We are Going to Barcode a Planet
Planning the International Barcode of Life Project – June 2007
Imagine... a world in which you can identify any:

- animal
- plant
- food
- fungus
- pest
- organism

This is the world that DNA barcoding will build

Making every species count
http://www.ibol.org

5 Years
500,000 Species
$100 million
Mobilize the iBOL Research Community Europe
Mobilize the iBOL Research Community China
Mobilize the iBOL Research Community India
Launch of iBOL

Guelph/Toronto, Canada - Sept 24-26, 2010

Nagoya, Japan - Oct 20, 2010
DNA Barcode Coverage – 2010

- 1 - 10
- >10
- >100
- >1000
iBOL Phase I – Target Achieved

Species (K)

iBOL-1 delivered coverage for about 5% of multicellular species
A Mission for Planetary Biodiversity
A Mission for Planetary Biodiversity

Census of All Species
Understand Biodiversity

Library of Life
Read Genomes

Global Biosurveillance
Protect Biodiversity

Census of All Species
Library of Life
Global Biosurveillance
Key Uncertainty: How Many Species Await Analysis?

10 Million
20 Million
or more?
Key Challenge for PBM: Projected Cost

$1 BILLION
A Key Discussion in Kruger
1. Could we achieve important goals?

2. Could we raise the funds?

3. Were you still keen to proceed?
MEMORANDUM OF UNDERSTANDING

I. Parties

WHEREAS the Australian Museum ("AM") is an Australian organization located at

6 College Street,
Sydney, NSW 2010,
Australia; and

WHEREAS AM seeks to inspire the exploration of nature and cultures, and is dedicated to
the promotion of science and technology in broader service to understanding natural history
in Australia; and

WHEREAS the Department of Primary Industries for and on behalf of the State of New
South Wales ("NSWDPI") is an Australian organization located at

161 Kite Street
Orange, NSW 2800,
Australia; and

WHEREAS NSWDPI is dedicated to fostering profitable and sustainable development of
primary industries in New South Wales, and underpinning their growth and biosecurity; and

WHEREAS the Botanic Gardens Trust ("BGT") is an Australian organization located at

Mrs Macquaries Rd,
Sydney, NSW 2000,
Australia; and

WHEREAS BGT seeks to inspire the appreciation and conservation of plants and their
critical role in the sustainability of our natural and urban environments; and
Our Moonshot – A Planetary Biodiversity Mission

iBOL-I
0.5M Species
Bio-surveillance

iBOL-II
1.5 Million Species

20 Million Species
Global Bio-surveillance
Our Moonshot – A Planetary Biodiversity Mission

iBOL-I
0.5M Species

iBOL-II
2.5 Million Species

Planetary Biodiversity Mission

20 Million Species
Global Bio-surveillance
Number of Species Barcoded

![Graph showing the number of species barcoded from 2004 to 2018. The graph indicates a steady increase in the number of barcodes and taxonomic units over the years.]
Plants – 0.5 million

Fungi – 5 million?

Animals – 15 million

20 Million Multicellular Species?
A Key Barrier for iBOL-II

How can we barcode more species than have been described?
Breaking a Key Barrier for iBOL-II

Develop a new method to register species
Automating the Registration of Animal Species

1. Mammals
2. Moths
3. Fishes
4. Butterflies
5. Aphids
6. Ants
7. Bees
8. Birds
9. Spiders
10. Beetles
Automating the Registration of Animal Species

Counting Sequence Clusters

1 2 3 4 5 6 7 8 9
Automating the Registration of Animal Species

Barcode Index Numbers

Names for Clusters:

- BOLD: AAF3452
- BOLD: AAC9004
- BOLD: AAB4123
- BOLD: AAA3260
- BOLD: AAA7781
- BOLD: AAF3512
- BOLD: AAF3297
- BOLD: AAF3129
- BOLD: AAE3351
Automating the Registration of Animal Species

- Go anywhere
- Collect specimens
- Gather barcode sequences
- Assign to Barcode Index Numbers

