Use of DNA Barcoding in Regulatory Science

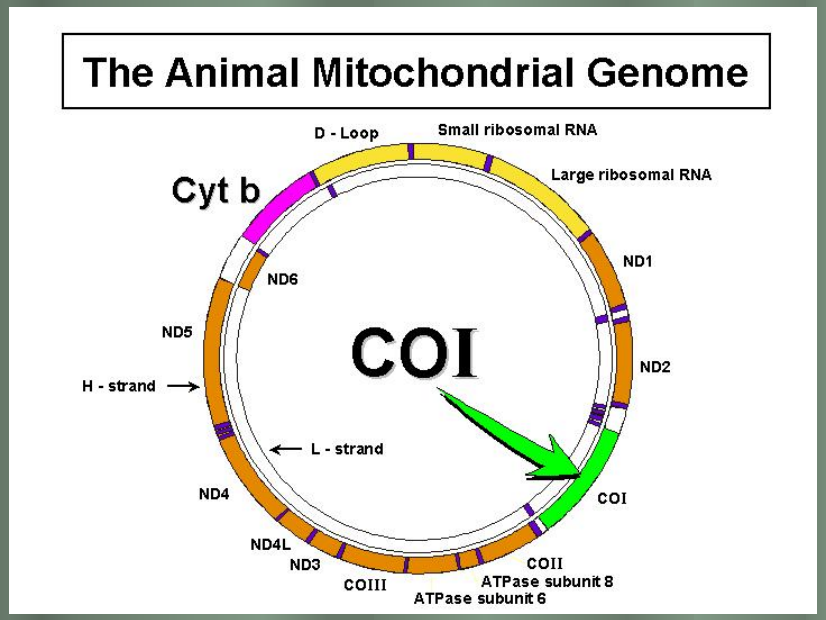
Haile Yancy, Ph.D.
U.S. Food & Drug Administration
Center for Veterinary Medicine

What is Barcoding?

A DNA barcode is a short gene sequence used to identify species taken from a standard position in the genome

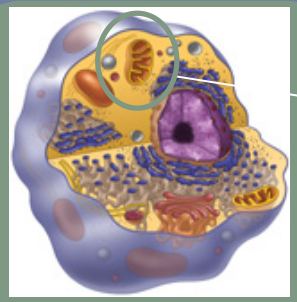
It is much like a barcode on a cereal box, except we use nucleotides instead of numbers to identify the species

Horizontal Genomics and COI

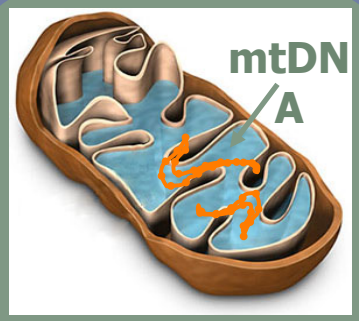


- Traditional genomics measures all genes for one species
- Horizontal genomics measures all species using one gene
- Gene of interest: COI (Cytochrome c Oxidase Subunit 1)

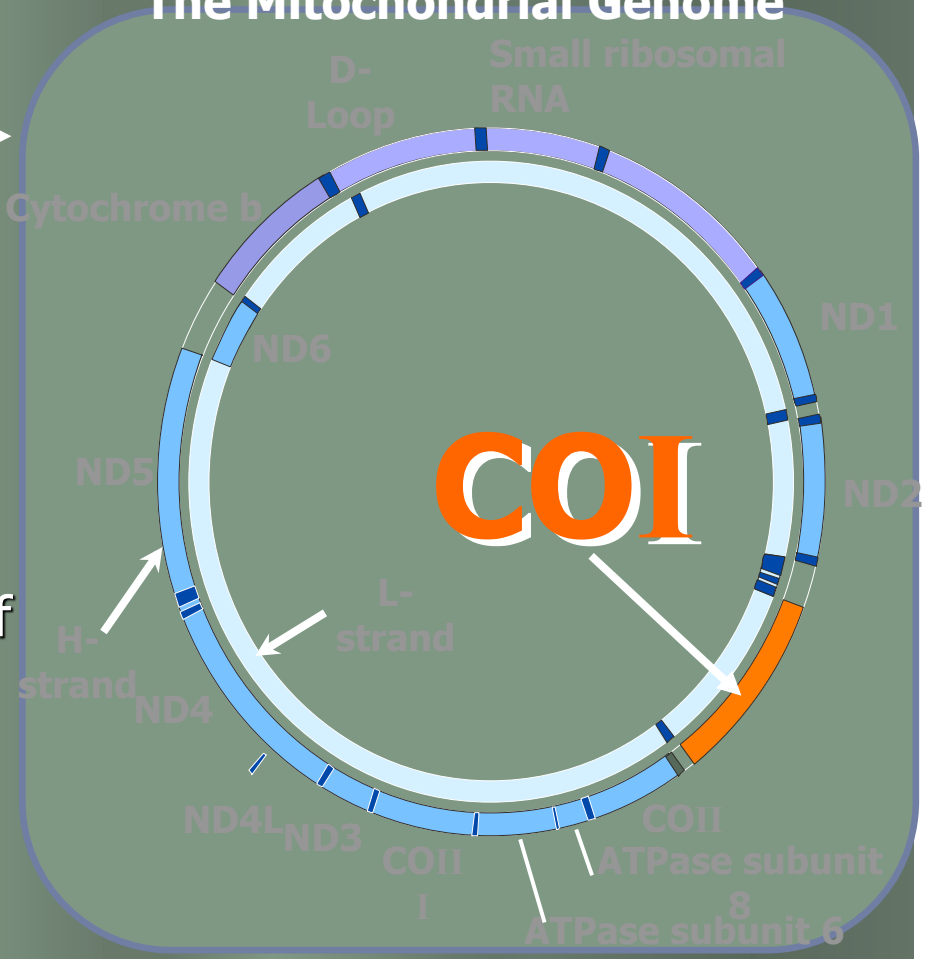
Cytochrome c oxidase subunit 1



Typical Animal Cell



Mitochondrion



- Eukaryotic and mitochondrial gene found in all animals
- No pseudogenes
- Slowest rate of amino acid change of any of the mitochondrial genes
- Enough sequence divergence to identify closely related species
- 5′ end is recovered easily from most animal phyla due to robust universal primers for COI

DNA Barcodes Invariable Across Life Stages, Processed Parts, Genders

What's in a Name?

Species Identification Matters

- Basic research on evolution, ecology
- Endangered/protected species
- Agricultural pests/beneficial species
- Disease vectors/pathogens
- Invasive species (e.g., in ballast water)
- Environmental quality indicators
- Managing for sustainable harvesting
- Consumer protection, ensuring food quality

