

CSL/JIFSAN Symposium on Food Safety and Nutrition:
Methods and Systems for Tracking, Tracing, and Verifying Foods

DNA Barcoding: Regulatory Applications at FDA

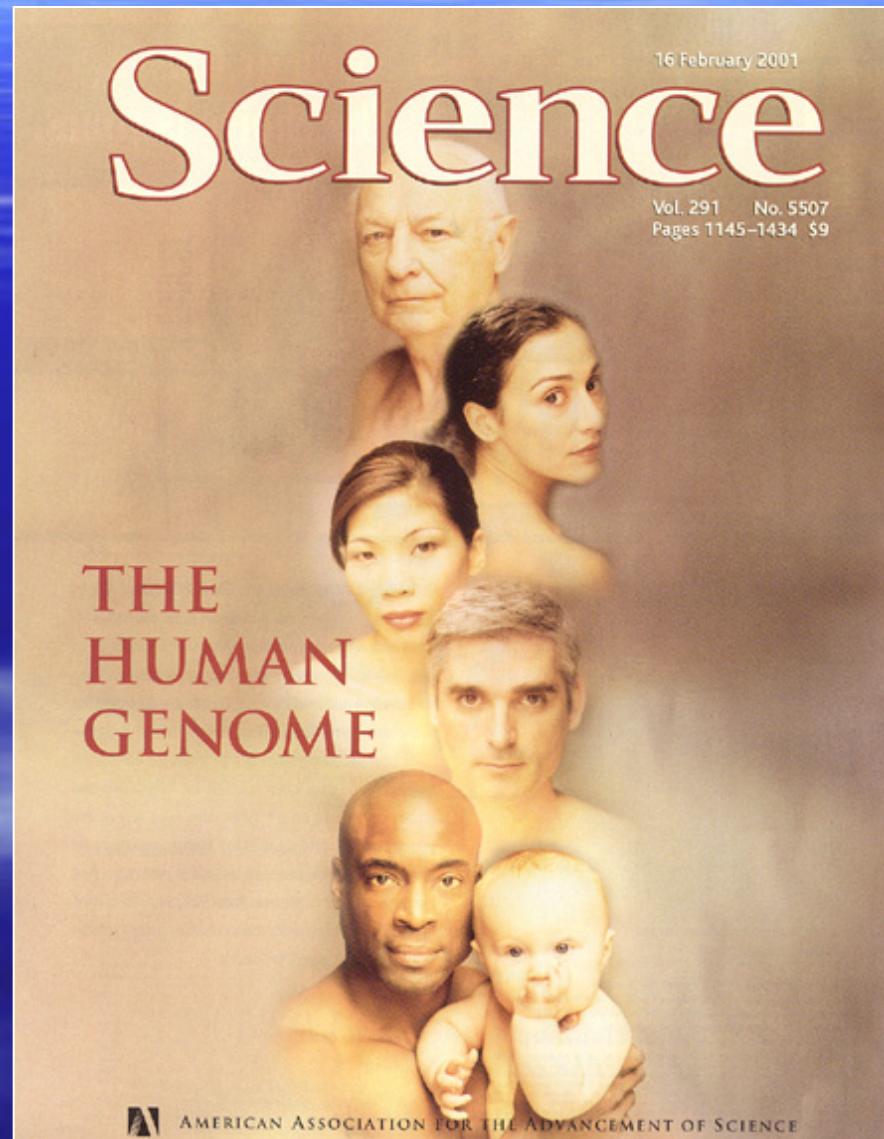
Haile Yancy, Ph.D.

