



TAXONOMY IN INDIA IN THE 21ST CENTURY: CALL FOR A DIGITAL REVIVAL

Sectional Editor: Colin Groves

Submitted: 29 June 2011, Accepted: 30 November 2011

V. M. Sathish Kumar*

*Zoological Survey of India, Southern Regional Centre, 130 Santhome High Road, Chennai-600 028, India
E-mail: drsathishkumarvm@gmail.com

Abstract

The Biodiversity of the planet is under threat due to various reasons. For the conservation and management of the bio-resources we need to precisely identify a species, so that the species can be accessed and managed accordingly. This is the very first basic step every taxonomist has to take. Though Traditional taxonomy has been used in systematics for over 250 years with the identification of more than 1.75 million species, in many countries as in India, traditional taxonomy is yet to adopt the latest molecular techniques available for a better and accurate identification of taxa. Enhanced taxonomic infrastructure and research tools can certainly allow Indian taxonomists to speed up the process of species exploration, description and classification which will enable the World scientific community to get access to India's flora and fauna dynamically without excuses of reduced funding or dearth of taxonomists for not mapping the biological diversity of India. This article overviews the latest digital advancements to revolutionize taxonomical research in India in the current century.

Keywords: biodiversity, molecular systematics, internet, ATI, CAT,

Introduction

The biological diversity of earth is estimated to harbour as many as 30 million species of which about 1.75 million have been described. With about 28.25 million species yet to be described, our gap in taxonomic understanding is truly an issue, especially considering that India, one of the 17 mega-diversity countries, contains global hotspots of biodiversity which are unusually enriched with endemic species. This is attributable to India's unique biogeographical location and diversified

climatic conditions. India holds only 2% of the World's total land surface but harbours over 7.43% of the world's species of animals. As emphasised by Heywood (1995), this precious natural bio-richness is vulnerable; most species are yet to be described, and the majority of current extinctions are going unrecorded -- species are dying out before we even learn of their existence, and they need to be conserved and managed accordingly. Taxonomic information is essential to achieve this; the more

complete our taxonomic knowledge, the better is the foundation for safeguarding our faunal wealth. Sampling, identifying, and making accurate identifications of biological specimens are among the compulsory first steps towards protecting and gaining from biodiversity. The Millennium Ecosystem Assessment (2005) pointed out the challenges of taxonomists during the very long history of classification while discovering and describing the millions of species, with mounting pressure imposed through steep rates of species extinction and worldwide disturbance and degradation of ecosystems, and this is precipitated by the dearth/dwindling number of trained taxonomists, above all in India.

This calls for the urgent participation of more taxonomists than ever, particularly given that an estimated 90 per cent of all species remain undescribed (Wilson, 2003), and retiring taxonomists are leaving numerous 'orphan' taxa behind and with few students entering into the field of taxonomy (Godfray, 2002; Wheeler *et al.*, 2004). Seberg (2004) affirmed that for describing all the existing taxa on Earth, traditional taxonomy at the current rate will require more than 940 years before all species will be described. To a large extent taxonomy in India is practiced very 'economically' as it costs much less to procure collection nets, glass containers and dissecting microscopes than to fund an expensive molecular biology laboratory equipped with micro-centrifuges, thermal cyclers, gel documentation systems and high end DNA sequencers. Yet under the present circumstances, it is very essential that the traditional taxonomic community in India should incorporate these modern tools to speed up the process of describing and identifying species. In this way India's rich flora and fauna could well be managed and taxonomy in India can survive and flourish in the twenty first century.

Use of Molecular Systematics: As noted by Blaxter *et al.* (2004), the present worldwide domination of molecular techniques over comparative morphology and the recent proposals for DNA-based taxonomy have indicated a way for describing species precisely. In the same year, Janzen (2004) proposed to identify species from their DNA, an exciting new tool for taxonomic research. The DNA barcode is a very short, standardized DNA sequence in a gene, and allows one to identify the species to which a plant, animal or fungus belongs. The Consortium for the Barcode of Life (CBOL) is encouraging international ventures that will facilitate people in

all countries to recognize and protect their biodiversity. The global library of DNA Barcodes has produced an open access library of reference barcode sequences which allows even non-taxonomists to identify specimens. The barcode of an unidentified specimen can be compared with the reference barcodes to pinpoint the identical species. Besides this, many other techniques such as Restriction Fragment Length Polymorphism (RFLP), Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphism (AFLP), Single Nucleotide Polymorphisms (SNPs), Minisatellites and Microsatellites are also in practice today. The National Centre for Biotechnology Information (NCBI) is an excellent paradigm for the sequence library. The Basic Local Alignment Search Tool (BLAST) finds regions of local similarity between sequences and so is extremely useful to find out various associations of a taxon.

