

Discussion Paper

Making *GeneBank Australasia* a Reality



Prof Andrew Lowe

Introduction

GeneBank Australasia is a concept that was developed at a workshop in December 2011, held as part of the International Australian Barcode of Life Conference. Thirty research leaders from 13 countries participated in the workshop and strongly supported a coordinated Australian initiative in DNA-barcoding and supporting institutional infrastructure. Since the workshop, Professor Andrew Lowe from the University of Adelaide and Professor Suzanne Miller from the South Australian Museum have had further discussions with research leaders from around Australia, which has reinforced the support for such a concept.

The vision for GeneBank Australasia is that it will be possible to cheaply and accurately identify the species of plants, animals and microbes so that we can:

- Make fast, accurate assessments of biodiversity impacts and find opportunities to manage and mitigate the impact of human development cheaply in the most practical and efficient ways possible.
- Better and more cheaply understand the threats to Australian agriculture, animals, fisheries and environment from exotic species.
- Conserve our native plants and animals more effectively and economically.

To realise these competitive advantages, Australia needs a deliberate strategy and coordinated investment.

We need to coordinate and fast-track the 'bar coding' of life forms; with a focus on economically important diseases in the region, the provenance of targeted food and forestry products, and Australasian plants and animals; and enhance national capabilities offered by the biodiversity collection institutes (museums and herbaria) and infrastructure programs, such as BioPlatforms Australia/Australian Genome Research Facility, the Atlas for Living Australia and the Terrestrial Ecosystem Research network.

Program Design

Achieving the objectives of GeneBank Australasia will require a number of steps listed in Table 1. Some of the activities require enabling science before they can be fully implemented.

Table 1. Steps required to establish the functionality of Genebank Australasia

Activity	Enabling Science
A1. Digitisation of existing collections information by collection institutes (museums and herbaria) and Atlas of Living Australia [ALA] - a priority in the DIIRSTE infrastructure roadmap	E1. New image capture and analysis technologies will be required for rapid species identification and information scanning (ALA and Australian Centre for Visual Technologies)
A2. Semi-automated field and institute data capture	E1. New image capture and analysis technologies will be required for rapid species identification and information scanning (ALA and Australian Centre for Visual Technologies) E2. Full implementation of AEKOS ecological field plot capture and information storage system (TERN)
A3. Establish tissue and DNA banks and undertake whole genome analysis of key species, sub-sampling representative material for molecular analysis	E3. Improved, rapid ID tools and establish biodiversity genomics centres in conjunction with AGRF
A4. Barcode a subset of representative material on a priority basis to provide an identification framework for multiple end-users	E4. Up-scaling DNA barcoding technology from scientific to commercial scale and cost with AGRF/BPA
A5. Construction of databases for DNA barcoding outputs	E5. Integrate systems with ALA, Barcode of Life Database (BOLD) and AEKOS

Steps A1 – A3 will require the coordination of biological data collection and inputs by the collection institutes, government agencies (including the National Reserve System), ALA and industries. Enabling Science Task E1 for rapid visual species identification and information scanning will be a task for ALA and

Australian Centre for Visual Technologies. Enabling Science Task E2 will build on existing investments in the Terrestrial Ecosystem Research Network (TERN) and Australian Ecological Knowledge & Observation System (AEKOS). Enabling Science Tasks E3 and E4 will establish national centres for biodiversity genomics and DNA barcoding, integrated with existing investments made by AGRF and BPA. Enabling Science Task E5 will establish national databases for DNA barcoding in conjunction with capabilities provided by the ALA, Barcode of Life Database (Canada) and the Terrestrial Ecosystem Research Network (TERN) and Australian Ecological Knowledge & Observation System (AEKOS)

GeneBank Australasia will create a DNA Barcoding Pipeline for Australia and neighbouring countries (e.g. Papua New Guinea). Essential to this strategy are specimen collections which have been undertaken systematically by the Collections Institutes (Herbaria, Museums, Seed Banks) for most of Australia's European History. These collections are central to the design of GeneBank, because they contain the best set of validated voucher specimens of Australasia's flora and fauna. The **Atlas of Living Australia (ALA)**, aggregates information on known species in Australasia from the Collections Institutes, as well as community groups, government departments, individuals and universities. Extra funds are required to provide ALA with a capability to capture high resolution images and information (labels, using Optical Character Recognition) associated with specimens (Task A1).