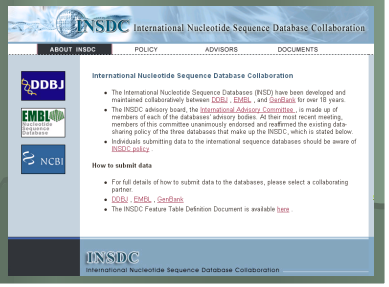
Creating the Reference Library

Identified Museum Specimens

Extract and Amplify DNA

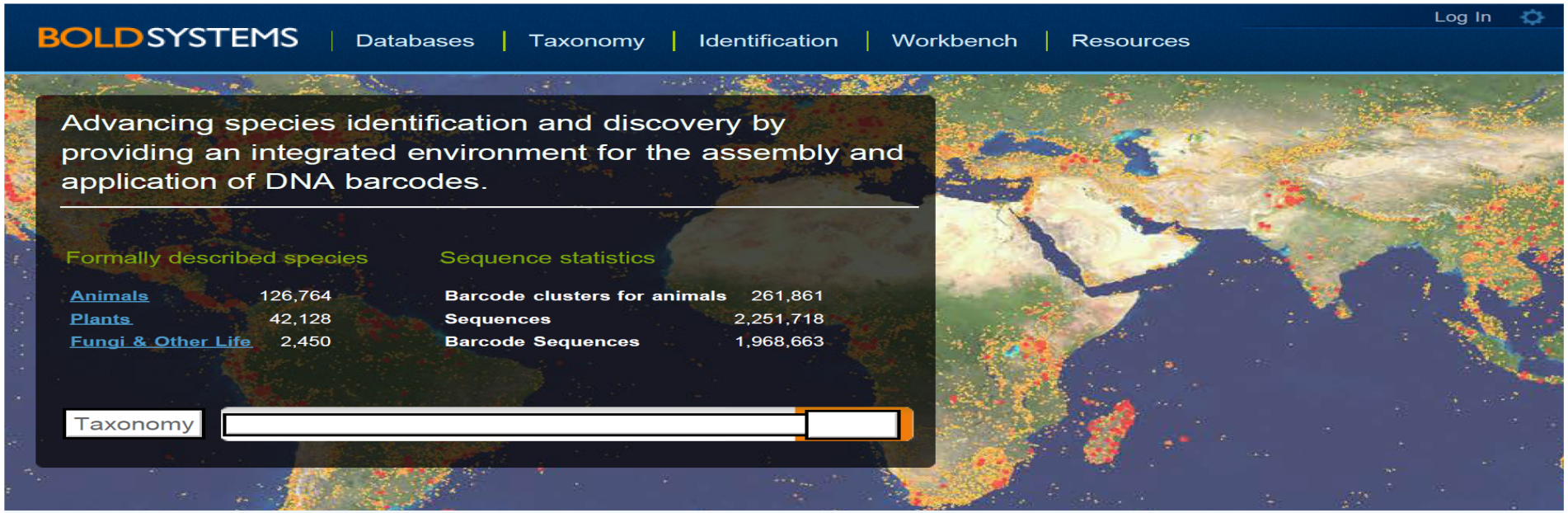
Sequence DNA

Publish Reference BARCODE Record



Required Elements for BARCODE

- Taxonomic identification to species
- Voucher specimen ID in standard format
- Name of barcode region (gene)
- Country/Ocean/Sea of origin
- Length, quality, 2 trace files
- Forward/reverse primer sequences, names



The Barcode of Life Database is designed to support the generation and application of DNA Barcode data. The platform consists of three main modules, a data portal, a database of barcode clusters, and data collection workbench. The Public Data Portal provides access to all public barcode data which consists of data generated using the Workbench module as well as data mined from other sources. The Barcode Index Number (BIN) system assigns a unique identifier to each sequence cluster of COI, providing an interim taxonomic system for species in the animal kingdom. The workbench module integrates secure databases with analytical tools to provide a private collaborative environment for researchers to collect, analyze, and publish barcode data and ancillary DNA sequences. This platform also provides an annotation framework that supports tagging and commenting on records and their components (i.e. taxonomy, images, and sequences), allowing for community-based validation of barcode data.

Public Data Portal

A data retrieval interface that allows for searching all 1.3M public records in BOLD using multiple search criteria including, but not limited to, geography, taxonomy, and depository. Search results can be summarized, plotted on high-res maps, and downloaded.

BIN Database

A searchable database of Barcode Index Numbers (BINs), sequence clusters that closely approximate species. This system allows for rapid validation and use of barcode data where taxonomic data are lacking or unverified.

Workbench

An integrated data collection and analysis environment that supports the assembly and validation of DNA barcodes and ancillary sequences. This application provides a secure storage environment for molecular research data but also allows for sharing with other users.

international BARCODE OF LIFE

Ontario
Ministry of Economic Development and Innovation

BOLDSYSTEMS | Databases | Taxonomy | Identification | Workbench | Resources

Haile Yancy

Project Console - FDA RFLP Dirty Twenty Two [DIRTT]

Print

PROJECT DATA VIEWS

View All Records
Record Search
User Console

PROJECT OPTIONS

List All Projects
Modify Project Properties
Register Primers
View All Primers

PUBLICATION

Submit to GenBank
Bibliography Submission
Project Summary - Specimens, Localities, and GenBank

UPLOADS

Specimen Data
Specimen Images
Sequences
Trace Files

DOWNLOADS

Sequences
Data Spreadsheets
Specimen Labels
Trace Files

SEQUENCE ANALYSIS

Taxon ID Tree
Distance Summary
Sequence Composition
Barcode Gap Analysis
Alignment Browser
Accumulation Curve
BIN Discordance Report
Diagnostic Characters

SPECIMEN AGGREGATES

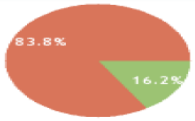
Distribution Map
Image Library

Progress

	Specimens (% complete)	Species (% complete)
COI-5P	73 / 99 (73.7%)	22 / 22 (100%)

Data Summary

BINs: 23
Countries: United States(95) Canada(4)
Taxonomy Breakdown:



- Insecta (class) (83)
- Mammalia (class) (16)

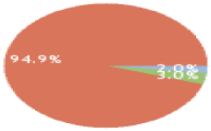
Sequence Data Report

Seqs lacking successful traces:	4	COI-5P[4]
Seqs with stop codons:	0	
Contaminated seqs:	0	
Problematic records flagged:	0	

Specimen Data Report

Specimen records:	99
Lacking geo reference:	98
Lacking country:	0
Lacking photographs:	31
Non-barcode compliant:	30

Specimen Depositories:



- U. S. Food and Drug Administration, Center for Veterinary Medicine (94)
- Bishop's University (3)
- Biodiversity Institute of Ontario (2)

Project Access

Manager	
Haile Yancy	hyancy@cvm.fda.gov
Users	
Yolanda L. Jones	yolanda.jones@fda.hhs.gov
Peter E. Miller	pemiller@uoguelph.ca
Nataly Ivanova	nivanova@uoguelph.ca
Chris D. Weland	cweland@uoguelph.ca

Project is publicly visible

Project Details

Title: FDA RFLP Dirty Twenty Two
Code: DIRTT
Description: Creating RFLP Fingerprints
Marker(s): COI-5P
Bounding Box Coordinates: Upper Left: N/A Lower Right: N/A

Most Recent Activities:

Show 25 entries Search:

Timestamp	Who	Action
Jun-20, 2012 15:15	Chris D. Weland	New-Images (3)
Jun-18, 2012 14:09	Chris D. Weland	Modify-Specimens (1)
Jun-18, 2012 14:09	Chris D. Weland	Modify-Specimens (1)
Jun-18, 2012 14:08	Chris D. Weland	Modify-Specimens (1)
Jun-18, 2012 14:07	Chris D. Weland	Modify-Specimens (1)
Jun-18, 2012 14:07	Chris D. Weland	Modify-Specimens (1)
Jun-18, 2012 14:07	Chris D. Weland	Modify-Specimens (1)
Jun-18, 2012 11:39	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 11:26	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 11:25	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 11:23	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 11:23	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 11:22	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 10:53	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 10:51	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 10:51	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 10:50	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 10:47	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 10:46	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 10:45	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 10:42	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 10:37	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 10:34	Yolanda L. Jones	Modify-Specimens (1)
Jun-18, 2012 10:19	Yolanda L. Jones	Modify-Specimens (1)
Jun-14, 2012 23:36	Nataly Ivanova	New-Traces (32)

