to many primary parasitoids, especially those belonging to a group of braconid wasps called the Aphidiinae. These lay their eggs in the aphid, typically before it is fully grown, and the parasitoid first-instar larva remains in a state of arrested development while its host feeds and grows to full size. The parasitoid then resumes development and consumes the aphid from the inside, causing it to become a husk, cemented to a plant surface. Inside this husk, or mummy as it is normally called, the primary parasitoid pupates. Aphid primary parasitoids are attacked by two guilds of secondary parasitoid, a taxonomically diverse group that lays its eggs in the primary after mummification, and the Alloxystinae. The latter do not attack the mummy, but live aphids containing larval parasitoids in arrested development. Using semiochemical signals, in a way that is not very well understood, the alloxystine is able to detect and distinguish a parasitized aphid, and then somehow to place its egg within the relative tiny body of the first parasitoid inside the aphid. The alloxystine first instar larva hatches and itself suspends development, resuming growth only when the primary parasitoid has caused the host to mummify (Sullivan 1988).

These wasps are important to us because for the last ten years we have been using aphids and their primary and secondary parasitoids as a model system to investigate general issues in commu-

nity ecology (e.g., Müller, *et al.* 1999). We are particularly interested in indirect effects: how the population dynamics of two aphids that feed on different host plants and so never interact directly may be coupled through their shared natural enemies (or, similarly, how two primary parasitoids that attack different aphid species may be linked by shared secondary parasitoids). To assess the potential for indirect effects we have built quantitative food webs each summer month of the last ten years of the aphid-parasitoid community in a field site in southern Britain (Müller *et al.* 1999). The web is used to design field manipulation experiments to test specific hypotheses about indirect effects (Morris *et al.* 2001). It is clearly important for us to get the taxonomy right.

Over ten years we have recorded about 40 aphid species from our site and about 30 primary parasitoid species. Of the secondary parasitoids, eight species attack the mummy, which leaves the Alloxystinae. At the start of the study, I thought that identifying the Alloxystinae would be relatively easy as there was a modern key to the British species. However, it was worrying that several long-series of reared specimens that seemed *prima facie* distinct (related hosts, morphology very similar) keyed out to the same species. Moreover, the only other alloxystine taxonomist in Europe maintained a much narrower species concept, though this was apparent only from notes and short papers as he had never attempted a Europe-wide revision. He used character states ignored in the British review which had sunk, without explanation, a number of species which he considered distinct. So which was correct, continental splitting or British lumping?

Frank van Veen joined my group from Holland and using a combination of molecular methods and careful study of morphology unambiguously, I believe, resolved this issue in favor of the continent. Sequencing the ITS region provided excellent distinguishing markers (van Veen *et al.* 2003), and based on this he was able to assess critically which morphological characters were most useful and hence write a traditional key. The British revision listed 16 species of *Alloxysta* from the United Kingdom while now we believe there are at least 18 in our food web (with at least another eight attacking tree-feeding aphids at our site that we do not include in the web). There were also some surprises, to our knowledge, the opening couplet of every key that has ever been written to *Alloxysta* begins with "wings fully developed/wings missing or shortened" or something similar, dividing the genus into winged or unwinged species. Frank found that four out of the five European unwinged "species" were in fact polymorphic, with all (two cases) or some (two cases) females winged. (The fifth species lives on the Arctic island of Svalbard; it has never been reared but presumably attacks the one species of aphid on the island, curiously, no primary aphid parasitoid has yet been found).

Van Veen's key initially included the species as codes, but he was soon able to assign about half the taxa to the few comparatively well known species, and with the help of Henk Evenhuis, the long-retired continental expert, to the species that he had studied over the years. But what of the rest? It was helpful that an American taxonomist had catalogued the names that had been applied to the world's Alloxystinae giving authors and, where known, the location of the type specimen. Many were telegraphically described by Keiffer in the early years of the twentieth century; while several species, secondary parasitoids of important pest aphids, had been described over and over again (Evenhuis had labored hard to sort out, successfully, many but by no means all, these problems). Were our species "v2, v3, ram1, br1, br2", etc. amongst these names floating in taxospace? More ambitiously, could we use our findings as a basis for a modern revision of the group?

