www.boldsystems.org

Adriana Radulovici

aradulov@uoguelph.ca

TAXONOMY

Kingdoms of Life Being Barcoded

SEARCH TAXONOMY

8,589,262	6,130,973	280,376
Specimen Records	Specimens with Barcodes	Species with Barcodes

Animals:

- Acanthocephala [858]
- Annelida [66285]
- Arthropoda [6934123]
- Brachiopoda [236]
- Bryozoa [2928]
- Chaetognatha [773]
- Chordata [609399]
- Cnidaria [20034]
- Cycliophora [326]
- Echinodermata [43320]
- Gnathostomulida [18]
- Hemichordata [167]
- Mollusca [162692]
- Nematoda [14543]
- Nemertea [2911]
- Onychophora [835]
- Platyhelminthes [17201]
- Porifera [4185]
- Priapulida [46]
- Rotifera [10761]
- Sipuncula [781]
- Tardigrada [2076]
- Xenoturbellida [7]

Plants:

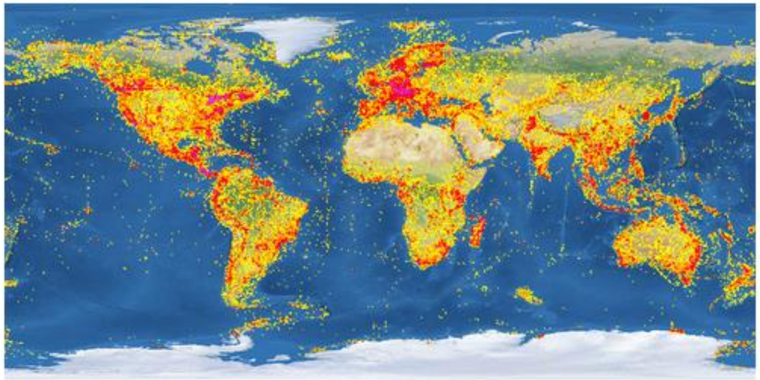
- Bryophyta [11973]
- Chlorophyta [12749]
- Lycopodiophyta [1142]
- Magnoliophyta [346771]
- Pinophyta [6909]
- Pteridophyta [10617]
- Rhodophyta [50017]

Fungi:

- Ascomycota [84997]
- Basidiomycota [57169]
- Chytridiomycota [277]
- Glomeromycota [3528]
- Myxomycota [226]
- Zygomycota [3136]

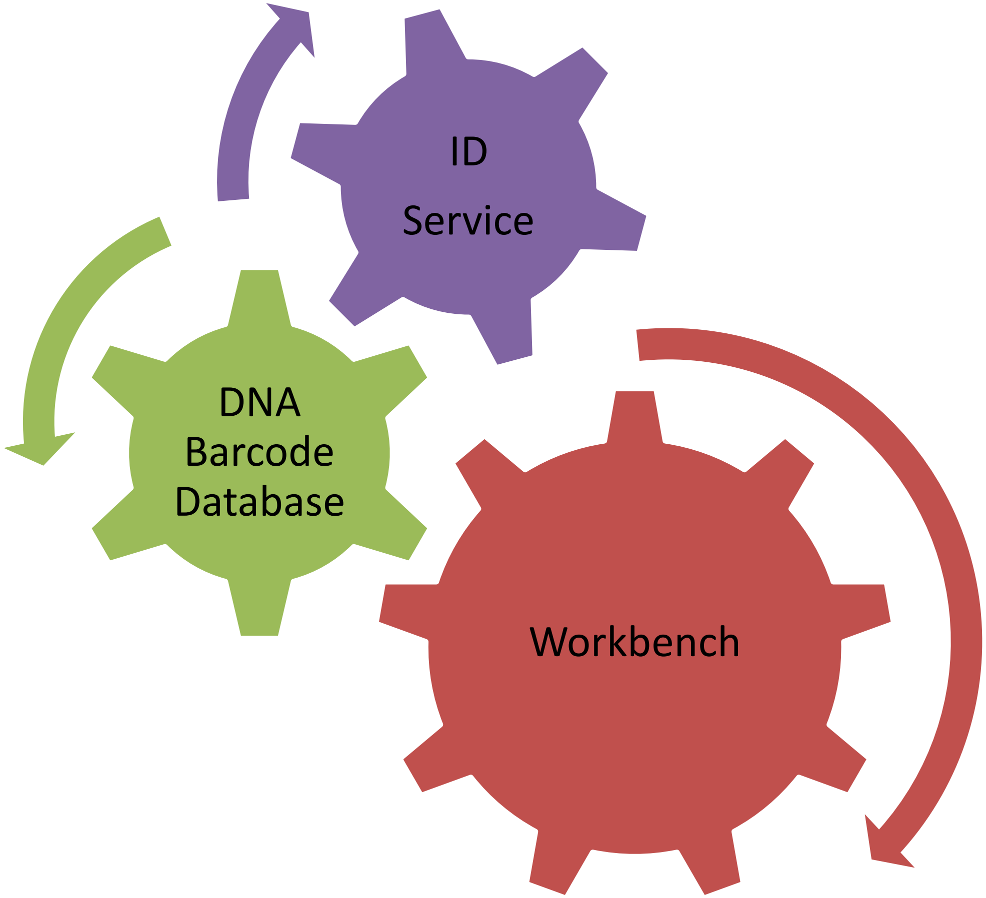
Protists:

- Chlorarachniophyta [67]
- Ciliophora [707]
- Heterokontophyta [6296]
- Pyrrophycophyta [2281]

