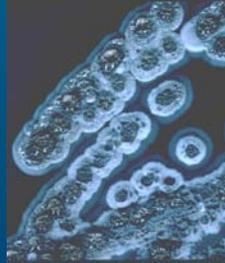


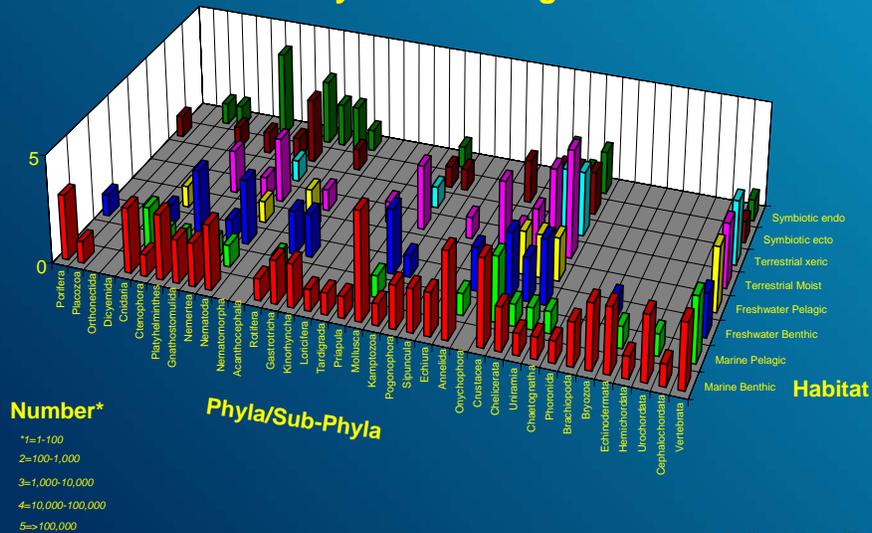
Towards a practical knowledgebase for marine genetic resources



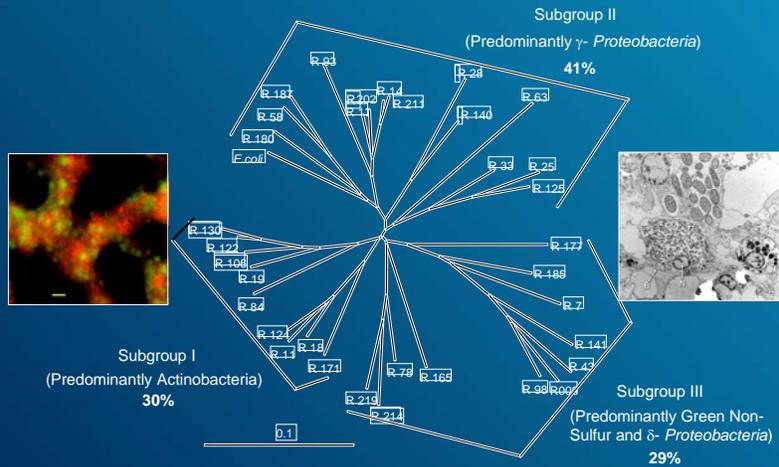
Libby Evans-Illidge

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

The seafloor is the most biodiverse place on earth - Based on diversity of macro-organisms



Superimposed with bacterial symbiont diversity



eg. Microorganisms found in the stable community within the Great Barrier Reef sponge *Rhopaloeides odorabile*. (Nicole Webster)



Marine Natural Products and Related Compounds in Clinical and Advanced

Table 1. Status of Marine-Derived Natural Products in Clinical and Preclinical Trials
name source status (disease) comment