The original descriptions and most early work (before Evenhuis) were essentially useless. Seldom were new species compared with others, and hardly ever were those characters described which today are recognized as the most helpful. The only way forward is to examine the types, distributed throughout the museums of Europe (with a few having found their way to the Smithsonian). Examining 100-year old alloxystine types is not fun (a referee of this article, correct-

ly, points out that this is a value judgement — but for the reasons that follow I claim it the status of a universal truth!). They look nothing like recent specimens, and usually have a washed out appearance that can make discerning critical characters almost impossible. Good bird watchers identify species by their “jizz”, the sum of the animal’s color, morphology and behavior, and entomologists who see many specimens of an insect species, even dead under the microscope, develop an equivalent skill. But while this is of huge value for recent specimens, it is rarely much use for studying old types. A different almost forensic approach is needed, the analogy seeming particular apt as often (>50%) the specimen is damaged, sometimes severely so. Moreover, frequently it is not clear what specimen actually is the type; it may be missing or poorly labeled, or several specimens may be mounted together, not infrequently of different species.

Sorting out this mess would be a huge undertaking, involving much travel, much scholastic antiquarianism, and little science. A depressing amount of entomological taxonomy, especially in Europe, consists of long and lengthy discussions of this type of taxonomic book-keeping (to avoid this, some of the best taxonomists I know work only in the tropics where they can be biologists rather than archivists). For Frank, pursuing this did not seem a great career move for a recently graduated student; in fact, the choice wasn’t there — no one would fund it. We never submitted a grant, but I did ask a senior science funding figure what would have happened if we did. His answer was clear — “why should we spend science money on a project that will be 90% book keeping and result in a technical publication in an obscure journal that will not be used by anyone except other taxonomists and which will be out of date in a decade”. Perhaps this overstates the case, but it is hard to argue that it is good value for money. I do not think the Alloxystinae will ever be revised as things stand at the moment.

It is dangerous to draw too wide conclusions from a single group (or, depending on your viewpoint, from what one critical reviewer called an unsubstantiated diatribe), though the Alloxystinae are far from unique in their intractability. But I derive two main lessons from this experience. First, the way we do taxonomy now, as encapsulated in the Zoological and Botanical Codes, may be part of the problem with attracting new funds to the field. The Linnaean system and the Codes have done a fabulous job in stabilizing nomenclature and have created a classification of the living world that is one of the triumphs of modern science. But these mechanisms were honed in an age of paper and post, and makes poor use of modern information technology. As I shall argue below it can be done better today, and the transition to something better can be carried out in a way that preserves the best of 250 years of Linnaean taxonomy but which jettisons some of the accreted historical baggage. Second, if taxonomy does not change to make it more relevant to the way that science is done now it will be replaced. Our modest use of molecular techniques helped immeasurably to sort out difficult issues in the Alloxystinae and the coming couple of decades will see radical advances in the speed, cheapness and accessibility of DNA sequencing. I think that there is a real possibility that massive parallel sequencing of genes from huge numbers of individuals will be able to create a classification from scratch. Whether this replaces or meshes with the Linnaean system depends on how the field evolves in the next ten years or so.

UNITARY TAXONOMIES

One of the difficult things today for consumers of systematics is the distributed nature of the taxonomy. By this I mean that the taxonomy of any particular group does not exist in a single location but is an ill-defined integral of all the papers, books and monographs on the taxon, backed up by types in museums and herbaria. To be an expert in a group is as much to know the topography of this information landscape as it is to be able to identify different organisms, recognize new

species, and determine their phylogenetic relationships. The complexity of this information excludes the majority of non-specialists from the subject, with two consequences. First, they can only use taxonomic resources when higher-level works such as accessible monographs, floras and field guides are available (and are reasonably up-to-date). For huge stretches of the animal and plant kingdoms they are absent. Second, the work of taxonomists impinges little on the daily lives of other biologists, who thus insufficiently appreciate its importance.