Showing 1 to 25 of 48 entries

First Previous 1 2 Next Last

Download Activity Report: Last Week | Last Month | Last 6 Months

Quality Stats

Sequence Quality Stats

	High (<1% Ns)	Medium (<2% Ns)	Low (<4% Ns)	Unreliable (>4% Ns)
COI-5P	100	0	0	0

Trace Quality Stats

	High(%)	Medium(%)	Low(%)	Failed(%)	Total
COI-5P	53.28	8.2	6.56	31.96	366

Sequence Length Distribution

Marker - COI-5P

BOLDSYSTEMS | Databases | Taxonomy | Identification | Workbench | Resources

Haile Yancy

Record List - FDA RFLP Dirty Twenty Two [DIRTT]

Print

Project Console
User Console

Options
List All Projects
Move Records to another Project
Bibliography Submission
Summary- Specimens, Localities, and GenBank
Add Records to Dataset

Downloads
Sequences
Data Spreadsheet
Specimen Labels
Trace Files

Sequence Analysis
Taxon ID Tree
Distance Summary
Sequence Composition
Barcode Gap Analysis
Alignment Browser
Accumulation Curve
BIN Discordance Report
BIN Breakdown Report
Diagnostic Characters

Specimen Aggregates
Distribution Map
Image Library

Record Flags Legend
GPS coordinates present for the record.
Record contains images of the specimen.
The number of traces present for the sample.
Sequence is Barcode Compliant.
Stop codons present in sequence.
Contamination present in sequence.
Flagged record, filtered from the ID engine.

Add Tags & Comments

Showing Records 1 to 99 | Page 1 | Records Per Page 100

Project Data : Select | COI-5P: 73 | Specimens: 99

Identification	Specimen Page	Sequence Page	Length [Ambig] COI-5P	Record Flags	Extra Info	Tags	BIN
Blatta orientalis	Oriental cockroach 3	DIRTT031-11	626 [0n]				BOLD:AAG9954
Blatta orientalis	Oriental cockroach 1	DIRTT029-11	658 [0n]				BOLD:AAG9954
Blatta orientalis	Oriental cockroach 2	DIRTT030-11	658 [0n]				BOLD:AAG9954
Blatta orientalis	Oriental cockroach	DIRTT004-10	658 [0n]				BOLD:AAG9954
Blattella germanica	German cockroach 1	DIRTT023-11	238 [0n]				
Blattella germanica	German Cockroach	DIRTT002-10	0				
Blattella germanica	German cockroach 3	DIRTT025-11	238 [0n]				
Blattella germanica	German cockroach 2	DIRTT024-11	238 [0n]				
Blattellidae	Brownbanded cockroach	DIRTT003-10	658 [0n]				BOLD:AAD2182
Blattellidae	Brownbanded cockroach 3	DIRTT028-11	0				
Blattellidae	Brownbanded cockroach 1	DIRTT026-11	658 [0n]				BOLD:AAD2182
Blattellidae	Brownbanded cockroach 2	DIRTT027-11	658 [0n]				BOLD:AAD2182
Calliphora vicina	Cosmopolitan blue bottle fly 3	DIRTT055-11	658 [0n]				BOLD:AAB6579
Calliphora vicina	Cosmopolitan blue bottle fly	DIRTT021-10	658 [0n]				BOLD:AAB6579
Calliphora vicina	Cosmopolitan blue bottle fly 2	DIRTT054-11	658 [0n]				BOLD:AAB6579
Calliphora vicina	Cosmopolitan blue bottle fly 1	DIRTT053-11	658 [0n]				BOLD:AAB6579
Calliphora vomitoria	Holarctic blue bottle fly	DIRTT022-10	0				
Calliphora vomitoria	Holarctic blue bottle fly 3	DIRTT058-11	566 [0n]				BOLD:AAA8931
Calliphora vomitoria	Holarctic blue bottle fly 2	DIRTT057-11	658 [0n]				BOLD:AAA8931
Calliphora vomitoria	Holarctic blue bottle fly 1	DIRTT056-11	658 [0n]				BOLD:AAA8931
Chrysomya megacephala	Oriental latrine fly 3	DIRTT061-11	658 [0n]				BOLD:AAA5667
Chrysomya megacephala	Oriental latrine fly	DIRTT011-10	0				
Chrysomya megacephala	Oriental latrine fly 1	DIRTT059-11	658 [0n]				BOLD:AAA5667
Chrysomya megacephala	Oriental latrine fly 2	DIRTT060-11	658 [0n]				BOLD:AAA5667
Cochliomyia macellaria	Secondary screwworm 3	DIRTT064-11	[0n]				
Cochliomyia macellaria	Secondary screwworm 1	DIRTT062-11	[0n]				
Cochliomyia macellaria	Secondary screwworm	DIRTT012-10	[0n]				
Cochliomyia macellaria	Seconday Screwworm Leg	DIRTT096-12	658 [0n]				BOLD:AAD7771
Cochliomyia macellaria	Secondary screwworm 2	DIRTT063-11	[0n]				
Cynomya cadaverina	Blue bottle fly	DIRTT013-10	0				
Cynomya cadaverina	Blue bottle fly 2	DIRTT066-11	658 [0n]				BOLD:AAB0868
Cynomya cadaverina	Blue bottle fly 3	DIRTT067-11	658 [0n]				BOLD:AAB0868
Cynomya cadaverina	Blue bottle fly 1	DIRTT065-11	658 [0n]				BOLD:AAB0868
Fannia canicularis	Little house fly 3	DIRTT049-11	658 [0n]				BOLD:AAF7101
Fannia canicularis	Little house fly 1	DIRTT047-11	658 [0n]				BOLD:AAF7101
Fannia canicularis	Little house fly	DIRTT009-10	658 [0n]				BOLD:AAY8764
Fannia canicularis	Little house fly 2	DIRTT048-11	658 [0n]				BOLD:AAF7101
Fannia scalaris	Latrine Fly 4	DIRTT097-12	658 [0n]				BOLD:ABX0715
Fannia scalaris	Latrine fly 1	DIRTT050-11	0				
Fannia scalaris	Latrine Fly 5	DIRTT098-12	658 [0n]				BOLD:ABX0715
Fannia scalaris	Latrine Fly 6	DIRTT099-12	0				
Fannia scalaris	Latrine fly 3	DIRTT052-11	[0n]				
Fannia scalaris	Latrine fly	DIRTT010-10	0				

Tools | Identification | Record Links | Details | BIN

BOLD SYSTEMS

Specimen - FDA RFLP Dirty Twenty Two [DIRTT]

Print

Edit Specimen

IDENTIFIERS

Sample ID:	Norway rat
Process ID:	DIRTT019-10
Institution Storing:	U. S. Food and Drug Administration, Center for Veterinary Medicine
Field ID:	021
Museum ID:	USNM 570865
Collection Code:	

TAXONOMY

Identification:	Rattus norvegicus (Berkenhout, 1769)
Rank:	Species
Identifier	
Identification Method:	
Identifier Institution:	
Identifier Email:	
Taxonomy Note:	

Rank	Current Record (Norway rat)	Others in BIN (BOLD:AAA8499)
Phylum:	Chordata	Chordata [80]
Class:	Mammalia	Mammalia [80]
Order:	Rodentia	Rodentia [80]
Family:	Muridae	Muridae [80]
Subfamily:	Murinae	Murinae [80]
Genus:	Rattus	Rattus [80]
Species:	Rattus norvegicus	Rattus norvegicus [80]