It has been suggested by Tautz *et al.* (2002 & 2003) that one way to remove the taxonomic impediment would be to convert to a taxonomy based on DNA sequences, rather than one based on morphology. The supporters of DNA barcoding (Tautz *et al.*, 2003; Hebert & Gregory, 2005) argue that the remarkable rise in sequencing facility and reduction in sequencing cost make it more extensively acceptable; the detection of new species is incredibly easy via analyses of one or a few standardized sections of the DNA molecule. Such sections, of course, do not embody all the genetic variations that may exist among any two species, but are adequate to 'type' different subspecific, specific and superspecific groups on the basis of phylogenetic resemblance. Adoption of a DNA-based approach for taxonomy and identification should be regarded as a community-wide priority for systematists because such a research program will support quicker and better identifications. The vast majority of all species known to science have been described on the basis of the morphological and not molecular characteristics; the progression of molecular techniques not only offers better ways to define taxa precisely but has also confirmed many well-founded species that were considered to be closely related but have been shown not to be so, and assuredly things will never be the same as they were. The pros and cons of 'DNA taxonomy' have been argued efficiently (Lipscomb *et al.*, 2003; Mallet & Willmott 2003) and it is not necessary to discuss them in detail here.

Use of Internet: Many would agree that the utilization of the Internet -- especially via the World Wide Web -- as proposed and first conceived in the 1980s by Berners-Lee (1999) is enormously helpful for making taxonomic decisions especially as of now; taxonomic information is certainly heading toward a Web-based system. For instance, a virtual, GenBank-like system for accessing morphological, audio and video data would be a fundamental step, because text-based descriptions alone will not deal with either the taxonomic hurdles or identification problems successfully. According to Scoble (2004), the internet is the fastest evolving medium for providing access to information currently distributed across the published paper-based literature, in unpublished archives, in curated collections and, increasingly, in personal or institutional databases. The Internet is no longer new to us early 21st century beings, but new internet technologies will continue to broaden access and render it more effective. New identification tools can allow anybody to make identifications, and numerous well-illustrated interactive online keys are now in use for a variety of taxa and the software for such practices is being rapidly enhanced and easily accessible to all.

Use of Automated taxon identification (ATI): As MacLeod (2008) noted that the automated identification of objects is not a new concept, only a concept that is new to many systematists, and this is part of computer-aided taxonomy. In reality, many active taxonomists still consider that automated identification systems are sort of fantasy story in science. Many users simply want to identify specimens accurately and discover which recognizable species it belongs to. Traditionally, systematists have not relied on molecular data to identify taxa, but have chosen the visual inspection of morphology and the comparison of these to type specimens. Thus, as noted by Riede (1993), with the use of automated identification systems, a taxon is identified automatically, so helping taxonomists in systematics and for the assessment of biodiversity monitoring and conservation.

Katsinis *et al.* (1984) carried out the earliest ATI on marine zooplankton using image processing, and as recent as September of last year David *et al.* (2010) correctly identified 94.5% of benthic invertebrate images by using the BugID ATI system from 9 larval stonefly taxa, even though small or damaged specimens were included in testing. Farr & Chesmore (2005) did the same for Coleoptera; and it has been done for Orthoptera (Chesmore &

Nellenbach, 2001; Dietrich *et al.*, 2003; Ohya & Chesmore 2003), cicadas (Ohya 2004) and mosquitoes (Campbell *et al.*, 1996). Bacteria have also been used in ATI research (Walker & Kumagai, 2000; Foreroa *et al.*, 2004). Comparatively little work has been carried out on "higher" animal groups such as vertebrates. Amphibians such as cane toads were identified with the help of acoustic inputs (Taylor *et al.*, 1996). Birds were also identified (Mills, 1995) using acoustic ATI. Some such acoustic studies are currently under way by Indian taxonomists as well. Early this year ATI have been developed and tested successfully for teleost fishes using an otolith contour online database, in which 1480 images of left saccular otoliths (sagittae) from 420 species and 72% of specimens were correctly identified at species level, and this percentage increased to 90% at genus level and reached 94% at family level (Parisi *et al.*, 2010). Identification of mammals includes Roe deer (Reby *et al.*, 1997, 1998a,b), domesticated cows (Jahns *et al.*, 1997), and false killer whale (Murray *et al.*, 1998a,b). Bats are another target for ATI; ultrasonic echolocation calls are generally species specific (Parsons & Jones 2000; Parsons 2001), and infrasound has recently been used for identification of elephant calls (Clemins & Johnson, 2002).