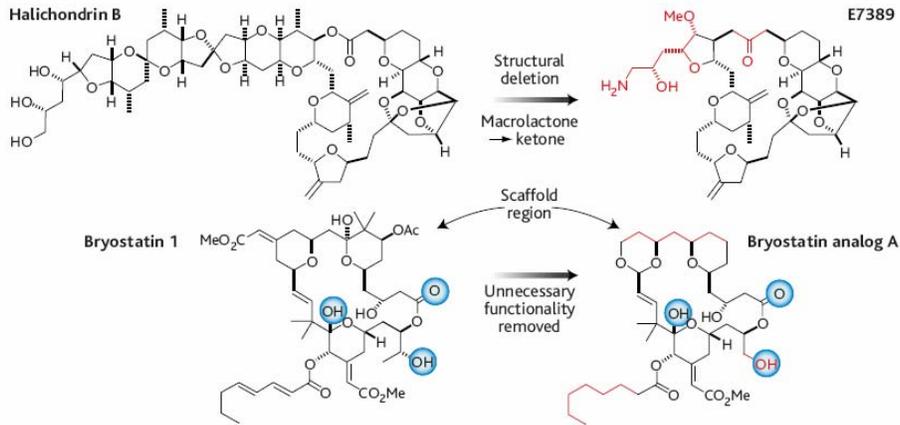
diclermin B	<i>Trididemnum solidum</i>	Phase I/III (cancer)	Phase II (cancer)	dropped middle 90s
dotsestatin 10	<i>Dolabella auricularia</i> (marine microbe derived; cyanophyte)		no further	trials
griroline	<i>Pseudodityssa cantharella</i>		Phase I (cancer)	discontinued
bengamide	<i>Jaspis</i> sp.		Phase I (cancer)	licensed Novartis
cryptophycins	<i>Nostoc</i> sp. & <i>Dysidea arenaria</i>		Phase I (cancer)	licensed to Lilly
bryostatins 1	<i>Bugula neritina</i>		Phase II (cancer)	licensed to GPC Bio
dotsestatin	<i>Dolabella</i>		Phase II (cancer)	melanoma, breast,
ecteinascidin	<i>Ecteinascidia turbinata</i>		Phase III/III (cancer)	licensed to JSJ
aplidine	<i>Apidium albicans</i>	Phase II (cancer)		
E7389 (Hali B)	<i>Lisodendoryx</i> sp.	Phase I (cancer)	Phase I (cancer)	Eisai
discodermolide	<i>Discodermia dissoluta</i>		Phase I (cancer)	licensed to Novartis
halahalide F	<i>Erylia rufescens</i> /Bryopsis sp.		Phase II (cancer)	licensed to PriMar
Spisuloseine	<i>Spisula polyryma</i>	Phase I (cancer)		Rho-GTP inhibitor
HTI-286	<i>Cymbastella</i> sp.		Phase II (cancer)	licensed to Wyeth
KRN-7000	<i>Agelas mauritianus</i>	Phase I (cancer)		
squalamine	<i>Squalus acanthias</i>	Phase II (cancer)		antiangiogenic
Laulimalide	<i>Cacospongia mycofilijensis</i>		preclinical (cancer)	preclinical (cancer)
Curacin A	<i>Lyngbya majuscula</i>		preclinical (cancer)	
viltreavamide	<i>Didemnum cuculliferum</i> & <i>Polysyncraton lithostrotum</i>		preclinical (cancer)	
diazonamide	<i>Diazona angulata</i>		preclinical (cancer)	
eleutherobin	<i>Eleutherobia</i> sp.	preclinical (cancer)	preclinical (cancer)	Just licensed
sarcodicyin	<i>Sarcodicyon rossum</i>		preclinical (cancer)	
peteroside A	<i>Mycale hentschellii</i>			
salicylhalimide	<i>Haliclona</i> sp.		preclinical (cancer)	
thiocoraline	<i>Micromonospora marina</i>		preclinical (cancer)	
variollins	<i>Kirkpatrickia variolosa</i>		preclinical (cancer)	
dicyodendrins	<i>Dicyodendrilla vorongiiformis</i>		preclinical (cancer)	
masoalide	<i>Luffariella variabilis</i>		Phase II	licensed to Taiho
IPL-576,092	<i>Petrosia contignata</i>	Phase II		licensed to Aventis
ziconotide	<i>Conus magus</i>		Phase III pain	licensed to WL
CGX-1160 plus	<i>Conus geographus, catus, victoriae</i>	Phase I (pain)		

Journal of Natural Products Reviews
David J. Newman* and Gordon M. Cragg
Natural Products Branch, Developmental Therapeutics Program, NCI-Frederick, P.O. Box B, Frederick, Maryland 21702
Received February 10, 2004; see also Fortman and Sherman CC 2005
© 38% from Australasia

Cytarabine (Ara-A,C) from *Cryptotheca crypta* now routine for leukaemia and lymphoma patients, Glaxo Smith Kline and Pfizer



Natural Products “Renaissance” (with a twist)



Paterson and Anderson, *Science* 21 October 2005



A significant cumulative effort in ocean exploration



‘Parents’ of marine science:

- Indigenous observations over millennia
- Aristotle 384-322BC
- Charles Darwin HMS *Beagle* 1831-36
- Challenger 1872-75