SPECIMEN DETAILS

Voucher Status:	DNA/tissue vouchered only
Tissue Descriptor:	
Sex:	
Reproduction:	
Life Stage:	
Extra Info:	
Note:	
Associated Taxa:	
Associated Specimens:	
Reference Link:	

ANNOTATION

Add Tags & Comments Comments: 0 Associated Tags: No Tags

PHOTOGRAPHS

License: Unspecified default All Rights Reserved
License Holder: Unspecified

Add Tags & Comments Comments: 0 Associated Tags: No Tags

Dorsal

GEOGRAPHY

Country:	United States
Province/State:	
Region/County:	
Sector:	
Exact Site:	
Lat/Lon:	

COLLECTION DETAILS

Collectors:	
Collection Event ID:	
Date Collected:	
Date Accuracy:	
Time Collected:	
Site Code:	
Habitat:	
Sampling Protocol:	
Coord. Source:	
Coord. Accuracy:	
Elevation:	
Elevation Accuracy:	
Depth:	
Depth Accuracy:	
Collection Notes:	

Public Data Portal - BIN Page

Barcode Index Number Registry For BOLD:AAA8499

Back to Last Page

Go to public records in this BIN

BIN DETAILS:

BIN URI:	BOLD:AAA8499	Average Distance:	0.37% (p-dist)
DOI:		Maximum Distance:	1.73% (p-dist)
Member Count:	86 [65 Public]	Distance to Nearest Neighbor:	8.61% (p-dist)
Founding Record:			

NEAREST NEIGHBOR (NN) DETAILS:

Nearest BIN URI:	BOLD:AAE7376	Average Distance:	0.79% (p-dist)
Member Count:	6	Maximum Distance:	1.61% (p-dist)
Nearest Member:	GBMA2187-09	Distance Variance:	0.44% (p-dist)
Nearest Member Taxonomy:	Chordata, Mammalia, Rodentia, Muridae, Murinae, Rattus, Rattus leucopus		

TAXONOMY:

Phylum:	Chordata [81]
Class:	Mammalia [81]
Order:	Rodentia [81]
Family:	Muridae [81]
Subfamily:	Murinae [81]
Genus:	Rattus [81]
Species:	Rattus norvegicus [81]

Add Tags & Comments Comments: 0 Associated Tags: No Tags

DISTANCE DISTRIBUTION:

PUBLICATIONS:

Kumarasamy,S,Sivarajan, Gopalakrishnan,K,Kathirvel, Shafton,A,Asher, Nixon,J,Jeremy, Thangavel,J,Jayakumar, Farms,P,Phyllis, Joe,B,Bina Mitochondrial polymorphisms in rat genetic models of hypertension. Mamm. Genome 2010-06-01;21(-1):299-306 (PDF)

Paulo B. Chaves, Vanessa G. Graeff, Mari Lia B. Lion, Larissa R. Oliveira and Eduardo Eizrik DNAbarcoding meets molecular scatology: short mtDNA sequences for standardized species assignment of carnivore noninvasive samples Molecular Ecology Resources 2011-08-30;11(5) (PDF)

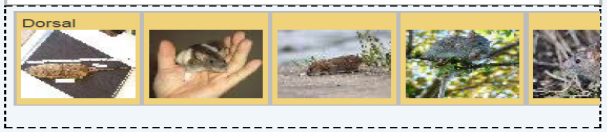
Min, X. J., & Hickey, D. A. DNA barcodes provide a quick preview of mitochondrial genome composition. PLoS ONE 2007-03-28;2(e325) (PDF)

Schlick,NE,Nancy E, Jensen-Seaman,MI,Michael I, Orlebeke,K,Kimberly, Kwitek,AE,Anne E, Jacob,HJ,Howard J, Lazar,J,Jozef Sequence analysis of the complete mitochondrial DNA in 10 commonly used inbred rat strains. Am. J. Physiol., Cell Physiol. 2006-12-01;291(6):C1183-92 (PDF)

Specimen Images:

GBMA0500-06 {Rattus norvegicus}
License: CC BYNC 30 (2012)
License Holder: Biopix: JC Schou, Biopix

Add Tags & Comments Comments: 0 Associated Tags: No Tags



Collection Sites:

Add Tags & Comments Comments: 0 Associated Tags: No Tags

BOLDSYSTEMS

Sequence - FDA RFLP Dirty Twenty Two [DIRTT]

Print

IDENTIFIERS

Sample ID:	Norway rat
Process ID:	DIRTT019-10
Identification:	Rattus norvegicus

COI-5P

SEQUENCE DATA

Genbank Accession:	
Translation Matrix:	Vertebrate Mitochondrial
Last Updated:	2011-08-12

Clear Sequence

NUCLEOTIDE SEQUENCE

Sequence: *658 bp*

```
AACCCTCTACCTATTATTTGGAGCCTGAGCAGGAATAGTAGGGACAGCTTTAAGTATTCT
AATTCGAGCTGAACTAGGACAGCCAGGCGCACTCCTAGGAGATGACCAAATCTATAATGT
CATCGTCACAGCCCATGCATTCGTAATAATTTTCTTTATAGTAATACCTATAATAATTGG
AGGCTTCGGGAACTGACTTGTACCACTAATAATTGGAGCCCCTGATATAGCATTCCCACG
AATAAATAACATAAGCTTTTGACTGCTTCCTCCATCATTTCTACTCCTTTTAGCATCCTC
CATAGTAGAAGCTGGAGCTGGAACAGGATGAACAGTATATCCCCCCTTAGCCGGAAACCT
AGCCCATGCTGGAGCATCCGTAGATTTAACTATTTTTCCCTCCACCTAGCCGGGGTGTC
TTCTATCTTAGGAGCTATCAACTTTATCACCACTATCATTAATATAAAACCCCCTGCTAT
AACCCAATATCAGACACCTCTCTTTGTATGATCCGTACTAATTACAGCCGTCCTACTACT
TCTCTCACTGCCAGTATTAGCAGCAGGTATCACTATACTCCTTACAGACCGAAATCTAAA
TACTACTTTCTTCGACCCCGCTGGAGGTGGAGACCCAATCCTTTATCAACACCTATTC
```

Composition:	A (186), G (106), C (175), T (191)
Ambiguous Characters:	0

Identify Sequence Using:

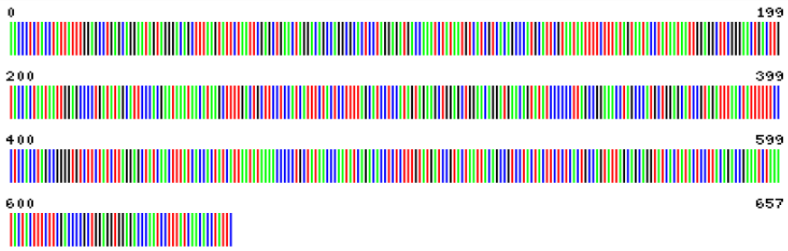
Full DB | Species DB | Published DB | Full Length DB

AMINO ACID SEQUENCE

Sequence: *233 residues*

```
TLYLLFGAWAGMVGTALSILIRAELGQPGALLGDDQIYNVIVTAHAFVMIFFMVMPMMIG
GFGNWLVPLMIGAPDMAFPRMNNMSFWLLPPSFLLLLASSMVEAGAGTGWTVYPPLAGNL
AHAGASVDLTIFSLHLAGVSSILGAINFITTIINMKPPAMTQYQTPLFVWSVLITAVLLL
LSLPVLAAGITMLLTDRNLNTTFFDPAGGGDPILYQHLF
```