Field-based information of species presence is being undertaken nationally through the **Multi-Scale Plot Network (MSPN)** of the **Terrestrial Ecosystem Research Network (TERN)**, which is then stored electronically in the **Australian Ecological Knowledge and Observation System (AEKOS)**. **Bush Blitz** is a nationally coordinated program to document the plants and animals in hundreds of properties across Australia's National Reserve System. These programs, together with ongoing activities within the Collections Institutes and research agencies, will create a vast stream of data and specimens for analysis (Task A2).

Tasks A3 will require the establishment of tissue and DNA banks in the Collections Institutes. Enabling Science E3 will require additional investment in improved, rapid identification and biodiversity genomics. A consortium of research institutions will be required to undertake this task.

A major gap in the GeneBank Australasia infrastructure at the moment is a rapid system of tissue sampling, DNA processing and sequencing (Figure 1). For Task A4, the **Australian Genome Research Facility** and **BioPlatforms Australia** will be important partners in GeneBank Australasia, for providing expertise and equipment required to establish a rapid DNA processing facility. BPA and AGRF will work with research institutions to implement Task E4, up-scaling DNA barcoding technology from scientific to commercial scale and cost.

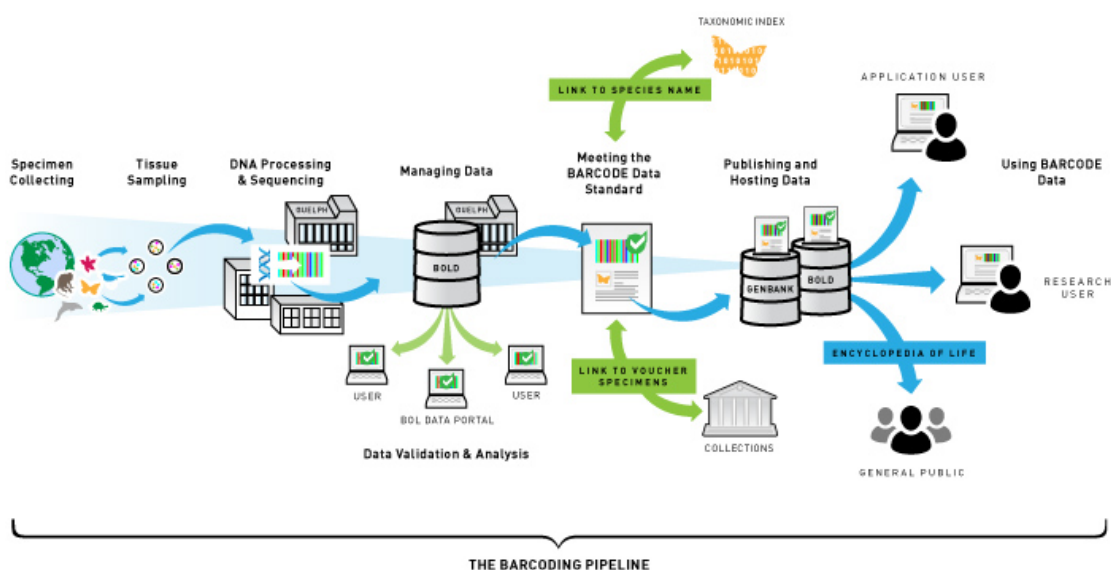


Figure 1. A Rapid DNA Barcoding System

The information generated by GeneBank Australasia will be stored by institutional databases and served up through the ALA, BOLD and AEKOS (Task A5), all of which will require upgrading to meet the new data requirements.

For each species, the following information will be stored in a standard data table:

- Species name (linked to ABRS taxonomic indices)
- Code and location of voucher specimen(s)
- DNA code and standards adhered to
- High resolution image of specimen and taxonomic features