Use of computer-aided taxonomy (CAT): Over the past decade several renowned taxonomists and a number of organizations at both national and international levels entrusted with cataloguing biodiversity (e.g. BioNET and INBIO) have added their voices to the rising demand for computer-aided taxonomy (CAT), and now the biological community is gifted to use tools like DAISY, SPIDA (Do *et al.*, 1999; Russell *et al.*, 2005) and ABIS (Schroder *et al.*, 1995) for the identification of various taxa.

Such tools would have a profound impact on both taxonomy and biodiversity, and it is very simple to correlate hand held devices linked via satellite to database libraries of taxonomic knowledge and images. The extensive usage of portable computers and mobile phones with high-resolution cameras, large memory and efficient processing power, offers easy execution of ATI on these devices; this has several major advantages for use in the field, such as automatic location recording with GPS and the capability to pass on images or other relevant information regarding species identification. A single such device would make it easy for the naturalist to identify whatever species she/he comes

across in the forest, for a farmer to find out whether a strange insect is a pest, and for an ecologist to determine what species he has surprisingly come across (MacLeod 2008). All the devices necessary for the achievement of such practices are already obtainable with the latest communications technologies, already available in many countries and expanding in others, and utilization of such systems will certainly contribute towards the betterment of taxonomic research in India.

Conclusion

Taxonomic research is speeded up with the advanced use of digital tools for species identification and to restructure many phylogenetic questions. A considerable number of potential applications are being realised by information revival. Easy to use interactive electronic keys with digital illustration, requiring no prior morphological vocabulary, are improving all the time and are downloadable to a laptop or handheld devices. Experts are easily accessed via the Web anywhere in the world to make crucial identifications, consult or learn (MacLeod 2008). Reviving taxonomy by no means indicates that its value should be negotiated, but it does denote that certain applications will be altered. For example, it will be possible to have high-quality taxonomy posted dynamically on the web without having to wait for a completed taxonomic revision and have it printed and published.

The wish for a new systematics does not ignore the old system of classification. It has been classical taxonomy, with its scholarly resources stocked by morphological taxonomists for 250 years, which paved the way for new methods of analysis. Taxonomy as a discipline requires data from numerous sources like comparative morphology, developmental biology, palaeontology, molecular genetics and comparative ethology. Every phase of taxonomy can benefit from and be accelerated by the development, adoption and application of new digital tools, from field inventories to the study of specimens and characters, analysis of cladistic relationships, erection of formal classifications, collection, curation and species identifications (Wheeler, 2007).

If most scientists agree upon anything pertaining to taxonomy, it is the lack of funding. Financial backing is necessary to educate fresh taxonomists and to offer facilities and resources to the taxonomists that already exist.

Finally, taxonomy should not remain at the level of only identifying dead specimens, and Indian taxonomists need to adapt to existing recent technologies and prepare for new developments; as we get through the twenty-first century, the demand for taxonomy is greater than ever before with impending impacts of global warming and accelerating rates of extinction. The managers of taxonomic institutes and universities in the country should encourage methods such as taxic identification within the field, essential data collection without killing specimens for study wherever possible. In addition alpha-taxonomists with adequate field experience and professional molecular expertise should be encouraged, and such facilities should be provided for solving taxonomic ambiguities. The backlog caused by reduced funding cannot serve as an excuse for long, especially when taxonomy is overdue for a revival from the way it is practiced in India, failing which thousands of species will remain unknown or unidentifiable, inaccessible to science and society. It is certain that some of the innovations is going to influence taxonomy in the next decade or so.

Thanks to information technologies and molecular genetics, we have innovative tools to do our work faster and better than ever before. At this juncture one simple question remains - are we the naturalists with adequate vision to tie together classical taxonomy with advanced digital tools for a better understanding of our biodiversity? I hope this article will hasten the process of revolution of taxonomic research in India. Future generations will look back with grief and anger if we fail.