Another difficulty for the non-specialist is divergent taxonomic opinion. In many groups there are competing views about how best to classify different taxa, and for some — groups of cacti, orchids and butterflies, for example — this can be a nightmare. When faced by criticisms from other biologists about this lack of consensus taxonomists often reply by pointing out, rightly, that every taxonomic scheme is a provisional hypothesis, and that the presence of alternative concepts is part of the healthy sorting out of ideas that occurs in all the sciences. Indeed, it is proof, if proof is needed, that taxonomy is a vibrant science and the discord and discourse should be encouraged. Criticisms of this type are sometimes characterized as naïve and stemming from a simplistic and old-fashioned view of taxonomy.

I fear that taxonomists who respond in this manner underestimate and in some cases patronize their critics. Most people who use taxonomies understand their provisional nature, and accept that they will change as knowledge accumulates (though they also tear their hair out when a name changes purely because some archaeological research in a museum has reinterpreted an overlooked type or discovered an early description). But I think it is perfectly reasonable to ask at any one time for a provisional consensus treatment that can be used consistently by consumers. This sometimes exists *de facto*, where there is a single authoritative revision or monograph, but the Codes have no mechanism for supplying a consensus, and give no guidance as to how the end-user should treat radically different taxonomic opinions.

My view is that for taxonomy to be better appreciated and better funded it must address the issues of its distributed nature, its lack of consensus, and elements of its historic baggage. It must do this in an evolutionary manner that preserves the best of its immense achievements. I am sure there are different ways this could be accomplished, but here I am going to discuss one idea which I have called “unitary taxonomy” (Godfray 2002a). I make no pretence that is the optimal or even a good candidate solution, but I hope that discussing its multiple failings might suggest a better way of proceeding.

My argument is that we need to move from the distributed taxonomies of today, only fully usable by those with access to large, specialized libraries, to unitary taxonomies fully accessible on the web. By unitary taxonomies I mean a systematic treatment of a significant taxonomic group of organisms at a global scale that includes the taxa descriptions and diagnoses that one finds in traditional taxonomic works, supplemented where necessary by all the kinds of information resources that the web can provide. The unitary taxonomy site would be a “one-stop shop” for taxonomic information on that group, and would contain both the current consensus taxonomy for the user community, as well as alternative hypotheses for future research by the taxonomic community.

How might this work in practice and, in particular, how would the initial transition from a distributed taxonomy to the “first web revision” occur? Let us suppose that one of our great museums offered to host the unitary taxonomy of the Culicidae, the mosquitoes — a group whose taxonomy is both difficult and of immense importance in disease control (I shall return later to issues of determining who should host what revision). The museum would then set up a committee to co-ordinate the preparation of the first web revision, the group acting rather like the editorial board of a scientific journal. The revision would consist of separate “pages” for each taxon — species, genus, etc. — which at the minimum would contain the information required for a current species or higher

category description. That is, a full description with the designation of a type specimen. An advantage over the current system is that the committee would almost certainly want to set a standard for descriptions so that information on certain characters is consistently available across all groups. One could envisage character matrices being automatically constructed, though other information would also be included. Mosquitoes, like other groups, have a large excess of names over valid species and the first web revision would also need to include "pages" for these synonyms, etc. linked to valid taxa. As this information accumulates for different genera, tribes and subfamilies it would be mounted on the web and offered for web-based, public refereeing by the mosquito-systematics community. The process of refereeing would be handled by the taxon committee who would determine the consensus taxonomy that would become the first web revision. Critically important though, the process of deciding on a consensus should not exclude divergent views. Any alternative hypothesis should by right be mounted on the unitary taxonomy site where it will be available for research and discussion and possible incorporation into a future revision.

How would taxonomy proceed after the transition to the web? Most importantly, the first web revision would determine the set of names and taxa that need to be considered in any future research. If I think I have found a new species I need only consider the taxa (and names) on the first web revision in deciding whether it is actually undescribed. At a stroke, much of the sterile part of modern taxonomy is dispensed with — the unprofitable searching through old literature and collections. After the first web revision, the complete relevant literature for the group will be at a single site. For sure, consultation of types will continue, though increasingly high quality photographs, produced by systems with perfect depth of focus, will be mounted on the web and actually be easier to use than the original specimen.¹ Suppose a new species is described on the web and then subsequently a 19th century description is unearthed. At present, the old name would take precedence (or would need to be formally suppressed), but in this system it would have no status beyond that of a historical footnote that might be included on the species web page.