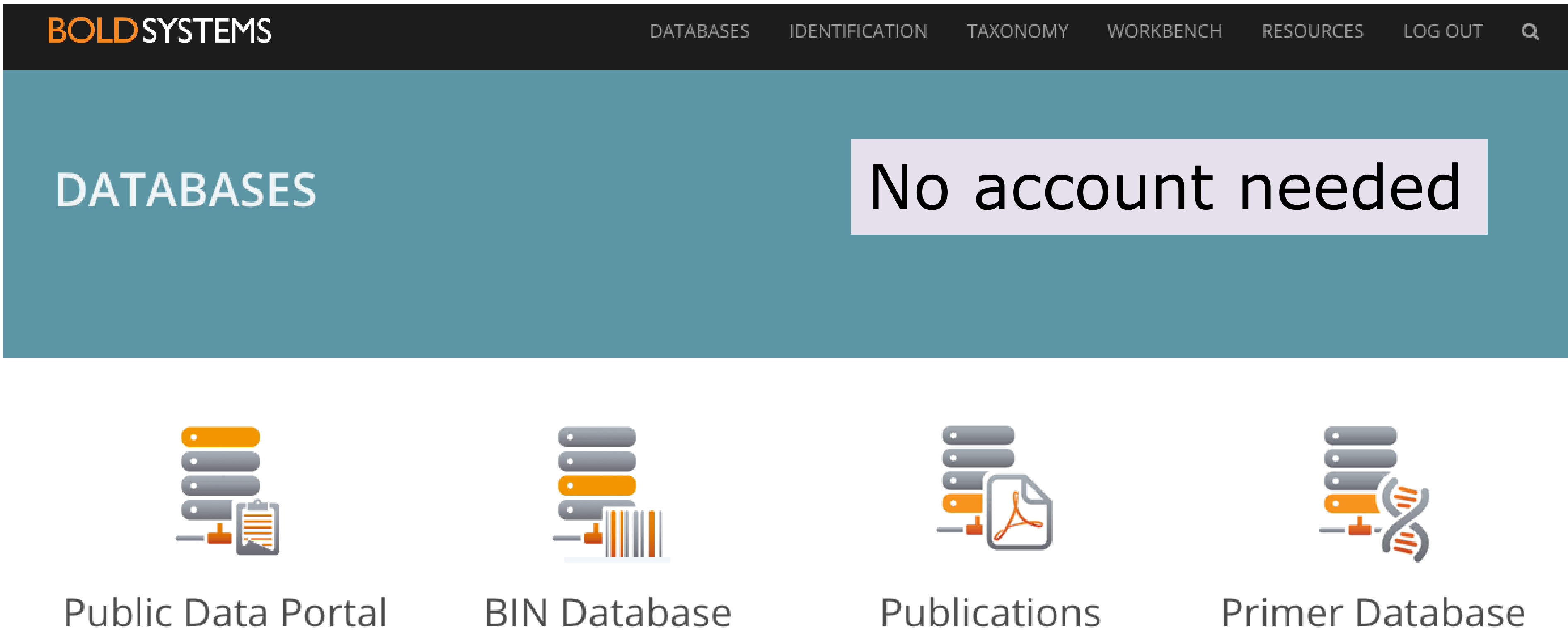


GTI DNA tech

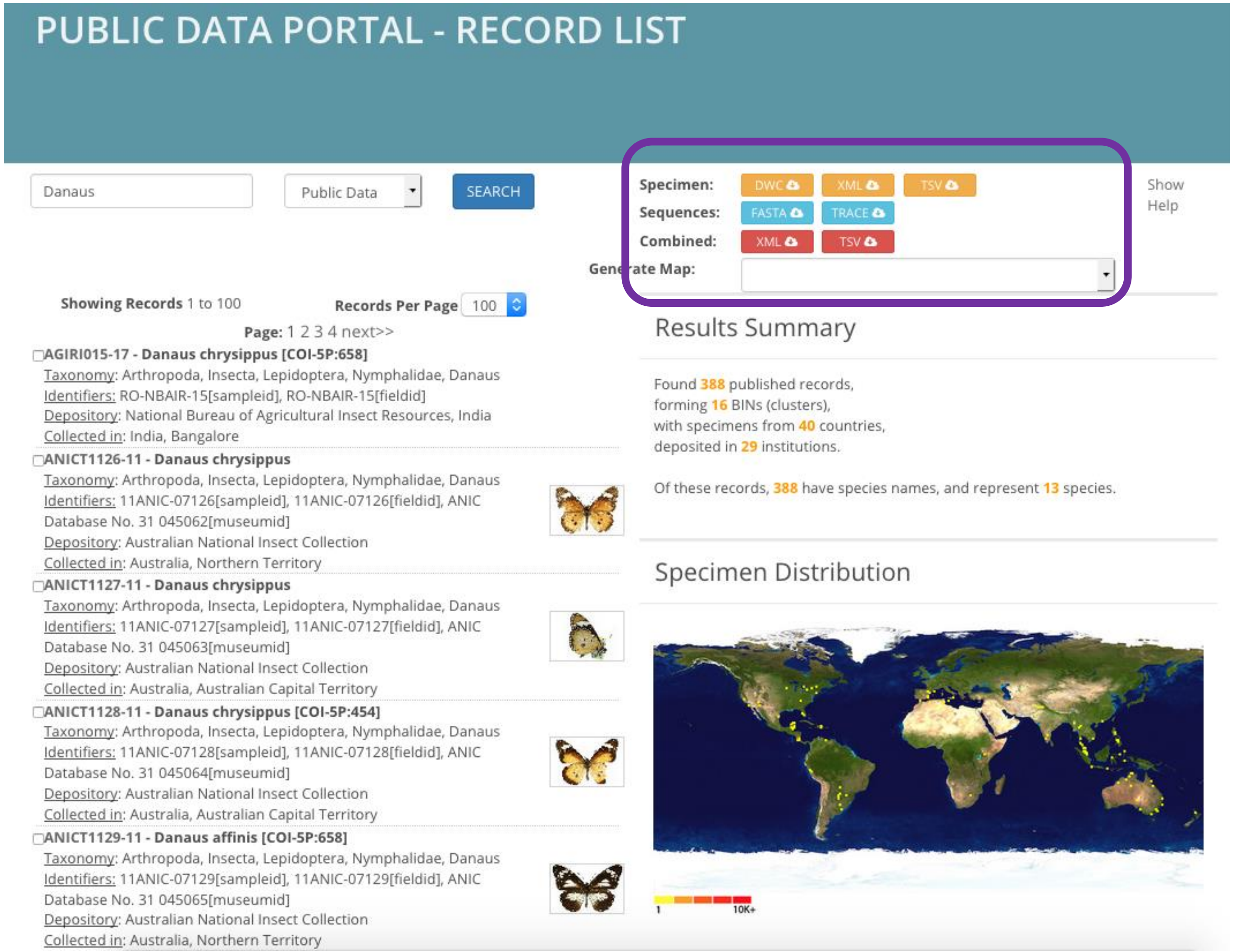
BOLD components and services



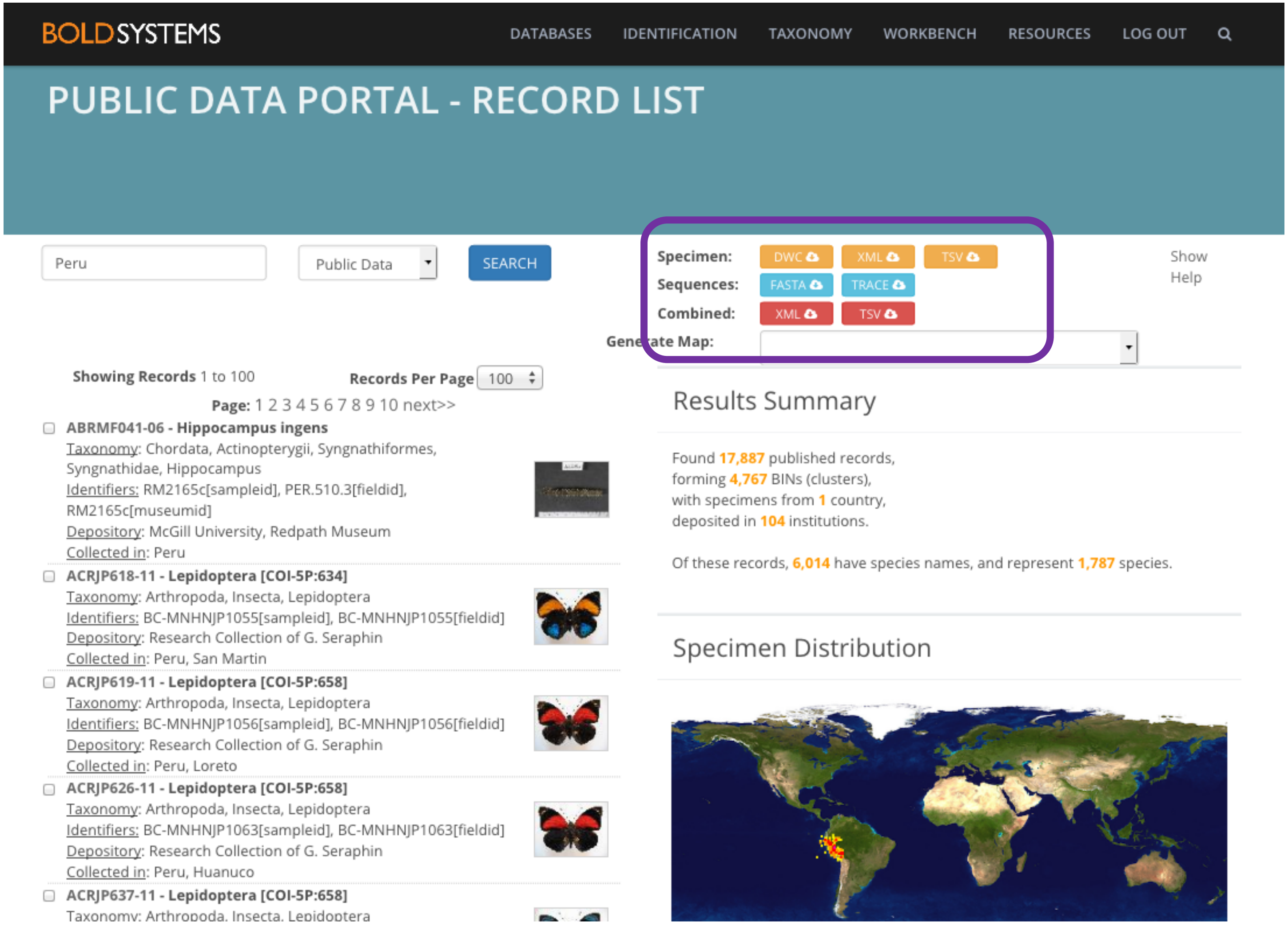
BOLD component databases



Data mining


PUBLIC DATA PORTAL - RECORD LIST
Danaus
Public Data
SEARCH
Specimen: DWC XML TSV
Sequences: FASTA TRACE
Combined: XML TSV
Generate Map:
Show
Help
Showing Records 1 to 100
Records Per Page 100
Page: 1 2 3 4 next>>
AGIRI015-17 - Danaus chrysippus [COI-5P:658]
Taxonomy: Arthropoda, Insecta, Lepidoptera, Nymphalidae, Danaus
Identifiers: RO-NBAIR-15[sampleid], RO-NBAIR-15[fieldid]
Depository: National Bureau of Agricultural Insect Resources, India
Collected in: India, Bangalore
ANICT1126-11 - Danaus chrysippus
Taxonomy: Arthropoda, Insecta, Lepidoptera, Nymphalidae, Danaus
Identifiers: 11ANIC-07126[sampleid], 11ANIC-07126[fieldid], ANIC Database No. 31 045062[museumid]
Depository: Australian National Insect Collection
Collected in: Australia, Northern Territory
ANICT1127-11 - Danaus chrysippus
Taxonomy: Arthropoda, Insecta, Lepidoptera, Nymphalidae, Danaus
Identifiers: 11ANIC-07127[sampleid], 11ANIC-07127[fieldid], ANIC Database No. 31 045063[museumid]
Depository: Australian National Insect Collection
Collected in: Australia, Australian Capital Territory
ANICT1128-11 - Danaus chrysippus [COI-5P:454]
Taxonomy: Arthropoda, Insecta, Lepidoptera, Nymphalidae, Danaus
Identifiers: 11ANIC-07128[sampleid], 11ANIC-07128[fieldid], ANIC Database No. 31 045064[museumid]
Depository: Australian National Insect Collection
Collected in: Australia, Australian Capital Territory
ANICT1129-11 - Danaus affinis [COI-5P:658]
Taxonomy: Arthropoda, Insecta, Lepidoptera, Nymphalidae, Danaus
Identifiers: 11ANIC-07129[sampleid], 11ANIC-07129[fieldid], ANIC Database No. 31 045065[museumid]
Depository: Australian National Insect Collection
Collected in: Australia, Northern Territory
Results Summary
Found 388 published records,
forming 16 BINs (clusters),
with specimens from 40 countries,
deposited in 29 institutions.
Of these records, 388 have species names, and represent 13 species.
Specimen Distribution
1
10K+

Data mining


BOLDSYSTEMS
DATABASES
IDENTIFICATION
TAXONOMY
WORKBENCH
RESOURCES
LOG OUT
PUBLIC DATA PORTAL - RECORD LIST
Peru
Public Data
SEARCH
Specimen:
DWC
XML
TSV
Sequences:
FASTA
TRACE
Combined:
XML
TSV
Show Help
Generate Map:
Showing Records 1 to 100
Records Per Page 100
Page: 1 2 3 4 5 6 7 8 9 10 next>>
ABRMF041-06 - Hippocampus ingens
Taxonomy: Chordata, Actinopterygii, Syngnathiformes, Syngnathidae, Hippocampus
Identifiers: RM2165c[sampleid], PER.510.3[fieldid], RM2165c[museumid]
Depository: McGill University, Redpath Museum
Collected in: Peru
ACRJP618-11 - Lepidoptera [COI-5P:634]
Taxonomy: Arthropoda, Insecta, Lepidoptera
Identifiers: BC-MNHNJP1055[sampleid], BC-MNHNJP1055[fieldid]
Depository: Research Collection of G. Seraphin
Collected in: Peru, San Martin
ACRJP619-11 - Lepidoptera [COI-5P:658]
Taxonomy: Arthropoda, Insecta, Lepidoptera
Identifiers: BC-MNHNJP1056[sampleid], BC-MNHNJP1056[fieldid]
Depository: Research Collection of G. Seraphin
Collected in: Peru, Loreto
ACRJP626-11 - Lepidoptera [COI-5P:658]
Taxonomy: Arthropoda, Insecta, Lepidoptera
Identifiers: BC-MNHNJP1063[sampleid], BC-MNHNJP1063[fieldid]