Acknowledgements

The author is thankful to Ted Taylor (University of Birmingham, UK), Domino Joyce (University of Hull, UK), K. Rema Devi (ZSI, Chennai) and K. Venkataraman (ZSI).

Literature Cited

- Berners-Lee, T., 1999. Weaving the Web: The Original Design and the Ultimate Destiny of the World Wide Web, Harper-Collins Inc., New York: 256.
- Blaxter, M., B. Elsworth, and J. Daub, 2004. DNA taxonomy of a neglected animal phylum: An unexpected diversity of tardigrades. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 271: S189-S192.

- Campbell, R. H., S. K. Martin, I. Schneider and W. R. Michelson, 1996. Analysis of mosquito wing beat sound, 132nd Meeting of the Acoustical Society of America, Honolulu.
- Chesmore, E. D. and C. Nellenbach, 2001. Acoustic methods for the automated detection and identification of insects. *Acta Horticulturae*, 562: 223-231.
- Clemins, P. J and M. T. Johnson, 2002. Automatic type classification and speaker verification of African elephant vocalizations. *Animal Behaviour*, Vrije Universiteit, Amsterdam.
- David A. L., G. Martínez-Muñoz, W. Zhang, N. Larios, L. Shapiro, R. Paasch, A. Moldenke, E. N. Mortensen, S. Todorovic and T. G. Dietterich, 2010. Automated processing and identification of benthic invertebrate samples. *Journal of the North American Benthological Society* Sep: Vol. 29, Issue 3: 867-874 doi: 10.1899/09-080.1.
- Dietrich, C., G. Palm and F. Schwenker, 2003. Decision templates for the classification of bioacoustic time series. *Information Fusion*, 4: 101-109.
- Do, M. T., J. M. Harp and K. C. Norris, 1999. A test of a pattern recognition system for identification of spiders. *Bulletin of Entomological Research*, 89: 217-224.
- Farr, I. J. and E. D. Chesmore, 2005. Acoustic detection and recognition of wood-boring insects. *Royal Entomological Society National Meeting: Entomology 2005*, 12-14 September, University of Sussex, Royal Entomological Society.
- Foreroa, M. G., F. Sroubek and G. Cristobal, 2004. Identification of tuberculosis bacteria based on shape and color. *Real Time Imaging*, 10: 251- 262.
- Godfray, H. C. J., 2002. Challenges for taxonomy. *Nature*, 417: 17-19.
- Hebert, P. D. N. and T. R. Gregory, 2005. The promise of DNA barcoding for taxonomy. *Systematic Biology*, 54: 852-859.
- Heywood, V. (Ed.), 1995. *Global Biodiversity Assessment*. Published for the United Nations Environment Programme. Cambridge University Press, Cambridge, UK: 1140.
- Jahns, G., W. Kowalczyk and K. Walter, 1997. An application of sound processing techniques for determining condition of cows. In *4th International Workshop on Systems, Signal and Image Processing*, May 28-30, Poznan, Poland, Institution of Electrical and Electronic Engineers, 1- 4.
- Janzen, D. H., 2004. Now is the time. *Philosophical Transactions of the Royal Society of London, Series B*, 359: 731-732.
- Katsinis, C., A. D. Poularikas and H. P. Jeffries, 1984. Image processing and pattern recognition with applications to marine biological images. In *SPIE 28th Annual International Technical Symposium on Optics and Electro-optics*, San Diego, Society of Photo-optical Instrumentation Engineers: 150-155.
- Lipscomb, D., N. Platnick and Q. D. Wheeler, 2003. The intellectual content of taxonomy: A comment on DNA taxonomy. *Trends in Ecology & Evolution*, 8: 65-68.
- MacLeod, N. (Ed.), 2008. *Automated taxon identification in systematics: theory, approaches and applications*. CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton: 368.
- Mallet, J. and K. Willmott, 2003. Taxonomy: Renaissance or Tower of Babel? *Trends in Ecology & Evolution*, 18: 57-59.
- Millennium Ecosystem Assessment, 2005. *Ecosystems and human well-being: biodiversity synthesis*. World Resources Institute. Washington, DC.
- Mills, H., 1995. Automatic detection and classification of nocturnal migrant bird calls. *Journal of the Acoustical Society of America*, 97: 3370-3371.
- Murray, S. O., E. Mercado and H. L. Roitblat, 1998a. Characterizing the graded structure of false killer whale (*Pseudorca crassidens*) vocalizations. *Journal of the Acoustical Society of America*, 104: 1679-1688.
- Murray, S. O., E. Mercado and H. L. Roitblat, 1998b. The neural network classification of false killer whale (*Pseudorca crassidens*) vocalizations. *Journal of the Acoustical Society of America*, 104: 3626-3633.
- Ohya, E., 2004. Identification of Tibicen cicada species by a principal components analysis of their songs. *Anais da Academia Brasileira de Ciencias*, 76: 441- 444.
- Ohya, E. and E. D. Chesmore, 2003. Automated identification of grasshoppers by their songs. *Annual Meeting of the Japanese Society of Applied Entomology and Zoology*, Iwate University, Morioka, Japan.
- Parisi-Baradad, V., A. Manjabacas, A. Lombarte, R. Olivella, Ò. Chic, J. Piera and E. García-Ladona, 2010. Automated Taxon Identification of Teleost