Taxonomy does not stand still and new research will require changes to the first web revision. I envisage new species and revisions being mounted on the web for public comment and refereeing, and then periodically a new consensus "current web revision", arbitrated by the taxon committee, being "published" (with rejected hypotheses still being available). This would replace but be linked to the previous set of revisions. A great advantage of web-based taxonomies is that species concepts can be followed both backwards and forwards through time. Thus, suppose that I, a non-taxonomist, publish a paper in which I refer to a particular species of mosquito. Today I might give its Linnaean binomial and author but were there a unitary taxonomy I would refer to its URL (or more likely a more sophisticated form of digital object identifier, DOI). The reference would link specifically to a taxon and revision edition. If in the future the species concept or taxonomic context changes then I can use the DOI to trace forward through time to discover what name people are currently referring to this species by. Similarly, given a species of interest it would be possible to go back through time to pick up references to it by whatever it used to be called. Today journals tend to be rather unclear and lax about how they demand species names to be referred to, while normally being very clear that references to DNA sequences, for example, have a database reference. A unitary taxonomy would provide a simple and valuable means of clarifying what the author means when he or she uses a binomial, while incidentally reinforcing the continuing relevance of taxonomy for the whole biology community.

Web revisions, as I have just described them, are essentially taxonomy as currently practiced.

¹ A unitary taxonomy could operate perfectly well with the current system of holotypes, paratypes, lectotypes, etc. I suspect that increasingly there will be a movement to choose new replacement modern types that are both fresh and undamaged for image capture and DNA archiving.

but placed within a modern digital information framework, and with the novel addition of the consensus taxonomy. Once such a unitary taxonomy had been set up, it could be extended very easily to provide much greater information resources. The most obvious extension is to provide more space for descriptions. Taxonomic works are often famously difficult to use, employing a recondite vocabulary, impenetrable to the non-specialist. Part of the difficulty is the genuine subtlety of species distinctions and the requirement for expertise in the group before safe diagnosis is possible. But much of the difficulty also stems from the requirement, imposed by the economics of publishing, for taxonomists to be brief, even telegraphic, and to use a minimum of expensive illustrations (ideally not in color). The marginal costs of extra storage on web servers is minimal and this will give taxonomists the space they need to be understood.

Descriptions are not the only information about a species that would be included on its web page. It will of course be straightforward to link to phylogeny projects, for example the "Tree of Life" initiative <<http://tolweb.org/tree/phylogeny.html>>. Many other sorts of data might be mounted there, and further resources might be accessed through links. Returning to our mosquito example, almost certainly one would want to accumulate information about early stages — eggs, larvae, and pupae — all of which are important for vector biologists. Cytology, and increasingly molecular biology, have played major roles in mosquito taxonomy, and for critical groups there would be illustrations of polytene chromosomes with inversion patterns, and sequence information for relevant loci. While the first web revision might just give fairly general information about geographic distribution, as museum collections are digitized and linked to GIS information one could easily imagine more sophisticated treatments of distribution emerging with unitary taxonomy web sites interrogating museum sites that link to them and automatically generating maps of known distributions. As authors elect (and journals require) species names to be linked to the web site then the unitary taxonomy will become ever more useful in accessing the primary literature concerning different taxa. Increasingly, taxonomists and the institutes they work for would move from being curators of specimens to curators of the body of information about their chosen groups.

The raw descriptions and associated data are of value to taxonomists and, to a more limited extent, other biologists, but to be maximally useful a unitary taxonomy site would need higher-level treatments of these basic data. A simple step would be to allow the selection of taxa from restricted geographic distributions so that to identify a mosquito from my garden in England I would not have to wade through Australian and African species. But the web site should also attract more sophisticated identification tools such as traditional and multi-access keys that can make use of the library of visual images stored at the site, and include as well the type of information that can be found in regional floras and field guides. Such resources would be directed at both the professional and amateur biologist, and at a yet higher level of abstraction one might want to encourage applications aimed at a very general audience: to enable a member of the public to work out the mosquito biting him or her on their porch, or to identify the pretty moth at the light. The ultimate aim would be to have a democratized site that anyone could enter at the appropriate level, and once in could then delve as deeply as they required. Apart from its obvious usefulness, such a site would again reinforce the contemporary relevance and importance of taxonomy.