US FDA Center for Veterinary Medicine

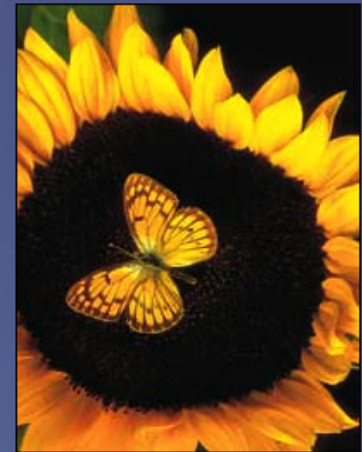


A DNA barcode is a short gene sequence used to identify species taken from a standard position in the genome

Conventional Genomics - All Genes, One Species



Horizontal Genomics - One Gene, All Species



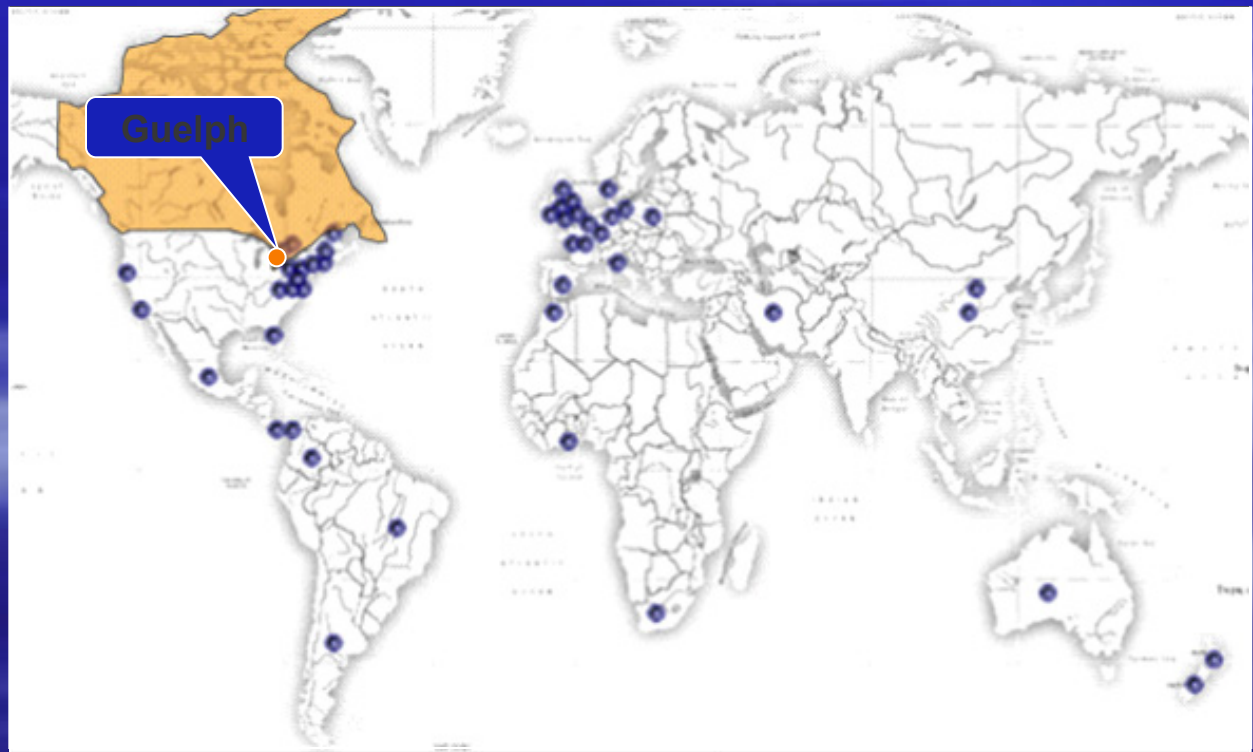