BARCODE OF LIFE DATA SYSTEMS
A BIOINFORMATICS PLATFORM FOR DNA BARCDOING OF ANIMAL, PLANT, AND FUNGAL SPECIES.
1 Barcode Index Number of 560,000
## Linnaean Species versus BINs

<table>
<thead>
<tr>
<th></th>
<th>Linnaean</th>
<th>BIN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Names</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Higher Taxonomy</td>
<td>Linnaean</td>
<td>Linnaean</td>
</tr>
<tr>
<td>Genus/species</td>
<td>Binomen</td>
<td>Alphanumeric</td>
</tr>
<tr>
<td><strong>Type Material</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holotype</td>
<td>Fixed</td>
<td></td>
</tr>
<tr>
<td>Paratype</td>
<td>Fixed</td>
<td></td>
</tr>
<tr>
<td>Other Specimens</td>
<td>Fixed</td>
<td>Dynamic</td>
</tr>
<tr>
<td><strong>Specimen Collaterals</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collection Dates</td>
<td>Fixed</td>
<td>Dynamic</td>
</tr>
<tr>
<td>Collection Locales</td>
<td>Fixed</td>
<td>Dynamic</td>
</tr>
<tr>
<td>Collector</td>
<td>Fixed</td>
<td>Dynamic</td>
</tr>
<tr>
<td>Image</td>
<td>Fixed</td>
<td></td>
</tr>
<tr>
<td><strong>Diagnostic Traits</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morphology</td>
<td>Fixed</td>
<td></td>
</tr>
<tr>
<td>Molecular</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Synonomy</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Linnaean Species versus GenoSpecies

<table>
<thead>
<tr>
<th>Names</th>
<th>Linnaean</th>
<th>GenoSpecies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher Taxonomy</td>
<td>Linnaean</td>
<td>Linnaean</td>
</tr>
<tr>
<td>Genus/species</td>
<td>Binomen</td>
<td>Alphanumeric</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type Material</th>
<th>Fixed</th>
<th>Dynamic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holotype/GenoType</td>
<td>Fixed</td>
<td>Fixed</td>
</tr>
<tr>
<td>Paratype/ParaGenoTypes</td>
<td>Fixed</td>
<td>Dynamic</td>
</tr>
<tr>
<td>Other Specimens</td>
<td>Fixed</td>
<td>Dynamic</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Specimen Collaterals</th>
<th>Fixed</th>
<th>Dynamic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collection Dates</td>
<td>Fixed</td>
<td>Dynamic</td>
</tr>
<tr>
<td>Collection Locales</td>
<td>Fixed</td>
<td>Dynamic</td>
</tr>
<tr>
<td>Collector</td>
<td>Fixed</td>
<td>Dynamic</td>
</tr>
<tr>
<td>Image</td>
<td>Fixed</td>
<td>Fixed</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnostic Traits</th>
<th>Fixed</th>
<th>Dynamic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphology</td>
<td>Fixed</td>
<td></td>
</tr>
<tr>
<td>Molecular</td>
<td></td>
<td>Dynamic</td>
</tr>
</tbody>
</table>

| Synonomy | | |
|----------|| |

How many specimens will need to be analyzed?
Amazon’s Biodiversity Warehouse
Global Insect Survey – 1.8 Million Specimens Analyzed
Global Insect Survey

153,988 BINs

Specimens (M)

BINs (K)
Challenge for iBOL-II: Sequencing Costs

20 million specimens 
\times \$5/\text{specimen} 

\$100 \text{ MILLION}
SEQUEL vs Sanger - Congruent, but Cost is 95% Less
Sequences 10,000 Specimens in 6 Hours
Specimen Barcodes: 5 million /year

Species Registration: 500,000/year
The Pace of Species Registration

- **Global Taxonomic Community (Past 2.6 Centuries):** 6,500 species per year
- **Global Taxonomic Community (Past Decade):** 15,000 species per year
- **One SEQUEL:** 500,000 species per year
Barcode Production