What are the downsides of unitary taxonomies (Knapp et al. 2002; Thiele and Yeates 2002; Scoble 2004)? Critics have accused it of being authoritarian and at variance with the *laissez faire* traditions of taxonomy. I hope the arguments above dispel some of these concerns, in particular the worry that alternative views might be excluded by a taxonomic clique. But it is true that one particular set of hypotheses will be designated the current consensus taxonomy. I argue that the increased engagement with the user community is a price worth paying for this extra burden (Godfray 2002b).

To my mind, a much more serious concern is the costs of maintaining and curating a unitary taxonomy web site. There are hardware costs in setting up the site, but probably more important are the salary costs of its maintenance. An advantage of the current system is that if no one is interested in the taxonomy of a particular group it can quietly snooze in the library; a unitary taxonomy is harder to put on one side. I believe that the task of creating and maintaining unitary taxonomies should be taken on by our major museums and herbaria, all of whom, virtually without exception, have enthusiastically embraced the web. It is a natural extension of many of their current activities. But, indubitably, for them to do it would require new funds.

Were unitary taxonomies (or a related idea) thought to be worth pursuing then the only sensible way of proceeding would be to experiment with one or a few taxa, perhaps of the size of the Culicidae. If the project failed, or the resulting resource was judged poor value for money, then no bridges are burnt and the original distributed mode of taxonomy can easily be re-adopted. But my hope is that the clear benefits of the unitary site would leverage the new monies for taxonomy that would be required to extend the project. I think it would generate the constituency that would support and militate for more taxonomy. Converting all of taxonomy to the web will take a very long time, but during this period both the current and unitary models can co-exist for different groups.

There are a series of other issues that would need to be overcome. Some body, the equivalent of those administering the Zoological and Botanical Codes, would need to approve an application from an organisation to mount a unitary taxonomy, and to ensure they have an appropriate taxon committee, long-term commitment, etc. To a certain extent this would be self-policing as a site would only be declared a unitary taxonomy after it had produced the necessary resources, but the central body would probably need to give the final approval. Another issue that is often raised is the problem of backup and ensuring that the information is not lost. While a real concern, I believe it easily dealt with; there are not hard copies of many of the genome-scale data bases of molecular biology.

A further criticism is that unitary taxonomies are elitist because they require access to the web. I think this is daft; distributed taxonomies are extraordinarily difficult to get hold of. Even if you can list all the relevant papers and books, they are frequently in obscure journals or very expensive monographs. Few people have easy access to the small number of taxonomy libraries that have a reasonably complete coverage of the literature, and, for those who don't, assembling the resources to do taxonomy is difficult and costly. A unitary taxonomy web site should be open access to everyone, and in my view should also be downloadable so that it can be used in the field where internet access is impossible. And as a final option, all or part of it can be printed out.

A more serious criticism is that of language. I think taxonomy is a science and today, inescapably, the *lingua franca* of science is English — unitary taxonomies, at least the taxon pages, should be in English and replace the current polyglot distributed model. At its worst, this means taxonomists learning one extra language instead of the several they typically require now (though as a monoglot English speaker I feel decidedly uncomfortable in advancing this argument). But, though the basic data would be in English, there is no reason why the higher-level components, the keys and material aimed at the general public, should not be in whatever language or languages are most appropriate.

Having discussed the costs of adopting unitary taxonomies, I want to finish this section by arguing that there are serious costs for maintaining the *status quo*. The first of these I have already alluded to: Taxonomy is failing to generate enough products that are of value to its end users, and many are, at best, indifferent to taxonomy and are failing to argue for new monies going to the subject. I believe that unless taxonomy produces resources that are valued by the rest of biology it will wither on the vine.