Depository: Research Collection of G. Seraphin
Collected in: Peru, Huanuco
ACRJP637-11 - Lepidoptera [COI-5P:658]
Taxonomy: Arthropoda, Insecta, Lepidoptera
Results Summary
Found 17,887 published records, forming 4,767 BINs (clusters), with specimens from 1 country, deposited in 104 institutions.
Of these records, 6,014 have species names, and represent 1,787 species.
Specimen Distribution

Application Programming Interface

API HOME

- Public Data API

  The BOLD Web Services provides the ability to query and retrieve data from the BOLD Public Data Portal.

- Taxonomy API

  The BOLD Taxonomy Web Service provides the ability to query and retrieve data from the BOLD Taxonomy Browser.

- ID Engine API

  The BOLD ID Engine Web Service provides the ability to query and retrieve results from the BOLD ID Engine.

- URL Generator

  The BOLD URL Generator for the APIs.

http://www.boldsystems.org/index.php/API_Public/combined?geo=Ontario&format=tsv

GTI DNA tech

BOLD database

Animals

192,000 species
+365,000 BINs

Fungi

17,700 species
+11,000 Interim names

Plants

67,200 species
+8,500 Interim names

Protists

2,560 species
+2,040 Interim names

8,609,727
Specimen Records

6,137,511
Specimens with Barcodes

BOLDSYSTEMS DATABASES IDENTIFICATION TAXONOMY WORKBENCH RESOURCES LOG OUT

IDENTIFICATION ENGINE

ANIMAL IDENTIFICATION [COI] FUNGAL IDENTIFICATION [ITS] PLANT IDENTIFICATION [RBCL & MATK]

The BOLD Identification System (IDS) for COI accepts sequences from the 5' region of the mitochondrial Cytochrome c oxidase subunit I gene and returns a species-level identification when one is possible. Further validation with independent genetic markers will be desirable in some forensic applications.

Historical Databases: Current Jul-2015 Jul-2014 Jul-2013 Jul-2012 Jul-2011 Jul-2010 Jul-2009

Search Databases:

- **All Barcode Records on BOLD (4,835,882 Sequences)**
  Every COI barcode record on BOLD with a minimum sequence length of 500bp (warning: unvalidated library and includes records without species level identification). This includes many species represented by only one or two specimens as well as all species with interim taxonomy. This search only returns a list of the nearest matches and does not provide a probability of placement to a taxon.
- **Species Level Barcode Records (2,738,067 Sequences/176,270 Species/64,832 Interim Species)**
  Every COI barcode record with a species level identification and a minimum sequence length of 500bp. This includes many species represented by only one or two specimens as well as all species with interim taxonomy.
- **Public Record Barcode Database (1,057,721 Sequences/86,122 Species/18,190 Interim Species)**
  All published COI records from BOLD and GenBank with a minimum sequence length of 500bp. This library is a collection of records from the published projects section of BOLD.
- **Full Length Record Barcode Database (1,713,726 Sequences/159,147 Species/56,298 Interim Species)**
  Subset of the Species library with a minimum sequence length of 640bp and containing both public and private records. This library is intended for short sequence identification as it provides maximum overlap with short reads from the barcode region of COI.