- fishes using an otolith online database-AFORO. *Fisheries Research*, 105 (1): 13-20.
- Parsons, S., 2001. Identification of New Zealand bats (*Chalinolobus tuberculatus* and *Mystacina tuberculata*) in flight from analysis of echolocation calls by artificial neural networks. *Journal of Zoology* (London), 253: 447- 456.
- Parsons, S. G. and Jones, 2000. Acoustic identification of 12 species of echolocating bat by discriminant function analysis and artificial neural networks. *Journal of Experimental Biology*, 203: 2641-2656.
- Reby, D., S. Lek, I. Dimopoulos, J. Joachim, J. Lauga and S. Aulagnier, 1997. Artificial neural networks as a classification method in the behavioural sciences. *Behavioural Processes*, 40: 35- 43.
- Reby, D., A. J. M. Hewson, B. Cargnelutti, J. M. Angibault and J. P. Vincent, 1998a. Use of vocalizations to estimate population size of roe deer. *Journal of Wildlife Management*, 62: 1342-1348.
- Reby, D., J. Joachim, J. Lauga, S. Lek and S. Aulagnier, 1998b. Individuality in the groans of fallow deer (*Dama dama*) bucks. *Journal of the Zoological Society of London*, 245: 79-84.
- Riede, K., 1993. Monitoring biodiversity, analysis of Amazonian rainforest sounds. *Ambio*, 22: 546-548.
- Russell, K. N., M. T. Do and N. I. Platnick, 2005. Introducing SPIDA-web: an automated identification system for biological species. *Taxonomic Database Working Group Annual Meeting*.
- Schroder, S., W. Drescher, V. Steinhage and B. Kastenholz, 1995. An automated method for the identification of bee species (Hymenoptera, Apoidea). *In Proceedings of the International Symposium on Conserving Europe's Bees, International Bee Research Association and Linnean Society*, London, International Bee Research Association, 6-7.
- Scoble, M., 2004. Unitary or unified taxonomy? *Philosophical Transactions of the Royal Society of London, Series B*, 359: 699-710.
- Seberg, O., 2004. The future of systematics: Assembling the Tree of Life. *The Systematist—Newsletter of the Systematics Association*, 2-8: 23.
- Tautz, D., P. Arctander, A. Minelli, R. H. Thomas and A. P. Vogler, 2002. DNA points the way ahead for taxonomy. *Nature*, 418: 479.
- Tautz, D., P. Arctander, A. Minelli, R. H. Thomas, R. H. and A. P. Vogler, A. P., 2003. A plea for DNA taxonomy. *Trends in Ecology and Evolution*, 18: 70-74.
- Taylor, A., G. Grigg, G. Watson and H. McCallum, 1996. Monitoring frog communities: An application of machine learning. *In Proceedings of the 8th Innovative Applications of Artificial Intelligence Conference (AAAI)*, 1564 -1569.
- Walker, R. and M. Kumagai, 2000. Image analysis as a tool for quantitative phycology, a computational approach to cyanobacterial taxa identification. *Limnology*, 1 (2): 107-115.
- Wheeler, Q. D., 2007. Digital innovation and taxonomy's finest hour. In *Automated Taxon Identification in Systematics: Theory, Approaches and Applications* (ed N. MacLeod), CRC Press, Boca Raton, FL, 9-23.
- Wheeler, Q. D., P. H. Raven and E. O. Wilson, 2004. Taxonomy: impediment or expedient? *Science*, 303: 285.
- Wilson, E. O., 2003. The encyclopedia of life. *Trends in Ecology and Evolution*, 18: 77-80.