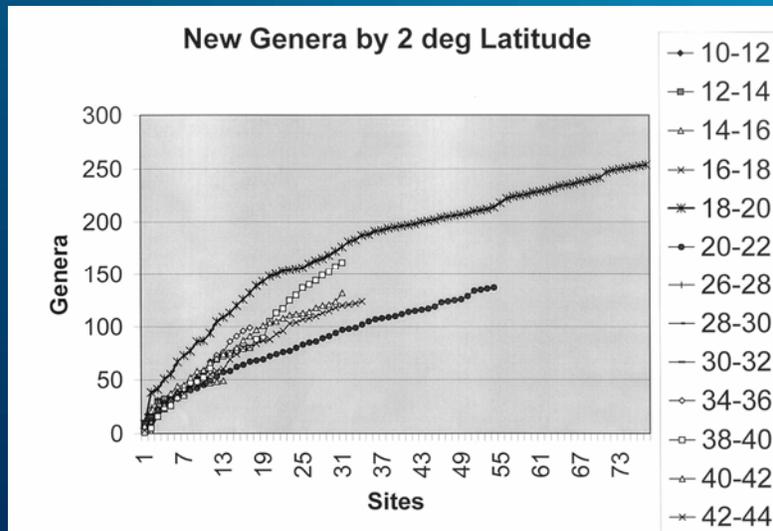


Marine R&D technology in the 21st Century



Australian Government AUSTRALIAN INSTITUTE OF MARINE SCIENCE

...but still so much remains unknown



Australian Government AUSTRALIAN INSTITUTE OF MARINE SCIENCE

Where is the data? How can we access it?

- Specialist data
- Published Literature
- Portals for metadata
- Networked datasets
- Integrated datasets

The published literature

Literature Databases

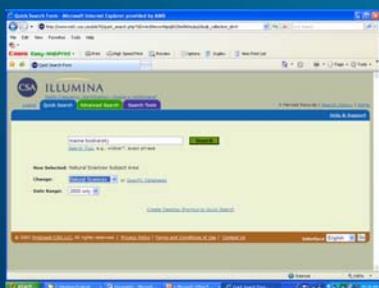
- Cambridge Scientific Abstracts
- Aquatic Sciences & Fisheries Abstracts
- Zoological Record
- Biosis
- MarineLit
- Patent databases

Citation & Indexing services

- Web of Science
- Google Scholar

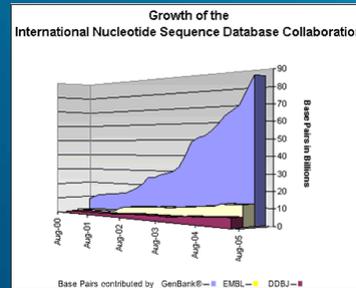
Citation data management

- Ref works
- Endnote



Data Networks/Repositories Nucleotide sequences

- **INSDC** (international Nucleotide Sequence Database Collaboration)
www.insdc.org
 - **GenBank[®], DDBJ, EMBL**
 - Annotated collection of all publicly available DNA sequences
 - Submission of sequences required by many journals prior to publication
 - Online submission, update and review
 - Country of origin identified
 - No restriction on use or distribution
 - 73078143 loci, 77248690945 bases, from 73078143 reported sequences (June 15 2007)
- **Data searching & analysis tools**
 - eg. BLAST



www.ncbi.nlm.nih.gov/Genbank/



Specialist data – find the people

Location of specimen and data holdings reflects historical location of the specialists



- eg. Scleractinian corals
- Major repository at Museum of Tropical Queensland, Townsville Australia
- Doorstep of Great Barrier Reef
- Home of 2 global experts - Dr Carden Wallace (MTQ) and Dr John Veron (AIMS, Townsville)



UN Atlas of the Oceans



<http://www.oceansatlas.org/>

- UN-Oceans coordination portal
- 14 Global partners and 8000 individual members.
- Peer reviewed.
- Features a Virtual Office, advanced Search, web statistics and monthly newsletter. FREE TO MEMBERS
- New marine genetics page



National data network example

atlas of LIVING AUSTRALIA

Name

- APNI
- Derivation

Description

Classification

- Barcode of Life
- Cladogram
- Close relatives

Illustrations

- Photographs
- Line Drawings

Identification tools

- Australian Tropical Rain Forest Plants

Usage

Horticulture

Ethnobotany

Pharmacology

Current Research

Bibliography

Castanospermum australe

Belongs to the Fabaceae family, commonly referred to as **Black Bean** or the **Moreton Bay Chestnut**.

Beautiful large evergreen tree with glossy dark green pinnate leaves and low spreading branches when grown in the open. During Spring it bears sprays of red and yellow pea-shaped flowers. The large cylindrical pods which are produced in autumn and have 3-5 large bean like seeds inside. The nectar produced by the flowers attracts birds, bats and butterflies.