Total Production = 20.5 Million
Raise Analytical Capacity at CBG

Specimens Processed (M)

<table>
<thead>
<tr>
<th>Year</th>
<th>Specimens Processed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>1</td>
</tr>
<tr>
<td>2016</td>
<td>1</td>
</tr>
<tr>
<td>2017</td>
<td>1</td>
</tr>
<tr>
<td>2018</td>
<td>2</td>
</tr>
<tr>
<td>2019</td>
<td>2</td>
</tr>
<tr>
<td>2020</td>
<td>3</td>
</tr>
<tr>
<td>2021</td>
<td>3</td>
</tr>
<tr>
<td>2022</td>
<td>4</td>
</tr>
</tbody>
</table>
Barcode Production

<table>
<thead>
<tr>
<th>Years</th>
<th>Cost per Specimen</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>$1.50</td>
</tr>
<tr>
<td>2020</td>
<td>$1.25</td>
</tr>
<tr>
<td>2021</td>
<td>$1.00</td>
</tr>
<tr>
<td>2022</td>
<td>$0.75</td>
</tr>
<tr>
<td>2023</td>
<td>$0.60</td>
</tr>
<tr>
<td>2024</td>
<td>$0.50</td>
</tr>
<tr>
<td>2025</td>
<td>$0.50</td>
</tr>
</tbody>
</table>

CBG

New Facility
A Smaller Challenge for iBOL-II: Sequencing Costs

20 million specimens x $1/specimen

$20 MILLION
The Stop Dan Movement: Specimens from Other Places

Specimens per Malaise Trap per Year vs Degrees of Latitude

- Northern Hemisphere
- Southern Hemisphere
- Costa Rica
- Urban

Locations:
- Finland
- Greenland
A Mission for Planetary Biodiversity

- Census of All Species
- Library of Life
- Global Biosurveillance

Understand Biodiversity
Read Genomes
Protect Biodiversity
Each block partitioned into 60 grid squares, each 2500 km²

4000 blocks of 150,000 km²
Biosurveillance Blocks – Canada
Biosurveillance – Canada Block #70

82 SITES, 3 MILLION SPECIMENS

Map of Canada with sites numbered from 1 to 82.
Biosurveillance – Canada’s Ecoregions

Map of Canada showing various ecoregions: Taiga Cordillera, Boreal Cordillera, Pacific Maritime, Montane Cordillera, Taiga Plains, Boreal Plains, Prairie, Boreal Shield, Arctic Cordillera, Canadian Arctic Archipelago, Hudson Plains, Mixed Wood Plains, and Atlantic Maritimes.
Biosurveillance for a Nation
Kruger National Park Malaise Program
Biosurveillance with Molecules – Machines & Methods

Extract DNA → PCR → Sequence → Assign each sequence to BIN
Analyze 1 Million Insects in a Week

250,000 as single specimens
750,000 as bulk samples
Single Specimens to Expand DNA Barcode Reference Library

Total Reads: 6 million
Specimens Analyzed: 250,000
BINs Revealed: 14,000
Bulk Samples to Map Species Distributions

Total Reads: 240 Million
Specimens Analyzed: 750,000
BINs Revealed: 19,000
## What Will iBOL-II Cost?