ILLUSTRATIVE BARCODE



SEQUENCING RUNS: Biodiversity Institute of Ontario

Run Date	Direction	Trace File	Seq Primer	Quality
PCR Primers: C_VF1LFt1/C_VR1LRt1				
2011-08-11	Reverse	DIRTT019-10[C_VF1LFt1,C_VR1LRt1]_R.ab1	M13R	high
2011-08-11	Forward	DIRTT019-10[C_VF1LFt1,C_VR1LRt1]_F.ab1	M13F	high
PCR Primers: LepF1/LepR1				
2011-08-18	Forward	DIRTT019-10[LepF1,LepR1]_F.ab1	LepF1	fail
2011-08-18	Reverse	DIRTT019-10[LepF1,LepR1]_R.ab1	LepR1	fail

View Trace Files | Download

ANNOTATION

Add Tags & Comments | Comments: 0 | Associated Tags: No Tags

SAMPLE REPORT FROM LIMS

BOLDLIMS Report Available.

Identification of a specimen

BOLDSYSTEMS

Sequence - FDA RFLP Dirty Twenty Two [DIRTT]

Print

IDENTIFIERS

Sample ID:	Norway rat
Process ID:	DIRTT019-10
Identification:	Rattus norvegicus

COI-5P

SEQUENCE DATA

Genbank Accession:	
Translation Matrix:	Vertebrate Mitochondrial
Last Updated:	2011-08-12

Clear Sequence

NUCLEOTIDE SEQUENCE

Sequence: *658 bp*

```
AACCCTCTACCTATTATTTGGAGCCTGAGCAGGAATAGTAGGGACAGCTTTAAGTATTCT
AATTCGAGCTGAACTAGGACAGCCAGGCGCACTCCTAGGAGATGACCAAATCTATAATGT
CATCGTCACAGCCCATGCATTCGTAATAATTTTCTTTATAGTAATACCTATAATAATTGG
AGGCTTCGGGAACTGACTTGTACCACTAATAATTGGAGCCCCTGATATAGCATTCCCACG
AATAAATAACATAAGCTTTTGACTGCTTCCTCCATCATTTCTACTCCTTTTAGCATCCTC
CATAGTAGAAGCTGGAGCTGGAACAGGATGAACAGTATATCCCCCCTTAGCCGGAAACCT
AGCCCATGCTGGAGCATCCGTAGATTTAACTATTTTTCCCTCCACCTAGCCGGGGTGTC
TTCTATCTTAGGAGCTATCAACTTTATCACCACTATCATTAATATAAAACCCCCTGCTAT
AACCCAATATCAGACACCTCTCTTTGTATGATCCGTACTAATTACAGCCGTCCTACTACT
TCTCTCACTGCCAGTATTAGCAGCAGGTATCACTATACTCCTTACAGACCGAAATCTAAA
TACTACTTTCTTCGACCCCGCTGGAGGTGGAGACCCAATCCTTTATCAACACCTATTC
```

Composition:	A (186), G (106), C (175), T (191)
Ambiguous Characters:	0

Identify Sequence Using:

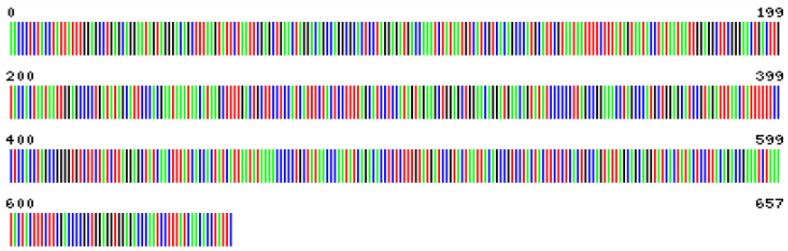
Full DB | Species DB | Published DB | Full Length DB

AMINO ACID SEQUENCE

Sequence: *233 residues*

```
TLYLLFGAWAGMVGTALSILIRAELGQPGALLGDDQIYNVIVTAHAFVMIFFMVMPMMIG
GFGNWLVPLMIGAPDMAFPRMNNMSFWLLPPSFLLLLASSMVEAGAGTGWTVYPPLAGNL
AHAGASVDLTIFSLHLAGVSSILGAINFITTIINMKPPAMTQYQTPLFVWSVLITAVLLL
LSLPVLAAGITMLLTDRNLNTTFFDPAGGGDPILYQHLF
```

ILLUSTRATIVE BARCODE



SEQUENCING RUNS: Biodiversity Institute of Ontario

Run Date	Direction	Trace File	Seq Primer	Quality
PCR Primers: C_VF1LFt1/C_VR1LRt1				
2011-08-11	Reverse	DIRTT019-10[C_VF1LFt1,C_VR1LRt1]_R.ab1	M13R	high
2011-08-11	Forward	DIRTT019-10[C_VF1LFt1,C_VR1LRt1]_F.ab1	M13F	high
PCR Primers: LepF1/LepR1				
2011-08-18	Forward	DIRTT019-10[LepF1,LepR1]_F.ab1	LepF1	fail
2011-08-18	Reverse	DIRTT019-10[LepF1,LepR1]_R.ab1	LepR1	fail

View Trace Files | Download

ANNOTATION

Add Tags & Comments | Comments: 0 | Associated Tags: No Tags

SAMPLE REPORT FROM LIMS

BOLDLIMS Report Available.

BOLDSYSTEMS | Databases | Taxonomy | Identification | Workbench | Resources

Haile Yancy

Specimen Identification Request

Print

Query: unlabeled_sequence

Top Hit: Chordata - Rodentia - Rattus norvegicus (100%)

Search Result:

The submitted sequence has been matched to *Rattus norvegicus*. This identification is solid unless there is a very closely allied congeneric species that has not yet been analyzed. Such cases are rare.

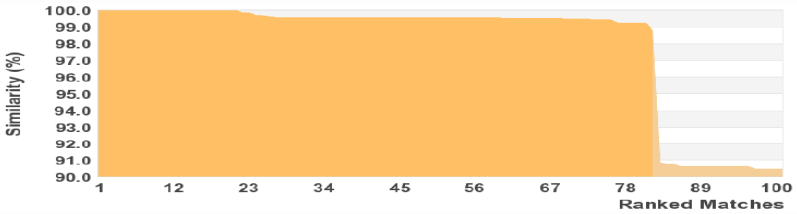
A species page and a BIN page are available for this taxon: Species Page | Top Match BIN Page

For a heirarchical placement - a neighbor-joining tree is provided: Tree Based Identification

Identification Summary:

Taxonomic Level	Taxon Assignment	Probability of Placement (%)
Phylum	Chordata	100
Class	Mammalia	100
Order	Rodentia	100
Family	Muridae	100
Genus	Rattus	100
Species	Rattus norvegicus	100

Similarity Scores of Top 100 Matches:



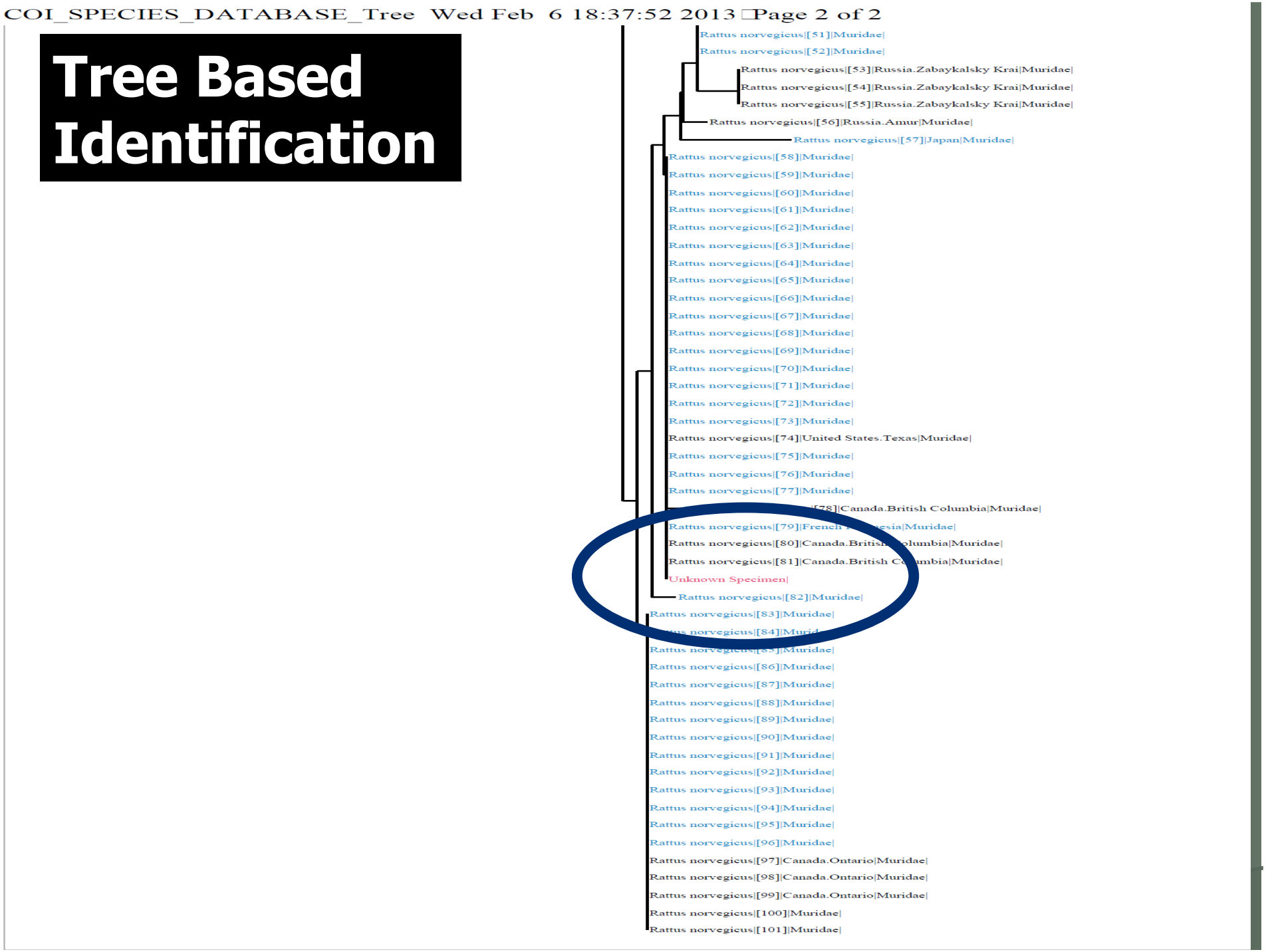
TOP 20 Matches :

Display option: Top 20

Phylum	Class	Order	Family	Genus	Species	Similarity (%)	Status
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Published
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Early-Release
Chordata	Mammalia	Rodentia	Muridae	Rattus	*norvegicus*	100	Early-Release

Sampling Sites For Top Hits (>98% Match):

Tree Based Identification



Identifying Evidence

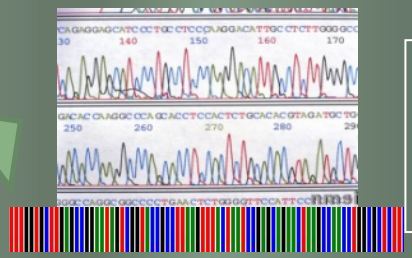
Identified Museum Specimens

Extract and Amplify DNA

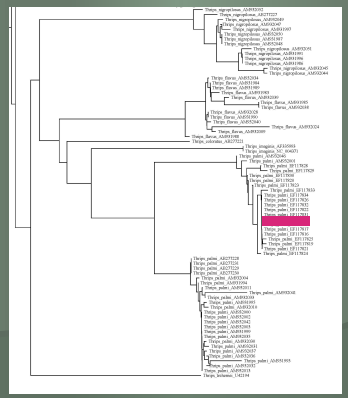
Sequence DNA

Publish Reference BARCODE Record

Confiscated Material



Unidentified DNA Barcode Sequence



Species Identification

Current FDA Barcoding Projects

- Fish
- Dirty 22 (insects and rodents)
- Plants
- Bushmeat

"Dirty 22" Pests

German cockroach	*E. coli, Salmonella spp., Shigella spp., Staphylococcus spp., Entamoeba histolytica*
Brownbanded cockroach	*E. coli, Salmonella spp., Shigella spp., Staphylococcus spp.*
Oriental cockroach	*E. coli, Salmonella spp., Shigella spp., Staphylococcus spp.*
American cockroach	*E. coli, Salmonella spp., Shigella spp., Staphylococcus spp., Entamoeba histolytica*
Pharaoh ant	*E. coli, Salmonella spp., Shigella spp., Staphylococcus spp.*
Thief ant	*E. coli, Salmonella spp., Shigella spp., Staphylococcus spp.*
House fly	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
Stable fly	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
Little house fly	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
Latrine fly	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
Cosmopolitan blue bottle fly	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
Holarctic blue bottle fly	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
Oriental latrine fly	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
Secondary screwworm	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
Blue bottle fly	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
Green bottle fly	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
Black blow fly	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
Redtailed flesh fly	*Shigella spp., Salmonella spp., E. coli, Campylobacter jejuni, Vibrio cholera, Cryptosporidium*
House mouse	*Salmonella*
Polynesian rat	*Salmonella*
Norway rat	*Salmonella*
Roof rat	*Salmonella*

FDA field labs have found the "Dirty 22," the 22 most common pests that FDA recognizes in contributing to the spread of food-borne pathogens, represent the highest risk under the Compliance Policy Guide (CPG), CPG 7120.18. Currently the labs are using microscopy.