Tree, to 35 m high with trunk up to 1.5 m diam., but usually smaller, mostly glabrous. Leaves 30-60 cm long; leaflets 9-17, a oblong-elliptic, 7-20 cm long, 3-5 cm wide, apex acuminate, margins entire, glabrous, upper surface glossy, lower surface paler and dull; petiole 3-6 cm long; lateral petioles 2-5 mm long. Racemes mostly 5-15 cm long; pedicels 20-25 mm long. Calyx c. 10 mm long. Corolla 30-40 mm long, orange to red. Pod mostly 10-20 cm long, 4-6 cm diam., glabrous; seeds 1-5, c. 30 mm diam., brown. Flowers spring.

Ecology

Grows in well developed rain forest, but is often found in gallery forest along creeks and rivers.

Source: Flora of NSW and Rain Forest Key

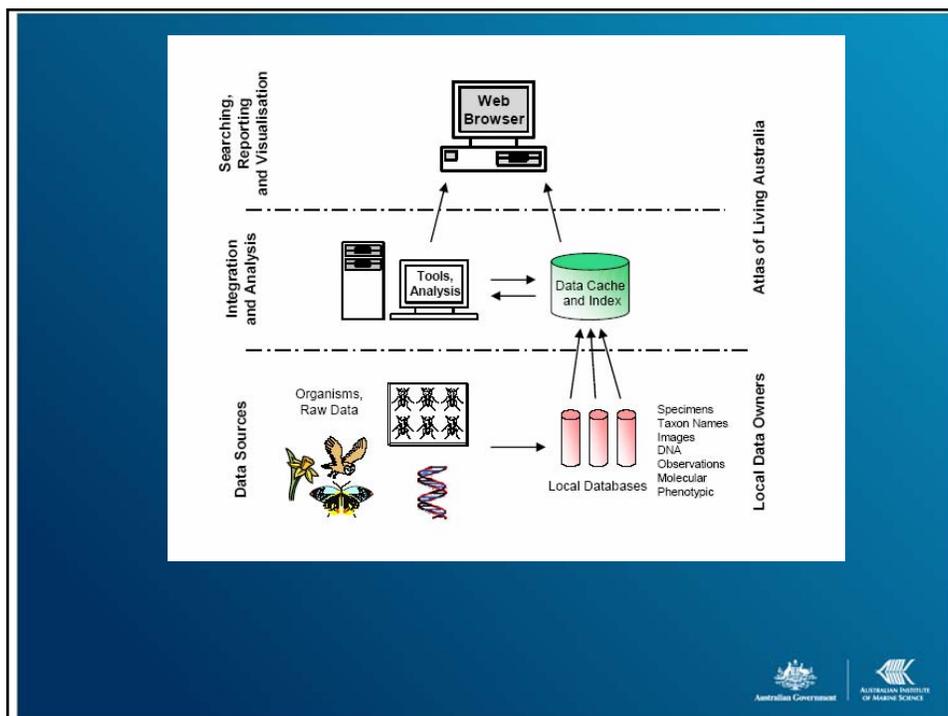
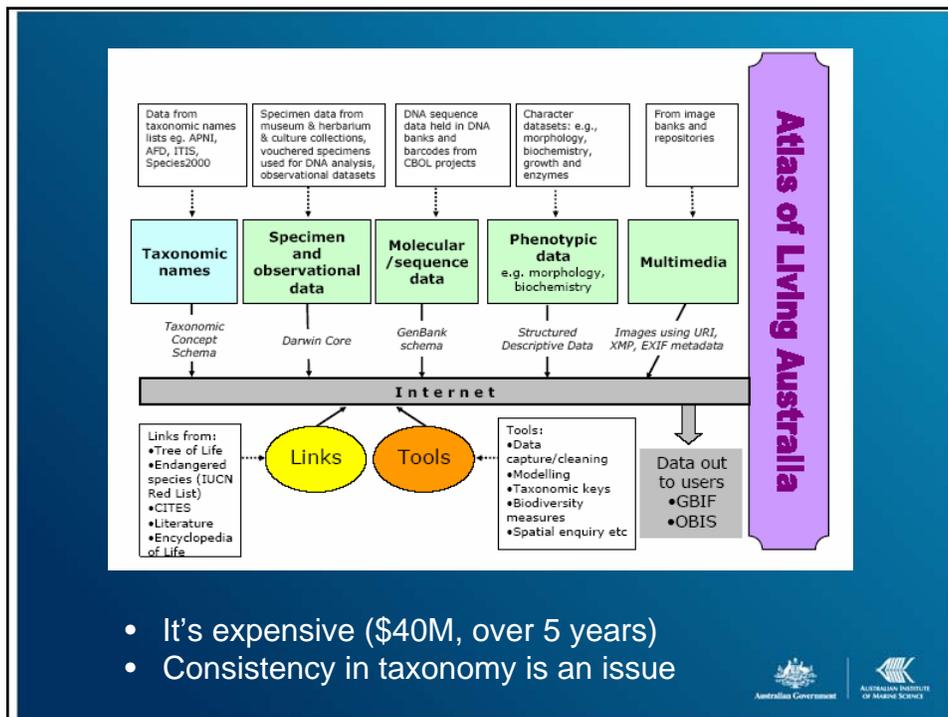