Enter sequences in fasta format:

```
GATCAAATTTATAATGTAATTGTTACTGCCCATGCTTTTGTAATAATTTTTTTCATAGTTATACCTATTA
TAATTGGAGGATTTGGTAATTGATTAGTACCTTTAATACTTGGAGCACCAGATATAGCATTTCCTCGAAT
AAATAATATAAGTTTTTGATTATTACCTCCTTCCCTTACATTATTATTAGTAAGTAGTATAGTAGAAAAT
GGAGCTGGAACAGGTTGAACAGTTTACCCTCCTCTTTCTTCTGTGATTGCTCATGGAGGTGCTTCTGTTG
ACTTAGCAATTTTTTCTCTTCATTTAGCTGGAATTTCTTCTATTTTAGGGGCTGTAAATTTTATCACAAC
AGTAATTAATATACGATCTACCGGAATTTCATTCGACCGAATACCTTTATTTGTTTGAGCAGTAGTTCTT
ACTGCACTACTTCTATTATTATCTTTACCAGTTTTAGCAGGAGCTAT
```

GTI DNA tech

Results Summary

Download

Query ID	Best ID	Search DB	Tree	Top %	Graph	Low %
unlabeled_sequence	*Ceratitis capitata*	COI SPECIES DATABASE		100.00		99.13

Query: unlabeled_sequence
Top Hit: Arthropoda Insecta - Diptera - *Ceratitis capitata* (100%)

Search Result:

A species level match could not be made, the queried specimen is likely to be one of the following:

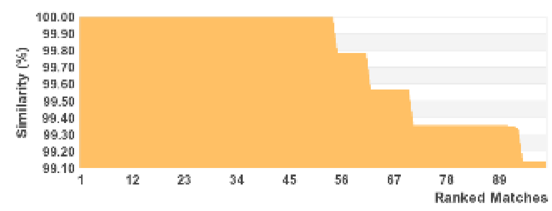
Ceratitis capitata
Ceratitis caetrata

For a hierarchical placement - a neighbor-joining tree is provided: TREE BASED IDENTIFICATION

Identification Summary

Taxonomic Level	Taxon Assignment	Probability of Placement (%)
Phylum	Arthropoda	100
Class	Insecta	100
Order	Diptera	100
Family	Tephritidae	100
Genus	*Ceratitis*	100

Similarity Scores of Top 99 Matches



Display: Top 20

Top 20 Matches

Phylum	Class	Order	Family	Genus	Species	Subspecies	Similarity (%)	Status
Arthropoda	Insecta	Diptera	Tephritidae	*Ceratitis*	*capitata*		100	Published
Arthropoda	Insecta	Diptera	Tephritidae	*Ceratitis*	*capitata*		100	Published
Arthropoda	Insecta	Diptera	Tephritidae	*Ceratitis*	*capitata*		100	Published

Results Summary

Download

Query ID	Best ID	Search DB	Tree	Top %	Graph	Low %
unlabeled_sequence	*Ceratitis capitata*	COI SPECIES DATABASE		100.00		99.13

Query: unlabeled_sequence
Top Hit: Arthropoda Insecta - Diptera - *Ceratitis capitata* (100%)

Search Result:

A species level match could not be made, the queried specimen i

Ceratitis capitata
Ceratitis caetrata

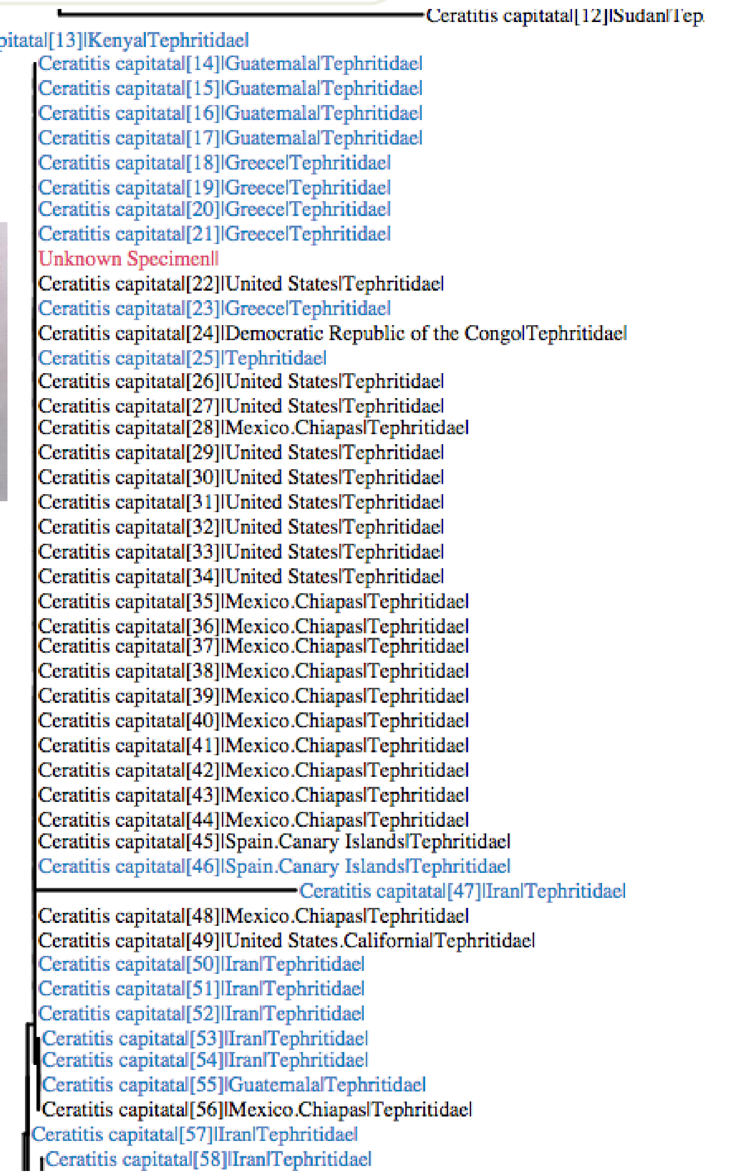
For a hierarchical placement - a neighbor-joining tree is provided:

Identification Summary

Taxonomic Level	Taxon Assignment	Probability of Placement (%)
Phylum	Arthropoda	100
Class	Insecta	100
Order	Diptera	100
Family	Tephritidae	100
Genus	*Ceratitis*	100

