

Pumping Indo-west Pacific sponges through the taxonomic pipeline: The Sponge Barcoding and SpongeMaps Projects

Hooper, J. N. A.^{1,2}, K. A. Hall¹, M. Ekins¹, S. Vargas³, D. Erpenbeck³, G.
Wörheide^{3,4} & G. Jolley-Rogers⁵

Natural Environments Program, Queensland Museum, South Brisbane 4101,
Australia¹

Eskitis Institute for Cell and Molecular Therapies, Griffith University, Mt Gravatt
Research Park, Nathan 4111, Australia²

Department of Earth and Environmental Sciences and GeoBio-CenterLMU, Ludwig-
Maximilians Universität München, Richard-Wagner-Strasse 10, 80333 München,
Germany³

Bayerische Staatssammlung für Paläontologie und Geologie, Richard-Wagner-Strasse
10, 80333 München, Germany⁴

CSIRO Plant Industry, Canberra 2601, Australia⁵

Phylum Porifera includes ~8,500 valid (named) species world-wide, distributed from ephemeral freshwater to deep-sea habitats, but this number represents potentially <30% of their actual diversity. Collections of these species from the Indo-west Pacific have escalated substantially over the past two decades due to pharmaceutical discovery and national bioregional planning (amongst others), such that for many bioregions they represent the most significant body of currently available knowledge. Yet most of these collections remain unidentified, or at best only partially processed (operational taxonomic units, or OTUs), and they now vastly outweigh the expertise available to resolve them via more complete taxonomy. Complicating this is the appalling problems associated with the traditional taxonomy of sponges (such as abundant phenotypic convergences, secondary losses, etc), with independent datasets becoming increasingly important. Recent molecular and chemical analyses continue to discover growing numbers of cryptic species, previously undetected morphologically.

In this way the Sponge Barcoding Project (www.spongebarcoding.org) – the first initiative to genetically barcode a non-bilaterian metazoan phylum – continues to build a comprehensive DNA database to facilitate the identification of unknown or partially identified specimens. DNA from approximately 20,000 museum specimens (predominantly from the western Pacific) is being extracted and amplified for the standard COI barcoding fragment, so far with ~25% mean amplification success rate. Notwithstanding some sponge-specific problems (such as frequent coamplification of non-target organisms), this methodology has already contributed considerably to the refinement of sponge systematics, evaluation of morphometric character importance, geographic phenotypic variability, and has demonstrated the utility of the standard barcoding fragment for Porifera (despite its conserved evolution within this basal metazoan phylum).

To provide biological (morphometric) context to these molecular barcodes, particularly those from OTUs still unresolved within the Linnaean classification, we recently created SpongeMaps (www.spongemaps.org) as a means to ensure there is a biodiversity informatics legacy of knowledge and expertise gained from one project for future projects. SpongeMaps integrates these diverse data (GIS specimen data;

species and OTUs; images; molecular, chemical, and other datasets) through an online, two-way iterative process, behind password protection. It provides an ability to interrogate existing data to better process new collections; a capacity to aggregate or differentiate populations across the Indo-west Pacific by creating new OTUs; automatic publication of online pages for individual species that contextualize GIS specimen data delivered by the online biodiversity databases (such as the *Encyclopedia of Life* and the *Atlas of Living Australia*); and to link external datasets for taxonomic hierarchy, specimen GIS and mapping, DNA sequence data, chemical structures, and images. This talk focuses mainly on the digitization and delivery of these sponge data using SpongeMaps.