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Cost ($ Million)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sampling</strong></td>
<td>Reference library: 20M specimens @$5</td>
<td>20 cash, 80 in-kind</td>
</tr>
<tr>
<td></td>
<td>Bio-surveillance: 10K samples @$1000</td>
<td>2 cash, 8 in-kind</td>
</tr>
<tr>
<td><strong>Imaging</strong></td>
<td>2 million GenoSpecies @$5 each</td>
<td>10</td>
</tr>
<tr>
<td><strong>Sequence Analysis</strong></td>
<td>Reference Library: 20M specimens @$1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Metabarcoding: 10,000 samples @$1000</td>
<td>10</td>
</tr>
<tr>
<td><strong>Informatics</strong></td>
<td>Data Entry</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>BOLD &amp; mBRAVE platforms</td>
<td>10</td>
</tr>
<tr>
<td><strong>Infrastructure</strong></td>
<td>Computational hardware</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Sequencing platforms</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Sampling Equipment</td>
<td>2</td>
</tr>
<tr>
<td><strong>Administration</strong></td>
<td>SSC &amp; BoD</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Communications</td>
<td>1</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>$179M</strong> ($88M in-kind, $91M cash)</td>
</tr>
<tr>
<td>Held</td>
<td>Source</td>
<td>Amount ($M)</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Informatics</td>
<td>CFREF</td>
<td>5</td>
</tr>
<tr>
<td>Core Facility</td>
<td>CFREF</td>
<td>8</td>
</tr>
<tr>
<td>Reference Library</td>
<td>Philanthropy</td>
<td>2</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td>$15 million</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Prospective</th>
<th>Source</th>
<th>Amount ($M)</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure</td>
<td>CFI</td>
<td>2</td>
<td>2019</td>
</tr>
<tr>
<td>Reference Library</td>
<td>Philanthropy</td>
<td>3</td>
<td>2020-2025</td>
</tr>
<tr>
<td>Bio-surveillance</td>
<td>NSERC</td>
<td>2</td>
<td>2019-2021</td>
</tr>
<tr>
<td>Bio-surveillance</td>
<td>Genome Canada</td>
<td>5</td>
<td>2020-2025</td>
</tr>
</tbody>
</table>
Other Funding Prospects - BiodivERsA

Biodiversity and its influence on animal, human and plant health

Eligible Nations:

<table>
<thead>
<tr>
<th>Country</th>
<th>Contribution (Euros)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>1M</td>
</tr>
<tr>
<td>Belgium</td>
<td>0.8M</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>0.45M</td>
</tr>
<tr>
<td>Estonia</td>
<td>TBD</td>
</tr>
<tr>
<td>France</td>
<td>2.3M</td>
</tr>
<tr>
<td>Germany</td>
<td>2.5M</td>
</tr>
<tr>
<td>Ireland</td>
<td>0.45M</td>
</tr>
<tr>
<td>Lithuania</td>
<td>0.1M</td>
</tr>
<tr>
<td>Poland</td>
<td>0.5M</td>
</tr>
<tr>
<td>Romania</td>
<td>0.5M</td>
</tr>
<tr>
<td>Slovakia</td>
<td>0.24M</td>
</tr>
<tr>
<td>Switzerland</td>
<td>1.6M CHF</td>
</tr>
</tbody>
</table>

Major Themes:
Relationship between biological diversity and animal, human and/or plant health: effects and underlying mechanisms

Type of Research Funded:
Collaborative research projects, generating new knowledge based on new primary data

Schedule:
Pre-proposal: November 16, 2018
Invitation to submit: Late January
Full proposal: March 15, 2018
"Tools and Technologies for Broad-Scale Pest and Disease Surveillance of Crop Plants in Low-Income Countries"

**Potential Topics:**
- Multidisciplinary approaches that leverage emerging research in data science, biology and other relevant fields
- Novel sensor-based strategies affordable to low-income country users.

**Details:**
- Application Deadline: November 14, 2018
- Maximum Request: $100,000 USD over 18 months
- Decision Date: April 14, 2019
- Possible Follow-up: If successful, chance to apply for $1M USD
Completing the Planetary Biodiversity Mission

Capacity

Hardware

Software

Community

Next Steps

2018–2025

1.5 M Species Barcoded
Bio-surveillance underway

2026–2045

20 M Species Barcoded
Global Bio-surveillance
Library of Life
Can We Afford It?

Humanity spends billions of Euros each year to advance understanding of our world.
Planetary Biodiversity $500 Million

New Horizons $700 Million
Arthropods in 0.5 hectares

Anthropocene

THE SIXTH EXTINCTION
AN UNNATURAL HISTORY
Colleagues at the Centre for Biodiversity Genomics
Colleagues on the SSC and BoD