

Towards a practical knowledgebase for marine genetic resources

Libby Evans-Illidge, Manager Bioresources Library, Australian Institute of Marine Science, PMB 3 Townsville MC, 4810, Queensland, Australia. e.evansillidge@aims.gov.au

The earth's most biodiverse habitats are marine. Around 75% of known eukaryotic phyla occur in the sea, many of them exclusively. This super-diversity of macro-organisms is overlaid with another order of magnitude in diversity of marine micro-organisms, often associated in complex symbiotic relationships. The combined diversity of marine genetic resources, which seems to be growing larger as new techniques emerge to measure it, is mirrored by an immense cache of molecules utilised by organisms to meet the challenges of life in even the most extreme environments. The resulting molecular arsenal is the result of 4 billion years of successful adaptation and evolution, and stands to benefit humankind through discovery of applications in parallel circumstances including in human health, biotechnology, and environmental change. Informatics to curate, catalogue, and facilitate maximum utilisation of the data generated by marine genetic resources research (biosystematics, microbiology, ecology, physiology, metabolite chemistry etc), is fundamental to maximising the benefits and maintaining investment in the research.

Perhaps one of the biggest obstacles to realising the potential of the globe's marine genetic resources is a continued incomplete understanding of the resource itself. The vast majority of the world's oceans remain inaccessible. For example, despite its relatively well resourced research effort, 70% of Australia's massive 16 million km² ocean territory remains unexplored at depths greater than 1500m which are largely beyond the national exploration capability. The costs associated with research on marine genetic resources especially in abyssal depths are considerable and above traditional sources of funding for the study of biodiversity, although funding associated with biodiscovery curation has assisted in some areas. However despite the obstacles and the scale of that which remains unknown, a global marine research effort has persevered since the famous Challenger expedition of 1872-1876, and today many hundreds of research organisations across the world undertake research on marine genetic resources.

The cumulative multi-century effort has produced collection and related information holdings at a range of formats and scales, from the ecosystem to the genome. This includes the knowledgebase of traditional owners of the resource; datasets, unpublished reports, and peer-reviewed journal articles by individual scientists; institutional collections and databases; and national and international networks. In many cases, due to the highly specialised nature of the taxonomy of a specific group, the information related to particular taxa reflects the locational history of the leading experts in that field. For example, the world's foremost scleractinian coral collection and associated database is held at the Museum of Tropical Queensland in Townsville Australia, alongside the world's most eminent experts in this field. Modern internet tools, compiled literature databases and metadata portals provide a *de facto*, practical method to at least identify the source, and in many cases search and access the contents of the collective knowledge base on marine genetic resources, despite its otherwise fragmented nature across all variables including locations, scientists, formats and taxa. In addition to this highly specialised and flexible (though inelegant) system, a networked project has commenced to compile and express global marine species data in a consistent format. The Census of Marine Life (www.coml.org) has engaged over 300 scientists in more than 50 countries in a ten-year initiative to 2010 to assess and explain the diversity, distribution and abundance of the world's oceans - past, present and future. Data from the CoML projects are being stored in the Ocean Biogeographic Information System (OBIS) (www.iobis.org), with a web-based interactive capability that enables mapping of records of each species. The Census also maintains the Biology pages of the UN Atlas of the Oceans (www.oceansatlas.org), and undertakes the marine arm of the Consortium for the Barcoding of Life (CBOL) (www.barcoding.si.edu).

Information about the location of marine species does not in itself provide an assessment of genetic resources and their potential, because molecules expressed by a species at any given time and location are the sum of its dynamic ecological setting. For example, bioactive secondary metabolites are energy-expensive to the organism producing them and will only be manufactured when needed, for example for defence or competition. The bioinformatics database associated with the Bioresources Library held at the Australian Institute of Marine Science, integrates species and site information with ecological observations, screening profile and chemistry. The result is a powerful tool for data-mining the results of 15 years of screening and collections data (10000 macroorganisms and 9000 microorganisms from 1500 sites around Australia), to identify trends in bioactivity with ecology and taxonomy. This presents a useful model for possible wider application. It facilitates analyses of the 'big picture'; to show that realisation of the potential of marine genetic resources is intimately linked to conservation not just of species in isolation, but in their fully dynamic and biodiverse context.