Linking specimens and DNA sequences through BOLD

ADRIANA RADULOVICI
BARCODE OF LIFE DATA SYSTEM v4

Advancing biodiversity science through DNA-based species identification.

EXPLORE THE DATA

DESIGNED TO SUPPORT THE GENERATION & APPLICATION OF DNA BARCODE DATA
BOLD components

- ID Service
- DNA Barcode Database
- Workbench
BOLD core data

Specimen Data

Specimen Images

Trace Files

Sequences

Custom Fields
Molecular markers in BOLD
DNA Barcoding map
Barcoded animal species: 25K (13%)
Barcode Index Number (BIN)

A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System

Sujeewan Ratnasingham¹, Paul D. N. Hebert¹,²

¹ Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, Canada, ² Department of Integrative Biology, University of Guelph, Guelph, Ontario, Canada

- Sequence based OTU designation by clustering all COI sequences in BOLD (once a month)
- Automatic, standardized, persistent
Species vs. BINs

- Lab issues: mis-labeling, cross-contamination
- Taxonomic issues: misidentification, synonymy
- Biology: cryptic species, young species, introgression, hybridization
BOLD:AAB9507

Anonyx sarsi

Max. distance: 1.08%
Nearest Neighbour: 3.05%

BOLD:AAA9153
**Platorchestia** (Amphipopda: Talitridae): 9 species, 23 BINs

<table>
<thead>
<tr>
<th>BIN</th>
<th># Sequences</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOLD:AAA2949</td>
<td>168</td>
<td>platensis</td>
</tr>
<tr>
<td>BOLD:AAB3402</td>
<td>62</td>
<td>platensis, monodi, pacifica</td>
</tr>
<tr>
<td>BOLD:AAD0041</td>
<td>30</td>
<td>pacifica, parapacifica, platensis</td>
</tr>
<tr>
<td>BOLD:ACH9802</td>
<td>3</td>
<td>pacifica</td>
</tr>
<tr>
<td>BOLD:ACI1345</td>
<td>1</td>
<td>pacifica</td>
</tr>
<tr>
<td>BOLD:ACQ1642</td>
<td>3</td>
<td>munmui</td>
</tr>
<tr>
<td>BOLD:ACB6373</td>
<td>21</td>
<td>monodi, munmui</td>
</tr>
<tr>
<td>BOLD:ACH9803</td>
<td>2</td>
<td>joi</td>
</tr>
<tr>
<td>BOLD:ACH9800</td>
<td>1</td>
<td>joi</td>
</tr>
<tr>
<td>BOLD:ACH2256</td>
<td>43</td>
<td>paludosus</td>
</tr>
<tr>
<td>BOLD:ACH9801</td>
<td>4</td>
<td>pachypus</td>
</tr>
<tr>
<td>BOLD:AAY4437</td>
<td>15</td>
<td>japonica</td>
</tr>
<tr>
<td>BOLD:ACI2547</td>
<td>4</td>
<td>japonica</td>
</tr>
<tr>
<td>BOLD:ACH7347</td>
<td>3</td>
<td>japonica</td>
</tr>
<tr>
<td>BOLD:ACH2809</td>
<td>2</td>
<td>japonica</td>
</tr>
<tr>
<td>BOLD:ACH9813</td>
<td>2</td>
<td>japonica</td>
</tr>
<tr>
<td>BOLD:ACI0930</td>
<td>2</td>
<td>japonica</td>
</tr>
<tr>
<td>BOLD:ACH2808</td>
<td>1</td>
<td>japonica</td>
</tr>
<tr>
<td>BOLD:ACH2810</td>
<td>1</td>
<td>japonica</td>
</tr>
<tr>
<td>BOLD:ACI2544</td>
<td>1</td>
<td>japonica</td>
</tr>
<tr>
<td>BOLD:ACI2545</td>
<td>1</td>
<td>japonica</td>
</tr>
<tr>
<td>BOLD:ACI2546</td>
<td>1</td>
<td>japonica</td>
</tr>
<tr>
<td>BOLD:ACI2549</td>
<td>1</td>
<td>japonica</td>
</tr>
</tbody>
</table>

Platorchestia (Amphipopda: Talitridae): 9 species, 23 BINs
Marine barcoding (25K species – 29K BINs)

Thousands

- WORMS
- BOLD_rec
- BOLD_COI
- BOLD_BINs

Chordata | Crustacea | Mollusca | Polychaeta | Echinodermata | Cnidaria | Porifera
---|---|---|---|---|---|---
0 | 50 | 60 | 10 | 5 | 0 | 0
DESIGNED TO SUPPORT THE GENERATION & APPLICATION OF DNA BARCODE DATA

BOLD is a cloud-based data storage and analysis platform developed at the Centre for Biodiversity Genomics in Canada. It consists of four main modules: a data portal, an educational portal, a registry of BOLD-generative species, and a data collection and analysis workbench. Please note that this version of BOLD is in beta and will contain bugs. Users can help address these bugs by testing the system and reporting bugs to support@boldsystems.org. This version is very different from the prior one but has access to all the same data.

DATA PORTAL
A data retrieval interface that allows for searching over 1.7M public records in BOLD using multiple search criteria including but not limited to, geography, taxonomy, and deposition.

EDUCATION PORTAL
A custom platform for educators and students to explore barcode data and contribute novel barcodes to the BOLD database.

BIN DATABASE
A searchable database of Barcode Index Numbers (BINs), sequence clusters that closely approximate species.

WORKBENCH
A data collection and analysis environment that supports the assembly and validation of whole barcodes and other sequence sets.

mBRAVE is now accepting all new registrants.
iBOL Consortium
4-5 February 2019

MARINE INVERTEBRATE DIGITIZATION WORKSHOP
BIOSCAN has four main objectives:

1. SCAN SPECIES COMMUNITIES AT 2,500 SITES
2. CODIFY SPECIES INTERACTIONS AT 2,500 SITES
3. EXPAND THE REFERENCE LIBRARY TO 2.5 MILLION SPECIES
4. SPEED THE UPTAKE OF DNA BARCODING TO PROTECT LIFE
Thank you