Distribution

Occurs in Cape York Peninsula (CYP) and North East Queensland (NEQ), and southwards to north eastern New South Wales. Altitudinal range in CYP and NEW from near sea level to 600m. Also occurs in New Caledonia and Vanuatu.

- Searchable database, common format
- all known Australian plant & animal species by 2015
- terrestrial and marine





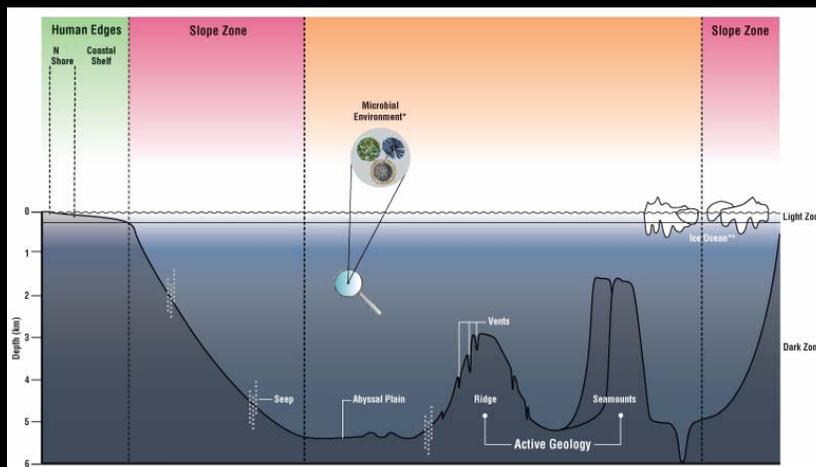


- Growing global network
- Assess and explain the diversity, distribution and abundance of marine life in the world's oceans
- past, present and future
- 2000 to 2010
- 50+ countries, 300+ scientists, 17 major projects
- Link to the Barcode of Life initiative



Oceans Present: Realm Projects

CoML defines its realms & zones in 2003 Baseline Report, *The Unknown Ocean*



* Microbial environment encompasses the entire world ocean.
 ** Ice oceans occur at both poles.

Theoretical Cross Section of the Ocean

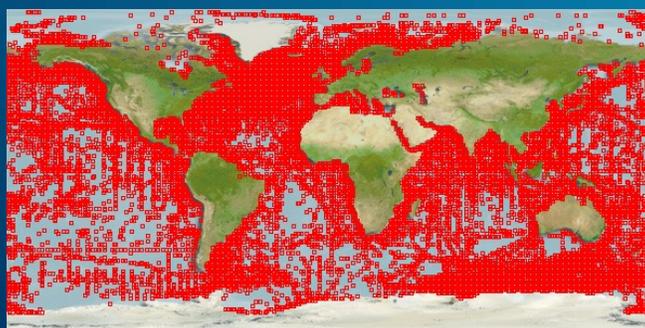
Ocean Realm Field Projects

Human Edges	Nearshore: NaGISA Reefs: CReefs Regional Ecosystem: GoMA Continental Shelf: POST
Hidden Boundaries	Continental Margins: COMARGE Abyssal Plains: CeDAMar
Central Waters	Zooplankton: CMarZ Top Predators: TOPP Ridges: MAR-ECO
Active Geology	Vents & Seeps: ChEss Seamounts: CenSeam
Ice Oceans	Arctic: ArcOD Antarctic: CAML
Microscopic Ocean	Microbes: ICOMM