Top 20 Matches

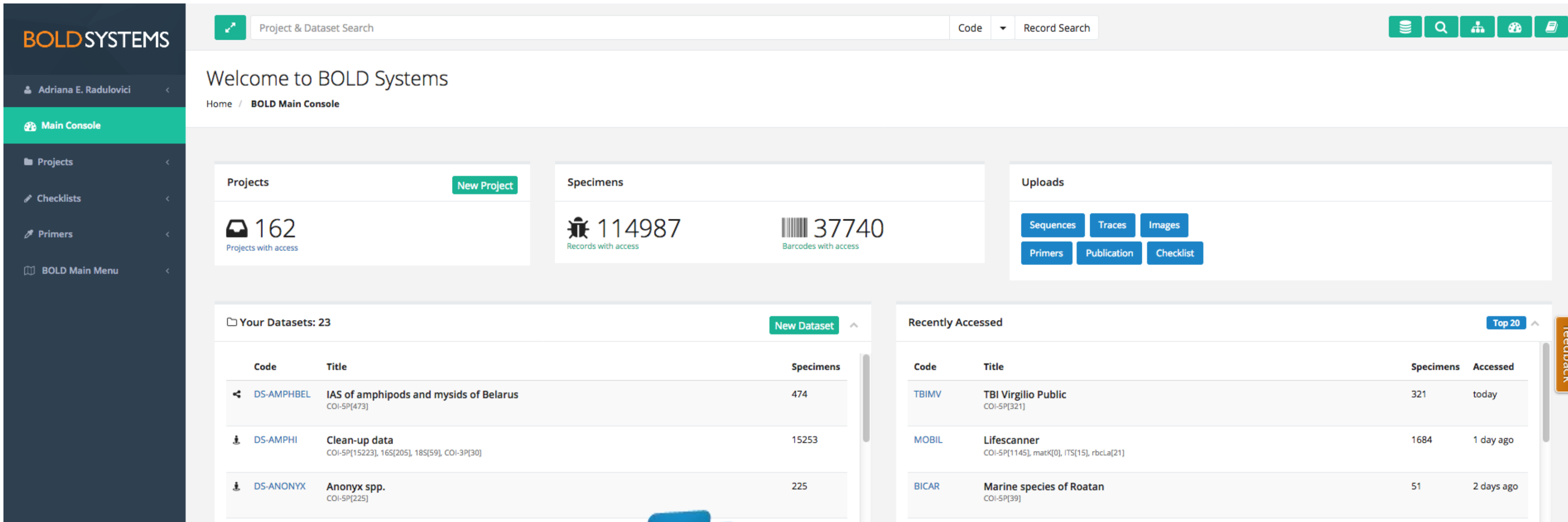
Phylum	Class	Order	Family	Genus	Sp
Arthropoda	Insecta	Diptera	Tephritidae	*Ceratitis*	*ca*
Arthropoda	Insecta	Diptera	Tephritidae	*Ceratitis*	*ca*
Arthropoda	Insecta	Diptera	Tephritidae	*Ceratitis*	*ca*

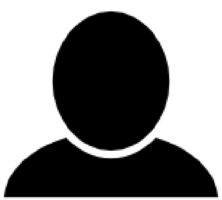


BOLD workbench

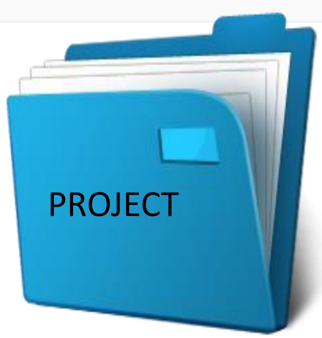
Username:

Password:

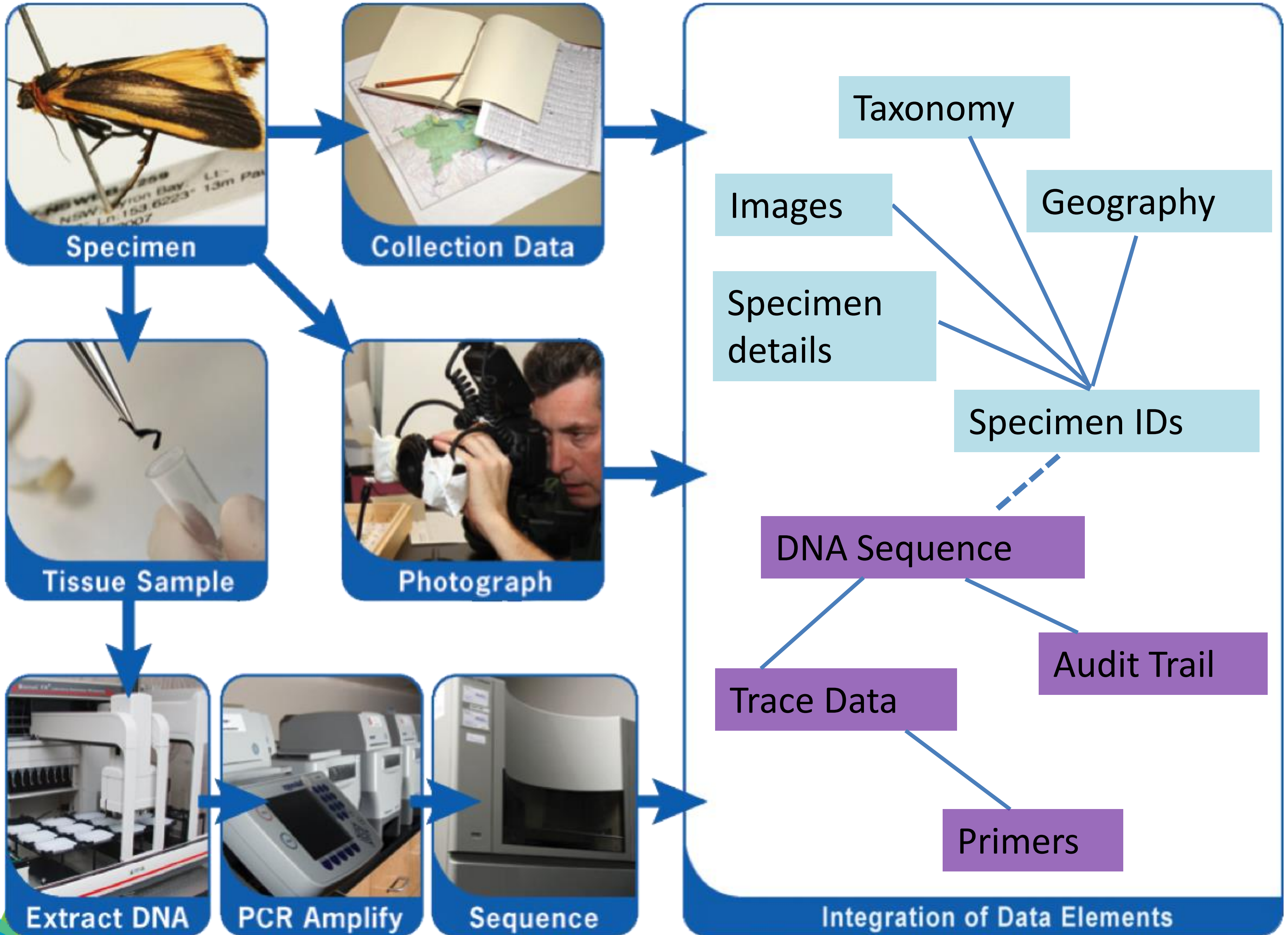


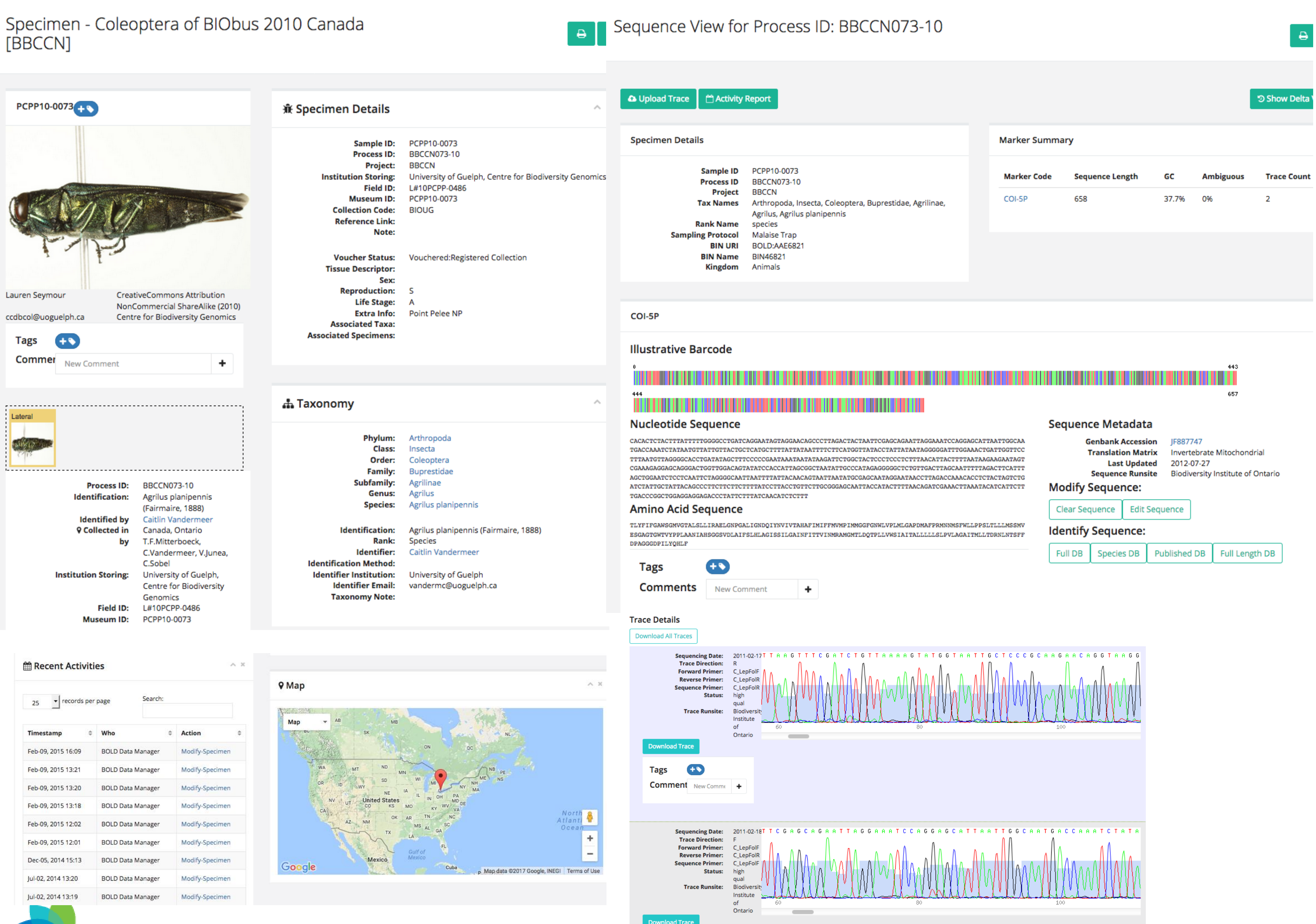


Project manager



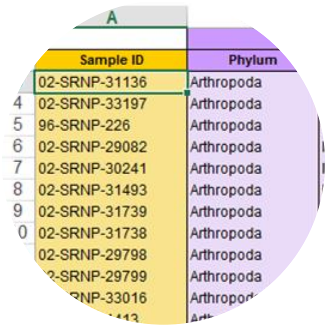
Collaborators (view and/or edit contents)


Specimen
Collection Data
Tissue Sample
Photograph
Extract DNA
PCR Amplify
Sequence
Taxonomy
Images
Geography
Specimen details
Specimen IDs
DNA Sequence
Audit Trail
Trace Data
Primers
Integration of Data Elements
GTI DNA tech


Specimen - Coleoptera of BIObus 2010 Canada [BBCCN]
PCPP10-0073
Lauren Seymour
ccdbcol@uoguelph.ca
CreativeCommons Attribution NonCommercial ShareAlike (2010)
Centre for Biodiversity Genomics
Tags
Comment
New Comment
Lateral
Process ID: BBCCN073-10
Identification: Agrilus planipennis (Fairmaire, 1888)
Identified by: Caitlin Vandermeer
Collected in: Canada, Ontario
by: T.F.Mitterboeck, C.Vandermeer, V.Junea, C.Sobel
Institution Storing: University of Guelph, Centre for Biodiversity Genomics
Field ID: L#10PCPP-0486
Museum ID: PCPP10-0073
Specimen Details
Sample ID: PCPP10-0073
Process ID: BBCCN073-10
Project: BBCCN
Institution Storing: University of Guelph, Centre for Biodiversity Genomics
Field ID: L#10PCPP-0486
Museum ID: PCPP10-0073
Collection Code: BIOUG
Reference Link:
Note:
Voucher Status: Vouchered:Registered Collection
Tissue Descriptor:
Sex:
Reproduction: S
Life Stage: A
Extra Info: Point Pelee NP
Associated Taxa:
Associated Specimens:
Taxonomy
Phylum: Arthropoda
Class: Insecta
Order: Coleoptera
Family: Buprestidae
Subfamily: Agrilinae
Genus: Agrilus
Species: Agrilus planipennis
Identification: Agrilus planipennis (Fairmaire, 1888)
Rank: Species
Identifier: Caitlin Vandermeer
Identification Method:
Identifier Institution: University of Guelph
Identifier Email: vandermc@uoguelph.ca
Taxonomy Note:
Recent Activities
25 records per page
Search:
Timestamp | Who | Action
Feb-09, 2015 16:09 | BOLD Data Manager | Modify-Specimen
Feb-09, 2015 13:21 | BOLD Data Manager | Modify-Specimen
Feb-09, 2015 13:20 | BOLD Data Manager | Modify-Specimen
Feb-09, 2015 13:18 | BOLD Data Manager | Modify-Specimen
Feb-09, 2015 12:02 | BOLD Data Manager | Modify-Specimen
Feb-09, 2015 12:01 | BOLD Data Manager | Modify-Specimen
Dec-05, 2014 15:13 | BOLD Data Manager | Modify-Specimen
Jul-02, 2014 13:20 | BOLD Data Manager | Modify-Specimen
Jul-02, 2014 13:19 | BOLD Data Manager | Modify-Specimen
Map
United States
Mexico
North Atlantic Ocean
Google
Map data ©2017 Google, INEGI Terms of Use
Sequence View for Process ID: BBCCN073-10
Upload Trace
Activity Report
Show Delta
Specimen Details
Sample ID: PCPP10-0073
Process ID: BBCCN073-10
Project: BBCCN
Tax Names: Arthropoda, Insecta, Coleoptera, Buprestidae, Agrilinae, Agrilus, Agrilus planipennis
Rank Name: species
Sampling Protocol: Malaise Trap
BIN URI: BOLD:AAE6821
BIN Name: BIN46821
Kingdom: Animals
Marker Summary
Marker Code | Sequence Length | GC | Ambiguous | Trace Count
COI-5P | 658 | 37.7% | 0% | 2
COI-5P
Illustrative Barcode
Nucleotide Sequence
Amino Acid Sequence
Sequence Metadata
Genbank Accession: JF887747
Translation Matrix: Invertebrate Mitochondrial
Last Updated: 2012-07-27
Sequence Runsite: Biodiversity Institute of Ontario
Modify Sequence:
Clear Sequence
Edit Sequence
Identify Sequence:
Full DB
Species DB
Published DB
Full Length DB
Tags
Comments
New Comment
Trace Details
Download All Traces
Sequencing Date: 2011-02-17
Trace Direction: R
Forward Primer: C_LepFolF
Reverse Primer: C_LepFolR
Sequence Primer: C_LepFolR
Status: high qual
Trace Runsite: Biodiversity Institute of Ontario
Download Trace
Tags
Comment
Sequencing Date: 2011-02-18
Trace Direction: F
Forward Primer: C_LepFolF
Reverse Primer: C_LepFolR
Sequence Primer: C_LepFolF
Status: high qual
Trace Runsite: Biodiversity Institute of Ontario
Download Trace

GTI DNA tech

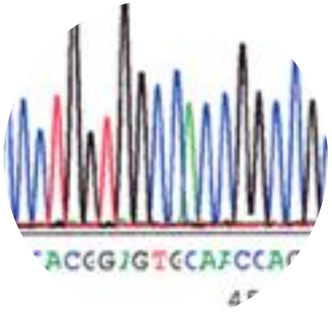
BOLD core data



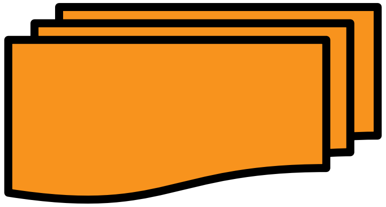
Specimen Data

Specimen Images

Sequences



Trace Files



Custom Fields

Data validation

Automated Validation

Issues

Seqs lacking successful traces

- COI-5P [21]