An optimist might counter this argument by saying that we will always need a system of classification and hence that, imperfect as it is, taxonomy as currently practiced will muddle along, very much as it has over recent decades. Why tinker with something that works? But there is an alternative to Linnaean taxonomy looming on the not too distant horizon. The speed, efficiency and cheapness of DNA sequencing has steadily increased over the last two decades at a rate that shows no sign of declining. Already there are arguments that much species-level identification can be automated and done by sequencing: the DNA bar-coding and related initiatives (Hebert et al. 2003, 2004; Stoeckle 2003). At the moment, routine sequencing to identify individual specimens is normally not cost effective but this will almost certainly change in the coming years, and it is not a complete flight of fancy to imagine a future when species identifications can be done routinely by machines in the field (Janzen 2004): the tricorder solution (recall the useful gadget in the *Star Trek* program). Ideally, this new technology will interdigitate seamlessly with current Linnaean-based methods, and if unitary taxonomies are adopted this will be particularly easy. I worry that if taxonomy continues as it is, in its current complex, distributed mode, then it will not mesh with DNA-based methods and be replaced by them. It will be easier for many groups to start from scratch and define species and higher taxa as clumps in sequence space than to try to relate them to the existing system. Would this matter? I think it would. We would throw much of the biological knowledge that we have accumulated over the last 250 years, a large amount of which we would have to discover anew. And we would also divorce sequence-based identification from the more traditional methods based on morphology, that will ensure continue, at least among the amateur consumers of taxonomy.

A final point: there are already a wide variety of interesting and important taxonomic resources on the web, though none I believe approaching a unitary taxonomy website as I have described here. To explore these sites a good point of entry is the Global Biodiversity Information Facility (GBIF) portal, <<http://www.gbif.org/>>.

NEW RANK-FREE TAXONOMIES

While unitary taxonomies seem revolutionary to some they are tame compared with some of the other alternatives to Linnaean taxonomy that are being advocated at the moment. Perhaps the most radical is the replacement of Linnaean binomials and taxonomic hierarchy with a rank-free system based purely on phylogenetic principles and regulated by a new set of naming rules, termed the PhyloCode <<http://www.phylocode.org>>. The logic behind this is to acknowledge the arbitrary nature of taxonomic ranks (including, depending on your point of view, species) and to incorporate rigorous phylogenetic definitions of different taxa. While acknowledging the arbitrary component of assigning ranks, and the importance of rigorous phylogenetic methodology, my view is that switching to such a system would be an unmitigated disaster that would risk the destruction of taxonomy.

Taxonomists are often unfairly caricatured as inward-looking and obsessed with changing names for reasons that have little to do with biology, and a lot to do with the formal rules of systematics. Such a wholesale revision of names as the PhyloCode envisages would confirm all the worst prejudices that outsiders have about the subject and would probably destroy its funding base. There are also technical objections to the PhyloCode (see, e.g., Forey 2002; Wheeler 2004) concerned with the instability of nomenclature as evolutionary relationships are resolved or new taxa are discovered that render old ones paraphyletic. In addition, an extremely important gift of the Linnaean system to the rest of biology is a hierarchical series of partitions of biodiversity, however imperfect that might be. An ecologist can work with the generic and family diversity of organ-

isms at one field site, and compare it with that at another. He or she knows that taxon distinctions are arbitrary, but this is relatively unimportant in comparative studies using the same taxa. Similarly, information can be organized at different levels; keys exist to the families of flowering plants, and many major garden plant encyclopedias are arranged as alphabetical lists of genera. Were taxonomists to abandon ranked taxa, they would be re-invented by ecologists, naturalists and gardeners. I suspect that there will continue to be a need for hierarchical partitions and their provision is best left to taxonomists. These all seem to me compelling arguments against kicking over the traces and abandoning the Linnaean system for something untried and with questionable general support. And from a phylogenetic perspective what matters is tree topology rather than the naming of nodes: it is not as if we cannot have our cake and eat it too. Linnaean taxonomies can be linked with phylogenetic hierarchies so that you have the benefits of both. Perhaps unsurprisingly, I would argue this is another benefit of a unitary taxonomy.