Bushmeat

Prevent the Spread of Infectious Disease by the Control of Illegal Importation of Bushmeat

Interagency working group (FDA,USDA,CDC, US Customs, Fish & Wildlife)

Building regulatory database containing ~1,000 species identified as bushmeat and ~ 2,000 closely related species and look-alikes that represent potential false-positive identifications

Global Impact Award Project

- Consortium for the Barcode of Life (CBOL) received $3 million grant from Google Foundation to barcode endangered species
- Catalyze and support the use of DNA Barcode evidence in prosecutions in six developing countries
- Provide training for:
  - Researchers, technicians
  - Inspectors, investigators, prosecutors, judges
- Build the Barcode Reference Library for:
  - ~2,000 CITES-listed species and
  - ~8,000 close relatives and look-alikes

Assay Development Using Barcode Database

Difficulty in using traditional barcoding for identification of species food and feed

- Samples contain multiple sources of DNA including multiple animals as well as filth such as feathers, feces, and insects
- Samples are degraded (cooked, processed, or neglected)

PCR product ~650 bp

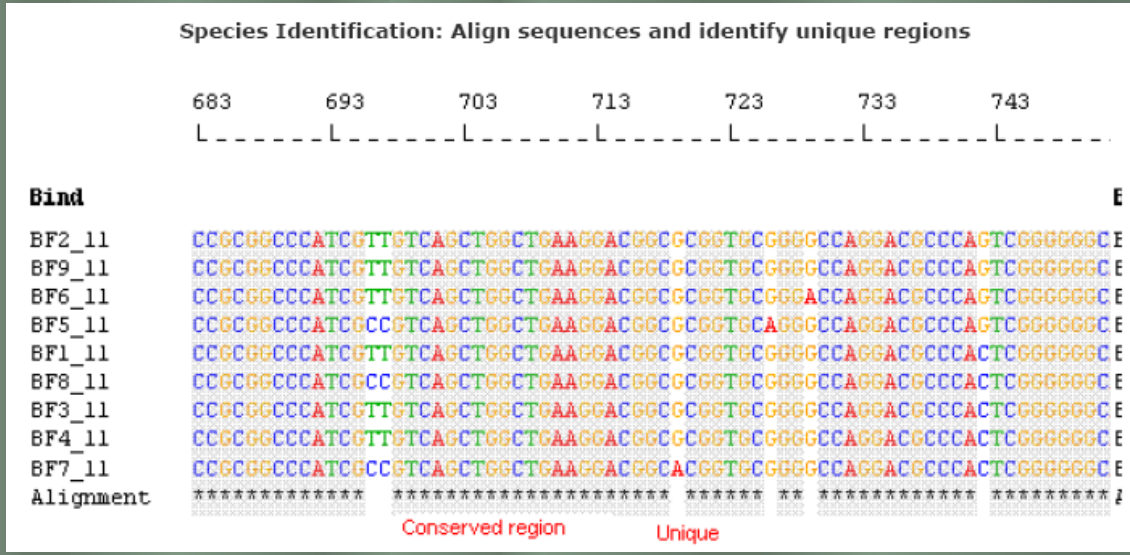
PCR product ~350-250 bp

PCR product ~250- 50 bp

Assay Development

- Microarray Assay
  - 40-80k species in single assay
  - Ability to analyze cooked, processed, or mixed samples
- Case Study
- Real- Time PCR Assay
  - Develop assay in 2-3 days
  - 3-4 species in single assay
  - Ability to analyze cooked, processed, or mixed samples
  - Results in ~1 hour

Allele ID

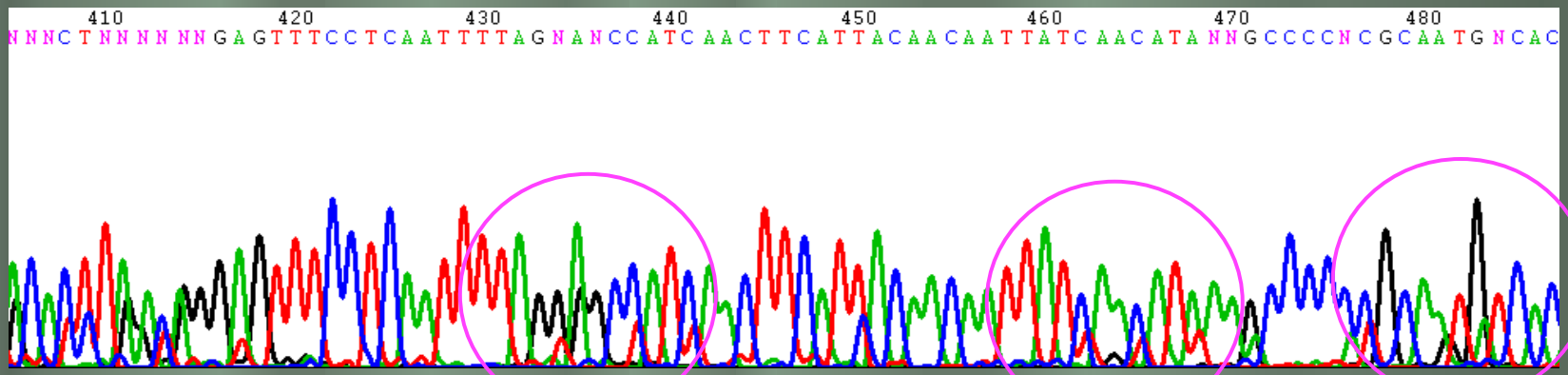


- Allele ID aligns sequences of different species to identify unique regions
- Species specific primers and probes are designed based on differences in sequence
- Taxa specific primers and probes can also be designed

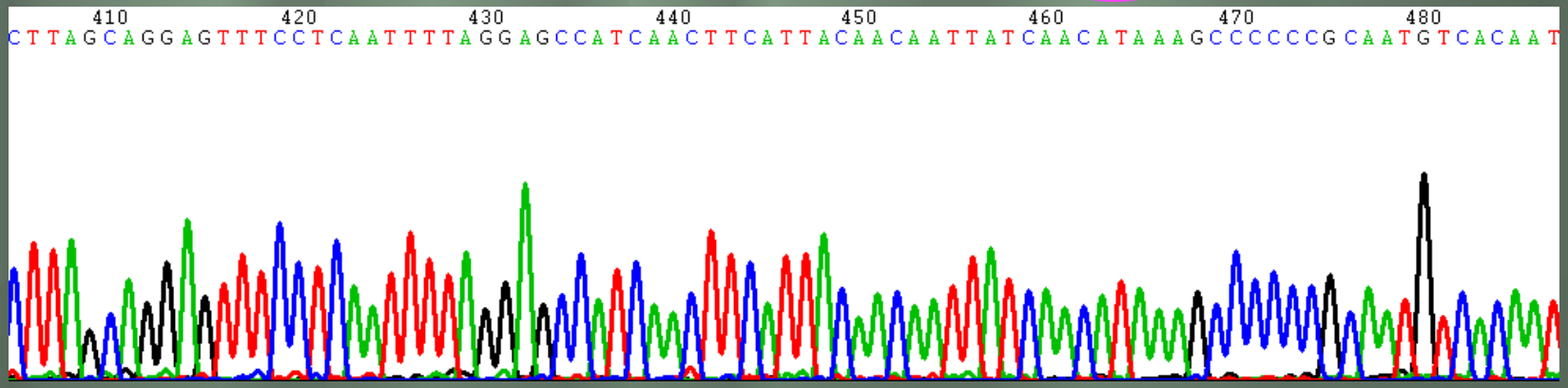
Microarray Assays for BSE "Mad Cow Disease"