Seqs with stop codons

- 0

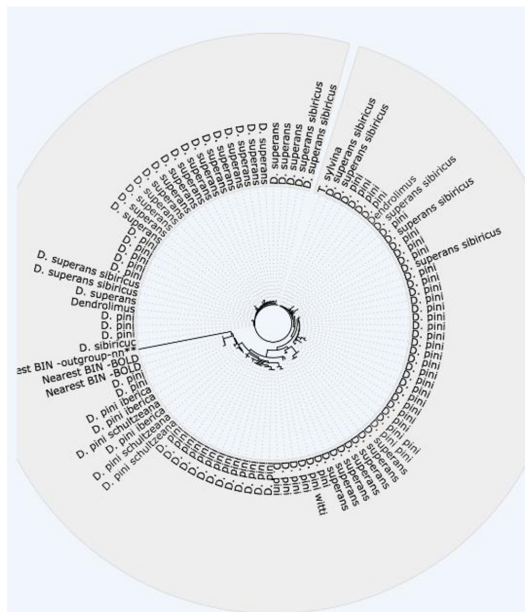
Contaminated seqs

- 0

Problematic records flagged

- 2

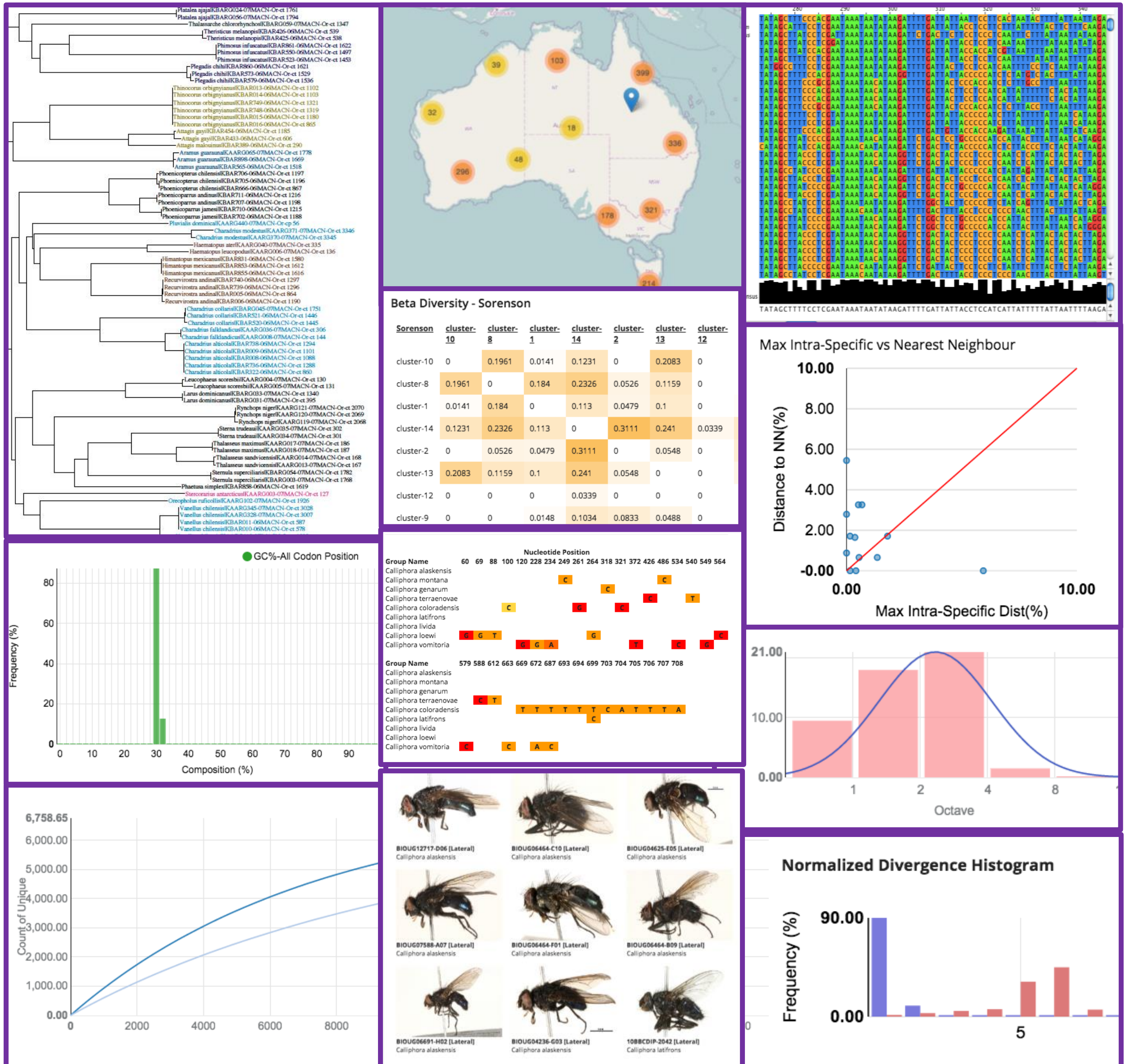
Integrated Validation



Independent Validation

Sample ID	Phylum	Class
2-SRNP-31136	Arthropoda	Insecta
2-SRNP-33197	Arthropoda	Insecta
6-SRNP-226	Arthropoda	Insecta
2-SRNP-29082	Arthropoda	Insecta
2-SRNP-30241	Arthropoda	Insecta
2-SRNP-31493	Arthropoda	Insecta
2-SRNP-31739	Arthropoda	Insecta
2-SRNP-31738	Arthropoda	Insecta
2-SRNP-29798	Arthropoda	Insecta
2-SRNP-29799	Arthropoda	Insecta
2-SRNP-33016	Arthropoda	Insecta
1-SRNP-1413	Arthropoda	Insecta
4-SRNP-9389	Arthropoda	Insecta
2-SRNP-29054	Arthropoda	Insecta
2-SRNP-31377	Arthropoda	Insecta
2-SRNP-29083	Arthropoda	Insecta
2-SRNP-33002	Arthropoda	Insecta
2-SRNP-31741	Arthropoda	Insecta
7-SRNP-9604	Arthropoda	Insecta
9-SRNP-5312	Arthropoda	Insecta
2-SRNP-31745	Arthropoda	Insecta
6-SRNP-227	Arthropoda	Insecta
2-SRNP-29080	Arthropoda	Insecta

BOLD: storage, validation, analysis & publication of DNA barcode data



Barcode Index Number (BIN)