CONCLUSIONS

For taxonomy to survive and prosper in the 21st century it needs to identify better its end users and its functions. For the taxonomy of living plants and animals, with which I have been most concerned here, the ultimate function of much of taxonomy is to enable further research in whole-organism biology, and of itself to produce a classification of the world's biodiversity with an account of their interrelationships and the tools required for their identification. There are major constituencies to support investment in both functions. The first constituency is other whole-organism biologists and possibly biologists in general. The second constituency is potentially everyone, from professional biologists to anyone who has ever wondered about the name or natural of a particular plant and animal. There should be a chorus clamoring for more taxonomy, with a much broader base than the relatively narrow interest groups that are currently arguing for greater funds. I have no illusions that unitary taxonomies as I have described them are the solution to these problems, but I do believe that some much better idea, that would convert the current distributed mode of taxonomy to a one-site form that is much more easier to use by the community, will be at least part of the answer.

LITERATURE CITED

- FOREY, P.L. 2002. PhyloCode — pain, no gain. *Taxon* 51:43–54.
- GAULD, I.D., AND B. BOLTON. 1988. *The Hymenoptera*. Oxford University Press, Oxford, England, UK. 332 pp.
- GODFRAY, H.C.J. 2002a. Challenges for taxonomy. *Nature* 417:17–19.
- GODFRAY, H.C.J. 2002b. Towards taxonomy's 'glorious revolution'. *Nature* 420:461.
- HEBERT, P.D.N., A. CYWINSKA, S.L. BALL, AND J.R. DEWAARD. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London, Series B: Biological Sciences* 270:313–322.
- HEBERT, P.D.N., S. RATSINGHAM, AND J.R. DEWAARD. 2004. Barcoding animal life: cytochrome C oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London, Series B: Biological Sciences* 270:S596–299.
- JANZEN, D. 2004. The future of taxonomy. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 359:731–732.
- KNAPP, S., R.M. BATEMAN, N.R. CHALMERS, C.J. HUMPHRIES, P.S. RAINBOW, A.B. SMITH, P.D. TAYLOR, R.I. VANE-WRIGHT, AND M. WILKINSON. 2002. Taxonomy needs evolution not revolution. *Nature* 419:559.
- MORRIS, R.J., C.B. MÜLLER, AND H.C.J. GODFRAY. 2001. Field experiments testing for apparent competition between primary parasitoids mediated by secondary parasitoids. *Journal of Animal Ecology* 70:301–309.
- MÜLLER, C.B., I.C.T. ADRIAANSE, R. BELSHAW, AND H.C.J. GODFRAY. 1999. The structure of an aphid-parasitoid community. *Journal of Animal Ecology* 68:346–370.

- SCOBLE, M.J. 2004. Unitary or unified taxonomy? *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 359:699–710.
- STOECKLE, M. 2003. Taxonomy, DNA, and the bar code of life. *BioScience* 53:796–797.
- SULLIVAN, D.J. 1988. Aphid hyperparasites. Pages 189–203 in A.K. Minks and P. Harrewijn, eds., *Aphids, Their Biology, Natural Enemies and Control*. Volume 2B. Elsevier, Amsterdam, The Netherlands.
- THIELE, K., AND D. YEATES. 2002. Tension arises from duality at the heart of taxonomy. *Nature* 419:337.
- VAN VEEN, F.J., R. BELSHAW, AND H.C.J. GODFRAY. 2003. The value of the ITS2 region for the identification of species boundaries between Alloxysta hyperparasitoids (Hymenoptera Charipidae) of aphids. *European Journal of Entomology* 100:449–453.
- WHEELER, Q.D. 2004. Taxonomic triage and the poverty of phylogeny. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 359:571–584.



Godfray, H. C. J. 2005. "Taxonomy as Information Science." *Proceedings of the California Academy of Sciences*, 4th series 56, Supplement I(16), 170–181.

View This Item Online: <https://www.biodiversitylibrary.org/item/265952>

Permalink: <https://www.biodiversitylibrary.org/partpdf/283002>

Holding Institution

California Academy of Sciences

Sponsored by

California Academy of Sciences

Copyright & Reuse

Copyright Status: In copyright. Digitized with the permission of the rights holder.

Rights Holder: California Academy of Sciences

License: <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Rights: <http://biodiversitylibrary.org/permissions>

This document was created from content at the **Biodiversity Heritage Library**, the world's largest open access digital library for biodiversity literature and archives. Visit BHL at <https://www.biodiversitylibrary.org>.