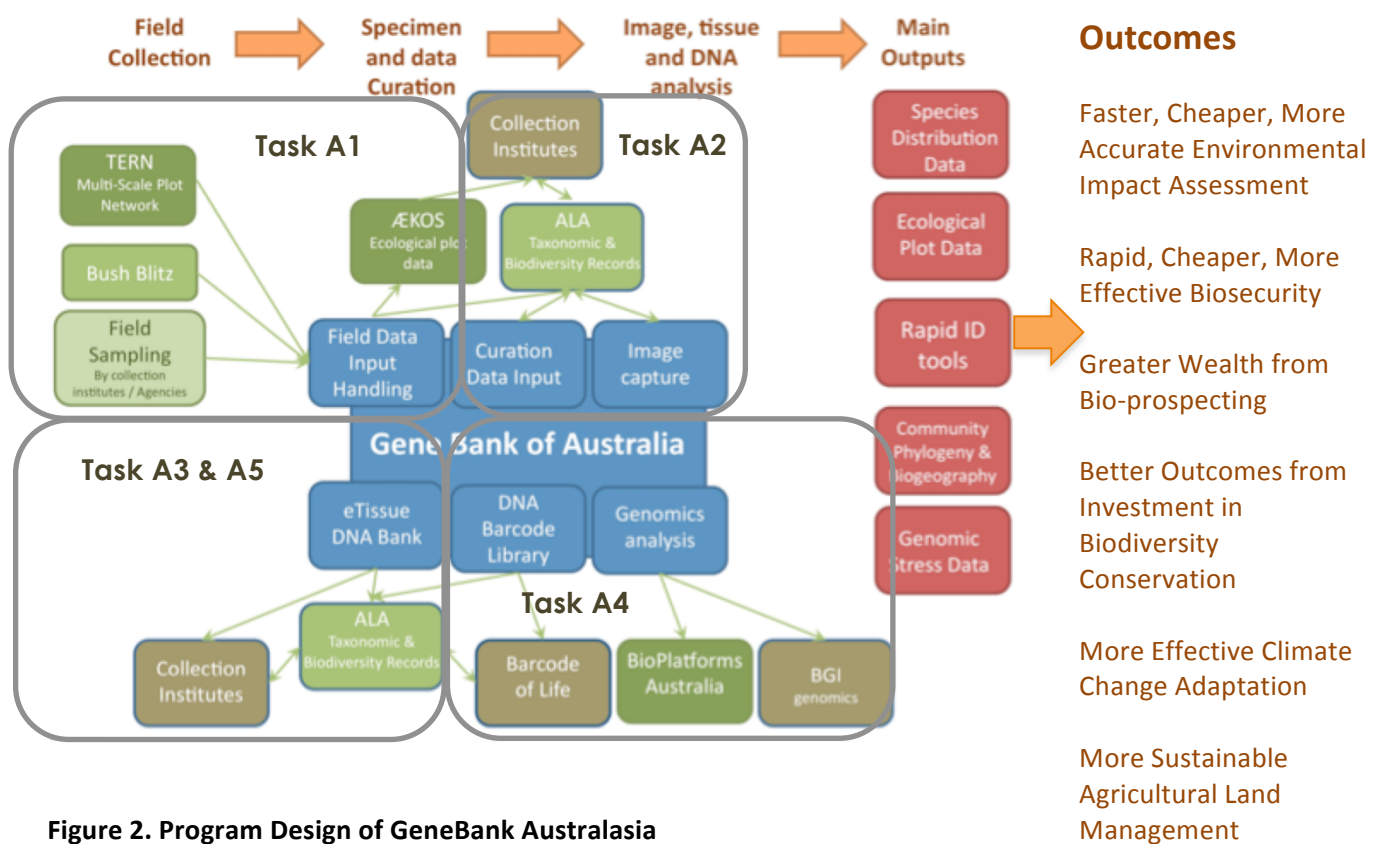


Figure 2. Program Design of GeneBank Australasia

Coordination and Governance

GeneBank Australasia requires coordination of existing capability and the deployment of new infrastructure, activities and expertise.

A governance model is proposed in Figure 3.