OPEN ACCESS Freely available online — PLOS ONE

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

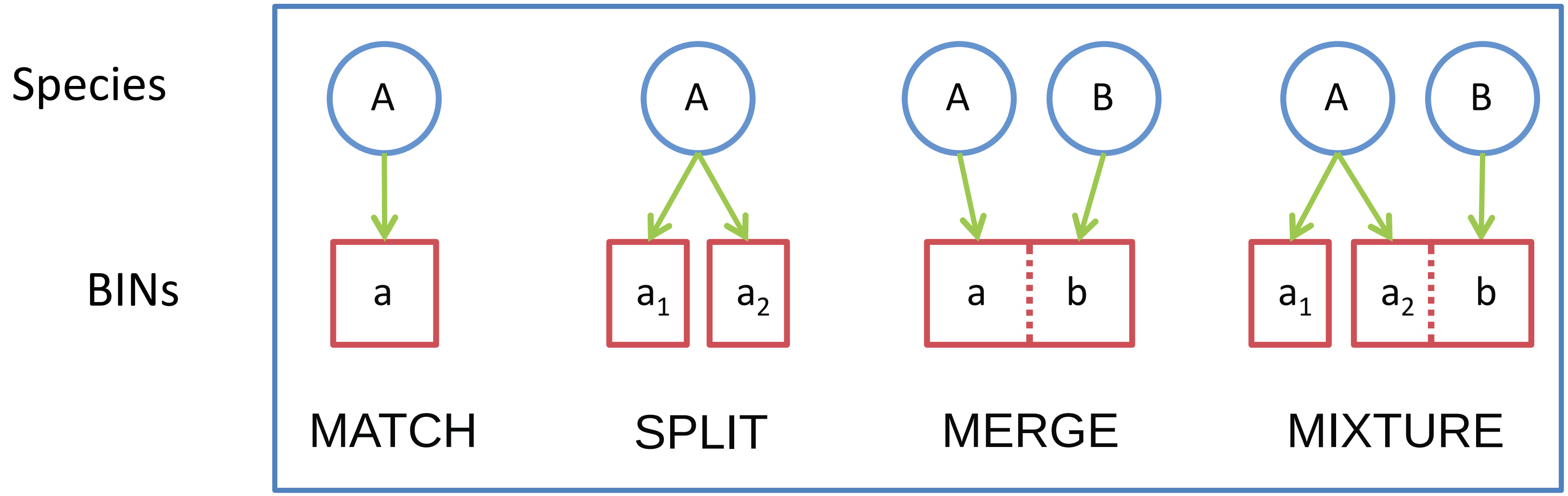
Sujeevan Ratnasingham[1]*, Paul D. N. Hebert[1,2]

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

BOLD:AAA3297

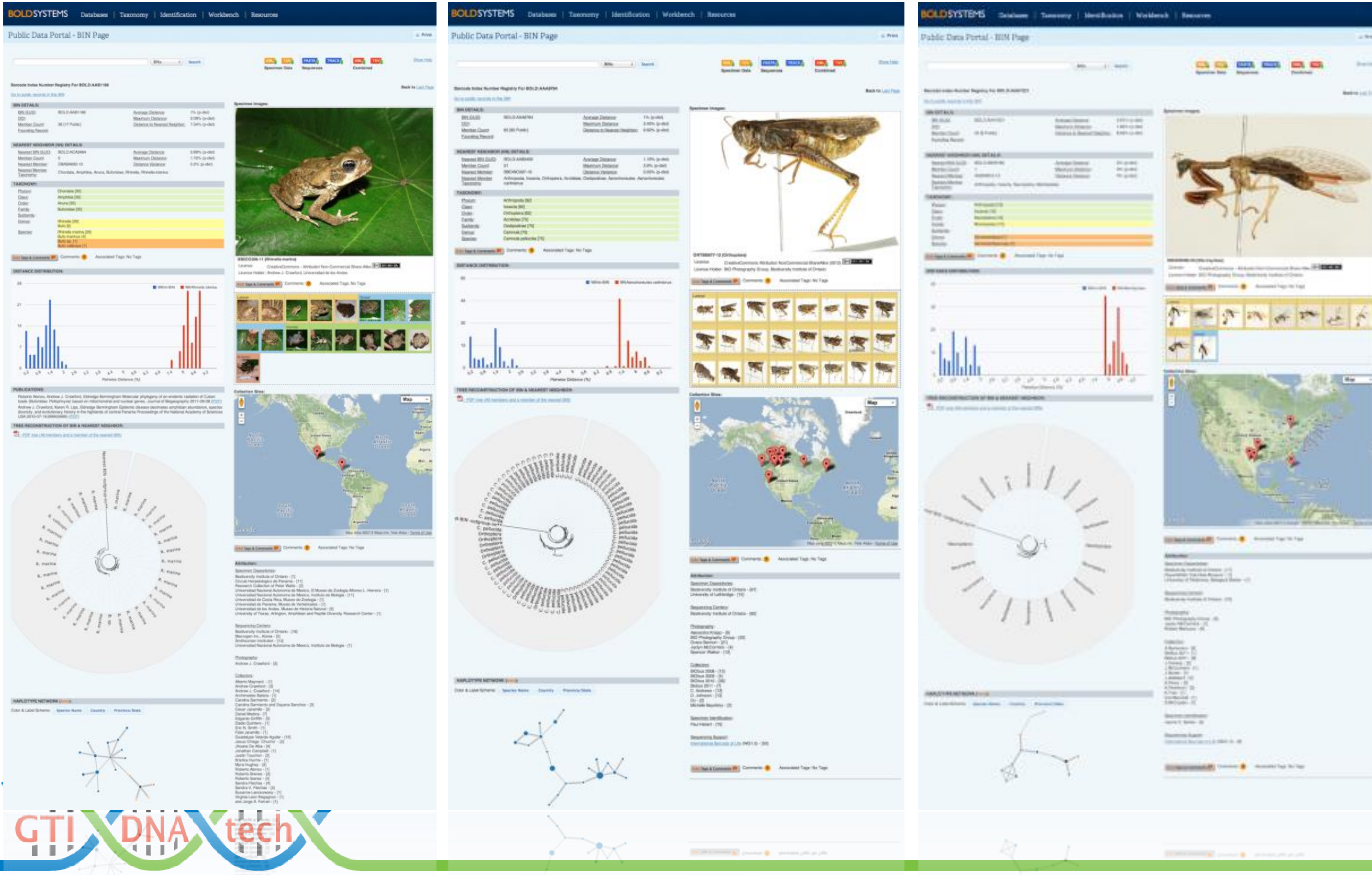
- **Sequence based OTU designation by clustering all COI sequences in BOLD (once a month)**
- **Automatic, standardized, persistent**

Species vs. BINs



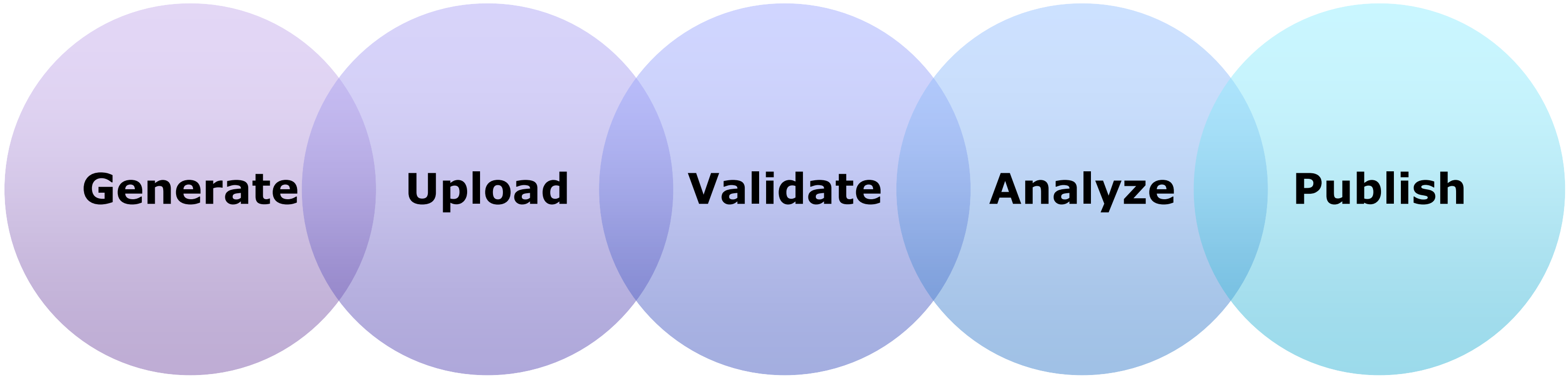
- Lab issues: cross-contamination, plate flips
- Taxonomic issues: misidentification, synonymy
- Biology: cryptic species, young species, introgression, hybridization

A unique page for each BIN



GTI DNA tech

Data life cycle



Data life cycle