Ocean Biogeographic Information System (OBIS) Portal www.iobis.org

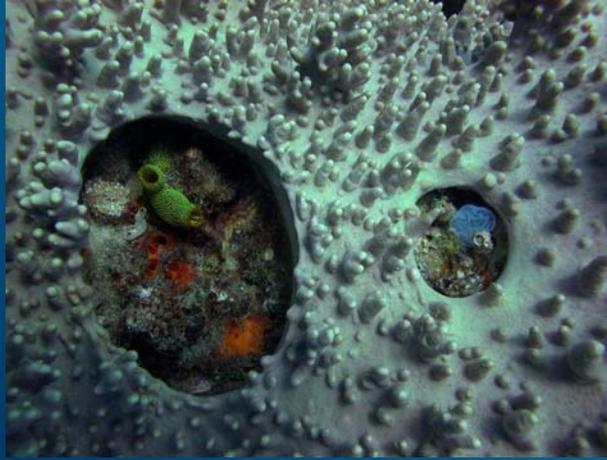
OBIS strives to document the oceans' diversity, distribution, and abundance of life.



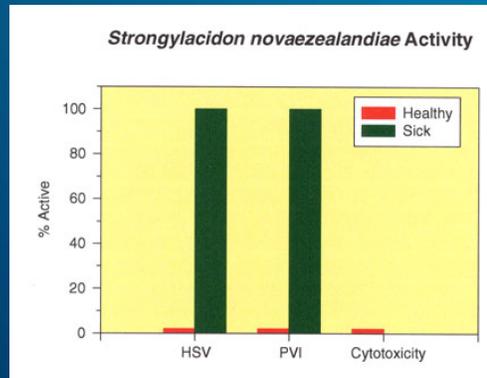
13.2 million records, 81000 species, 210 databases
Searchable on name or geography
Consistency in taxonomic assignment is an issue



Assessment of Marine Genetic Resources - You need to know more than 'where they occur'



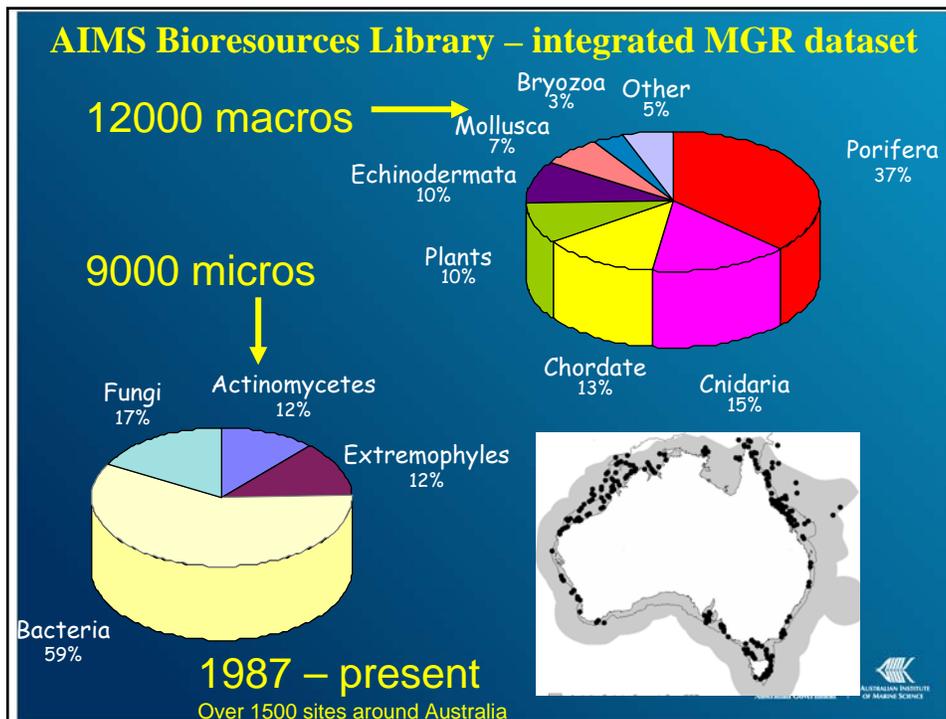
Marine genetic resources are the integrated result of dynamic, complex ecology.



eg.