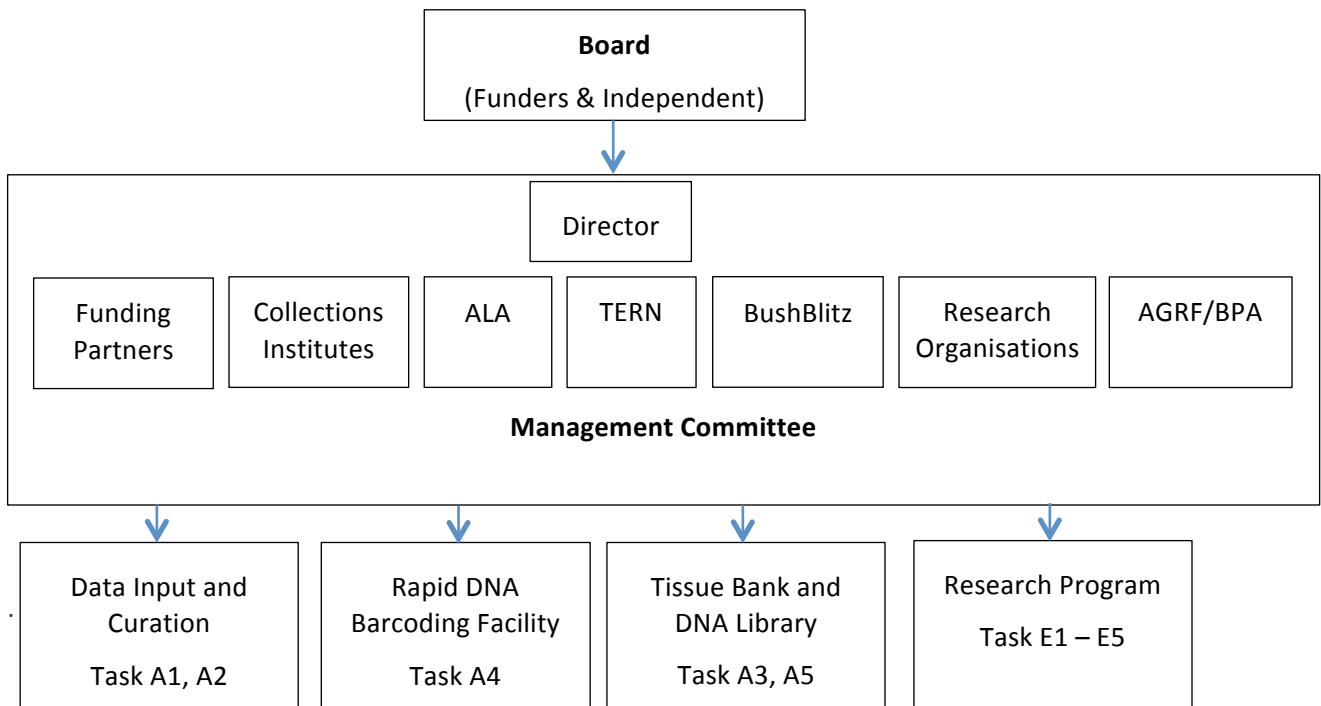


Figure 3. Management Structure of GeneBank Australasia

The GeneBank Australasia Board will have an independent Chair and will be composed of independent members and members from the funding partners. The Board will be responsible for funding decisions, appointment of the Director and strategic direction of GeneBank Australia.

A Management Committee will include representation from the partners, and will be responsible for the coordination of activities and works programs across the partner agencies. The Director will be a member of the Committee and will be ultimately responsible to the Board for the implementation of the strategy.

The Management Committee will coordinate four major works programs (Figure 3):

- Data Input and Curation (Tasks A1 and A2): National coordination of field and specimen data entry into ALA database (?); through Collections Institute, TERN, Bush Blitz and Research Organisations.
- Creation of a National, Rapid DNA Barcoding Facility (Task A4); in partnership with BioPlatforms Australia and the Collections Institutes and Research Organisations.
- Creation of a National Tissue Bank and DNA Library (Tasks A3 and A5); in partnership with BioPlatforms Australia and the Collections Institutes and Research Organisations
- Research Program (Tasks E1 – E5); with all partners

The Delivery Partners will be responsible for the delivery of GeneBank Australia activities. They will sign a Partnership Agreement to become members of GeneBank Australia, which will describe:

- Institutional form of GeneBank Australasia (suggest limited liability company)
- Governance
- Cost recovery formulae
- Communications, branding and marketing of GeneBank Australasia

- Rules of membership and participation
- Roles and responsibilities of each of the partners
- Reporting requirements for GeneBank Australia activities
- Management of intellectual property

Indicative Budget

The following budget is proposed over five years:

Item	Program	Delivery Timeframe	Delivery Partners	Estimated Cost (7yrs)
Governance		2014 - 2020	All	\$500,000
Director and Management		2014 - 2020		\$2,500,000
Image capture and Curation Data Input (A1)	Data Input & Curation	2014 - 2020	ALA, Collections Institutes, TERN, Bush Blitz, Research Organisations	\$20,000,000
Coordination and Extension of Field Input Data (A2)	Data Input & Curation	2014 - 2020	ALA, Collections Institutes, TERN, Bush Blitz, Research Organisations	\$10,000,000
eTissue DNA Bank (A3)	Tissue Bank and DNA Library	2016 - 2020	BioPlatforms Australia, Research Organisations, Collections Institutes	\$15,000,000
Genomic Analysis (A4)	Rapid DNA Barcoding Facility	2015 - 2020	BioPlatforms Australia, Research Organisations, Collections Institutes	\$25,000,000
DNA Barcode Library (A5)	Tissue Bank and DNA Library	2016 - 2020	BioPlatforms Australia, Research Organisations	\$5,000,000
Research Program (E1 – E5)	Research program	2014 - 2020	All partners	\$15,000,000
TOTAL				\$93,000,000