- Creating probe assays based on barcode sequences

Cow and Pig

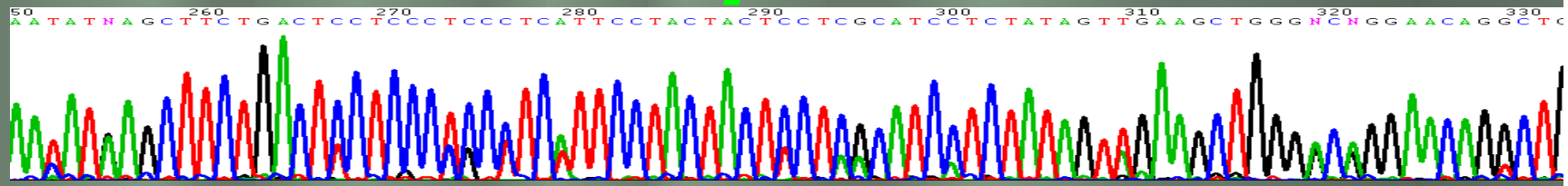


Cow

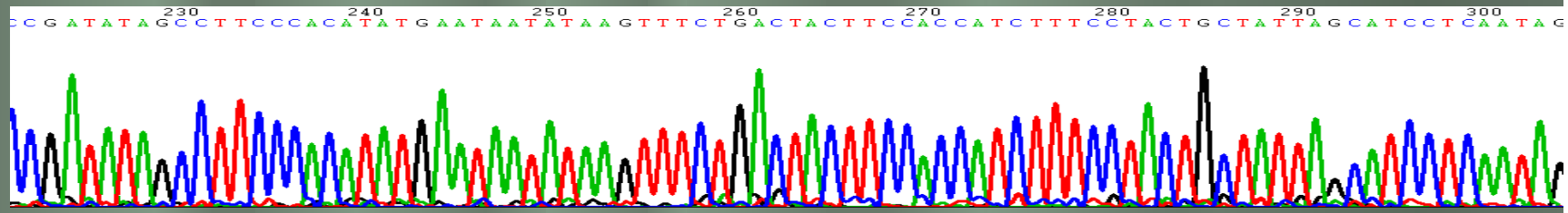


Microarray Process

Cow



Pig



Allele ID

Oligo (5 probes per species)

Create Zoo Chip

Unknown Sample

Label

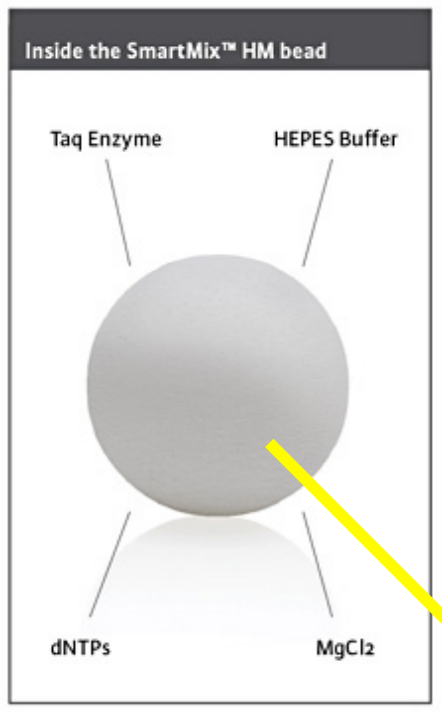
Alexa Fluor 647

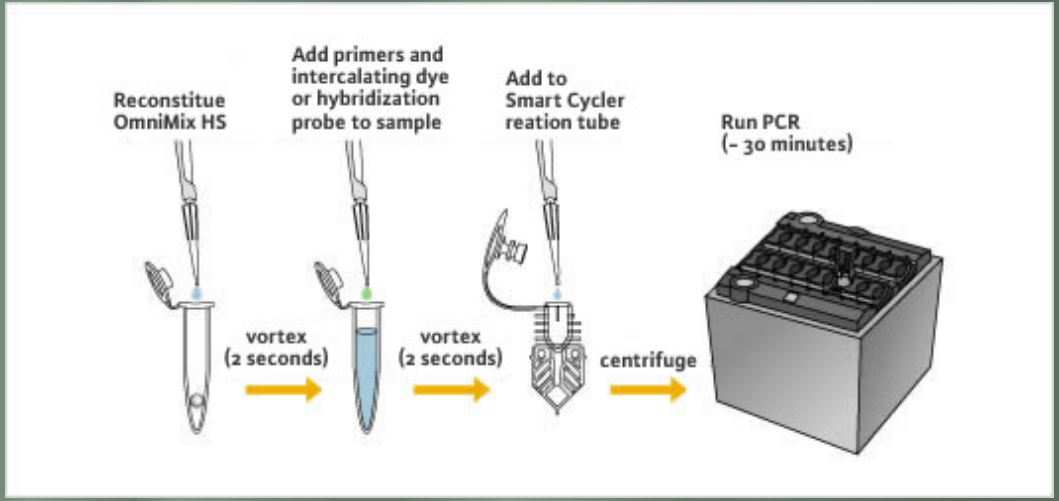
HYB

Identify Species

(Pig and Cow)

PCR Assay

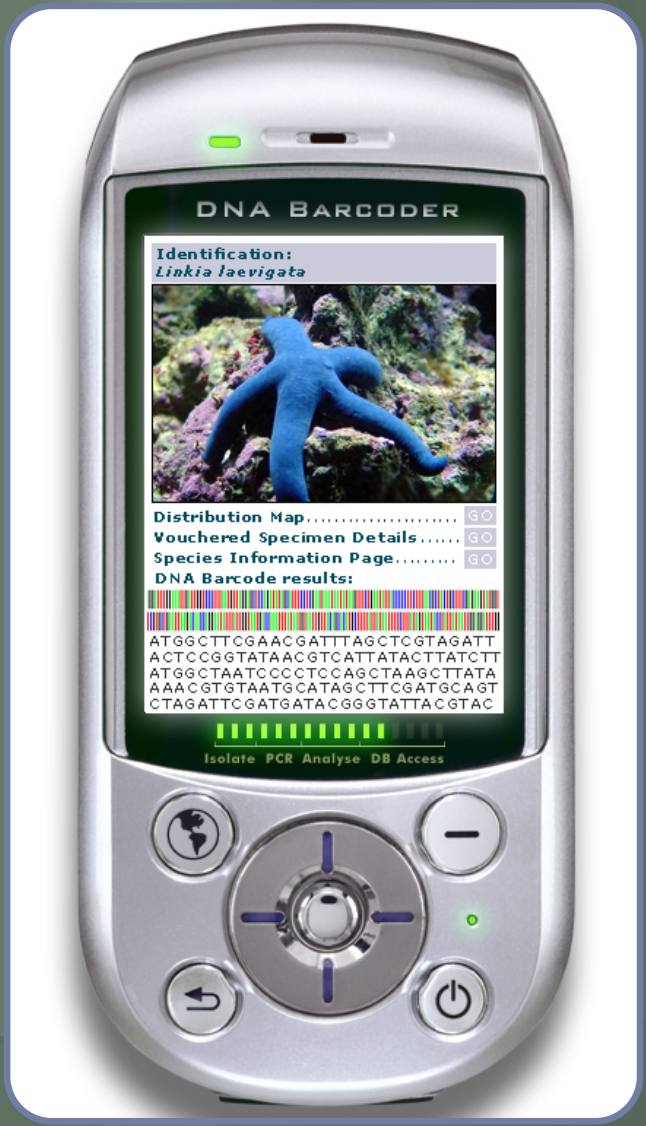




Species specific probes (4)

In the "not-so-distant" Future

CVM is collaborating with several private and government groups to develop a hand held device that has the ability to extract DNA, amplify DNA, and identify a species (using PCR, probes or sequence)!



PARTNERS IN FUNDING AND IN-KIND SUPPORT

GenomeCanada OntarioGenomicsInstitute Parks Canada Agriculture and Agri-Food Canada Fisheries and Oceans Canada ROM Royal Ontario Museum Canadian Museum of Nature Musée canadien de la nature

Canada Foundation for Innovation Fondation canadienne pour l'innovation Gordon and Betty Moore Foundation University of Guelph Environment Canada Natural Resources Canada NSERC CRSNG Alfred P. Sloan Foundation

Canadian Barcode of Life Network

RESEARCH PARTNERS

44 Researchers from 16 universities, 7 government departments, and 2 museums

Thank You