- Activity of interest (HSV and PVI) in this example is only produced by sick sponges
- Note high selectivity to the target (and absence of cytotoxicity)





COLLECTION DATA

- Precise location
- Depth
- Size
- Best taxonomy (voucher back-up)
- Imagery
- Coded descriptions
 - habitat
 - morphology
 - substrate
 - colour(external/internal)
 - epibionts
 - symbionts
- Descriptive notes and observations
- Site information (location, habitat description, community composition)

Micros:

- Details of source and selective isolation

Australian Government | Australian Institute of Marine Science

Bioinformatics – begins with sample info

The screenshot displays a web-based database interface with several overlapping windows:

- Frozen Bulk Details:** A table with columns: ID, Date, Location, Collected, and Comments. It shows one entry with ID 1000001.
- History of Frozen Bulk:** A table with columns: ID, Name, Date, Date In, Status, and Add Date. It shows one entry with ID 1000001.
- Macro Sample Details:** A detailed form with fields for Name, Date, Time, Location, Depth, Station, and various other attributes.
- Voucher Details:** A table with columns: ID, Voucher, and Date. It shows one entry with ID 1000001.
- Macro Sample List:** A table with columns: ID, Name, Date, Date In, Status, and Add Date. It shows one entry with ID 1000001.
- Macro Sample List:** A table with columns: ID, Name, Date, Date In, Status, and Add Date. It shows one entry with ID 1000001.
- Macro Sample List:** A table with columns: ID, Name, Date, Date In, Status, and Add Date. It shows one entry with ID 1000001.

Two photographs of coral are included: one showing a close-up of a branching coral colony, and another showing a larger, more complex coral structure in an underwater environment.

The screenshot displays a web-based database interface with three overlapping windows showing taxonomic details:

- Eubacteria Details:** A table with columns: DATE, MEDIA, COL_MEDI, PURP_MEDI, TEMP, INOC_TIME, CULTURE, OPTICAL, SIZE, STRAINS, and COMMENTS. It shows one entry with DATE 2000-01-01.
- Fungi Details:** A table with columns: DATE, MEDIA, COL_MEDI, PURP_MEDI, TEMP, INOC_TIME, CULTURE, OPTICAL, SIZE, STRAINS, and COMMENTS. It shows one entry with DATE 2000-01-01.
- Actino Details:** A table with columns: DATE, MEDIA, COL_MEDI, PURP_MEDI, TEMP, INOC_TIME, CULTURE, OPTICAL, SIZE, STRAINS, and COMMENTS. It shows one entry with DATE 2000-01-01.

The Australian Government and Australian Institute of Marine Science logos are visible in the bottom right corner.

...integrated with other research outputs

Create report | IC50 curve | Go back

MNUM	CA % control SE	CA	NO5 % control SE	NO5	PHYLUM	CLASS	ORDER	FAMILY
120	7	115	9	107	PORIFERA	DEMOSPONGIAE	RETROSDIA	RETROSDIAE
50	115	1	107	6	PORIFERA	DEMOSPONGIAE	DICTYOCERATIDA	SPONGIDAE
52	92	5	91	6	PORIFERA	DEMOSPONGIAE	DICTYOCERATIDA	DYSZEDAE
51	98	10	103	3	PORIFERA	DEMOSPONGIAE	DICTYOCERATIDA	DYSZEDAE
52	6	11	114	4	PORIFERA	DEMOSPONGIAE	HADROMERIDA	SPRASTRILLIDAE
44	100	13	115	7	PORIFERA	DEMOSPONGIAE	POECLOSCLERIDA	MICROCONDIAE
45	76	8	88	7	PORIFERA	DEMOSPONGIAE	POECLOSCLERIDA	MICROCONDIAE
47	111	13	117	4	PORIFERA	DEMOSPONGIAE	POECLOSCLERIDA	MYCALEAE
43	95	17	95	4	PORIFERA	DEMOSPONGIAE		
43	108	2	106	6	PORIFERA	DEMOSPONGIAE		
55	112	15	110	7	PORIFERA	DEMOSPONGIAE		

Pharmacology Search

Calcium Channel:

r/Nos Assay:

MNUM:

Phylum:

Class:

Order:

Family:

Genus:

Species:

Local name:

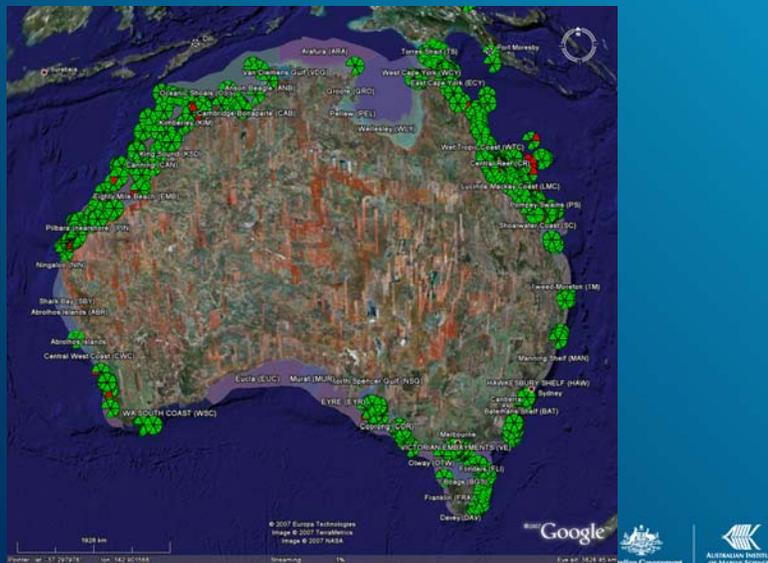
IC50 CURVES MNUM: 17605

CALCIUM

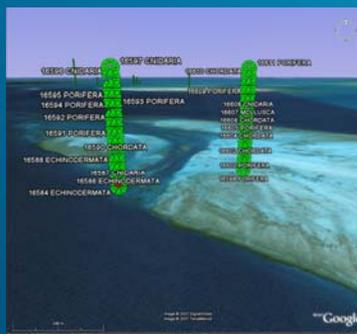
Australian Government | AUSTRALIAN INSTITUTE OF MARINE SCIENCE

Integrated tool for data mining.

eg. Regional and taxonomic patterns in anti-microbial activity



Analysis at a range of scales



Google Earth interface facilitates good visualisation of multi-variate data

Understanding the chemical ecology Apply data-mining to enhance biodiscovery

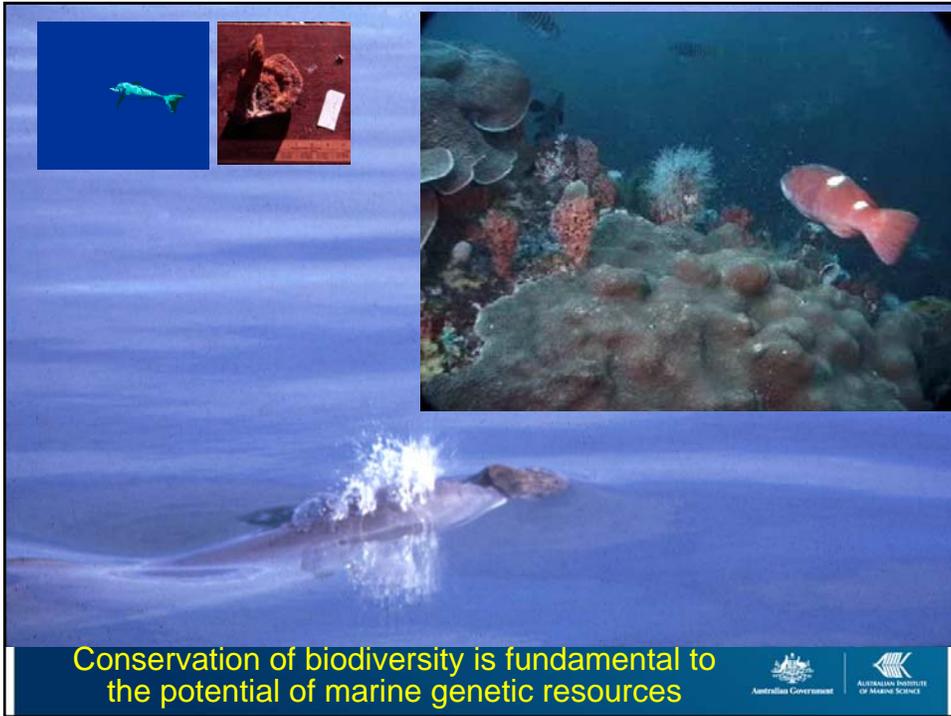
- Elaborate leads
 - identifying other material with similar taxonomy/ecology/ screening profile/chemistry
 - Naturally occurring analogues
 - Re-supply without re-collecting
- Predict results of future screening based on past profile
 - Compile list of pre-leads
 - Targeted biodiscovery with ex-situ material



Summary

- Tools exist to access marine biodiversity and genetic resources data that is in the global public domain
- Major networking projects underway to bring together independent geo-referenced datasets (CoML, ALA)
- Consistency in taxonomy is a big issue in networking independent datasets
- Integrated informatics is the ideal
 - Integrate biodiversity data (ecosystems, species (macro and micro), genomes (and meta-genomes)), with natural products research outputs (instrument outputs, structures of compounds, proteins, enzymes etc), and screening results
 - Powerful datamining tool to assess and understand marine genetic resources
 - 'Commercial in Confidence' issues, but even a compilation of non-commercially sensitive information is useful





Conservation of biodiversity is fundamental to the potential of marine genetic resources