Barcoding Products and Life



CONSORTIUM FOR THE BARCODE OF LIFE

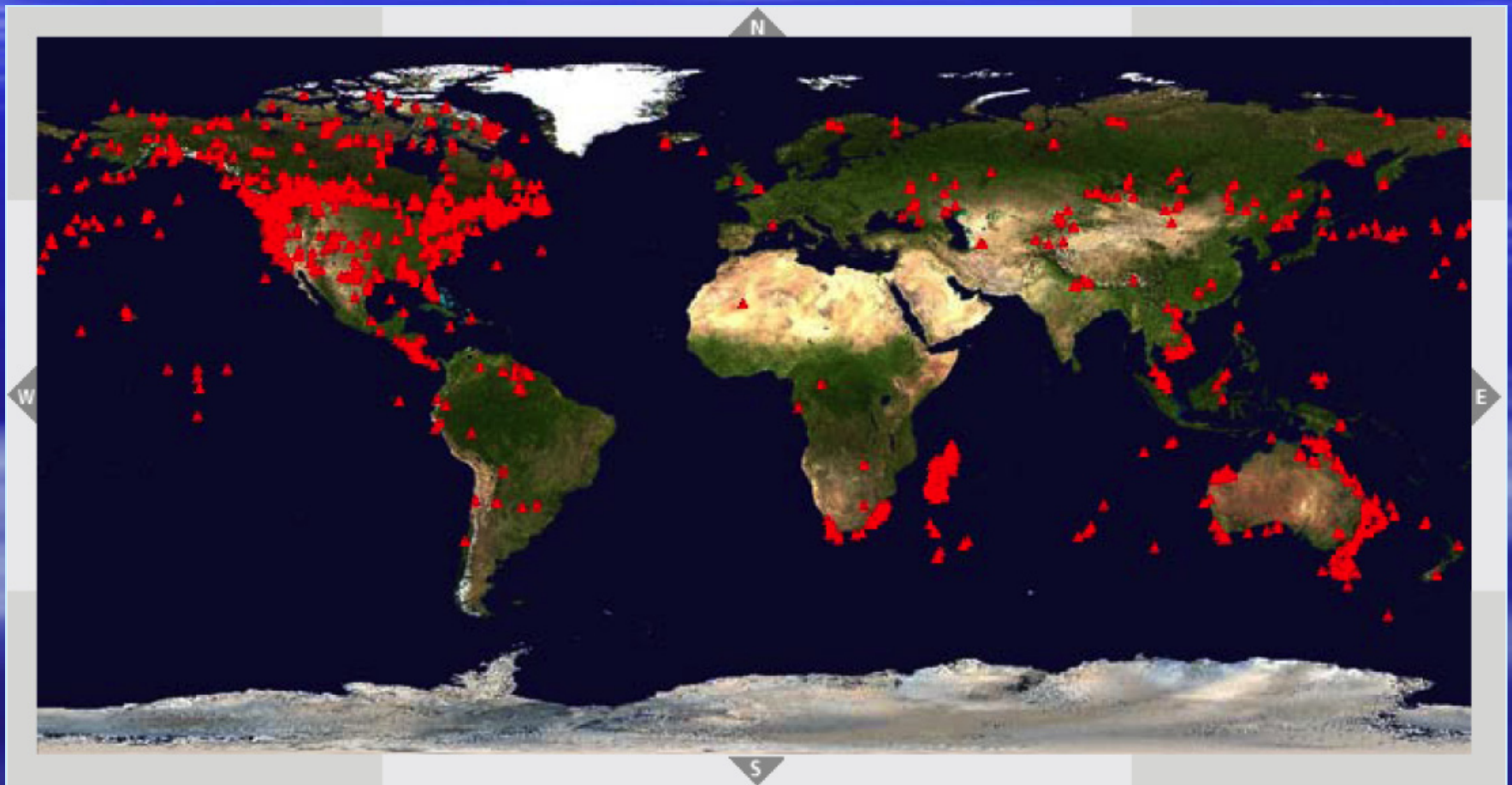


- CBOL launched April, 2004.
- Active memberships in 50 countries (and growing).
- Projects to barcode all birds and all fishes on Earth.
- Estimated cost to barcode all animal life: \$1 billion.

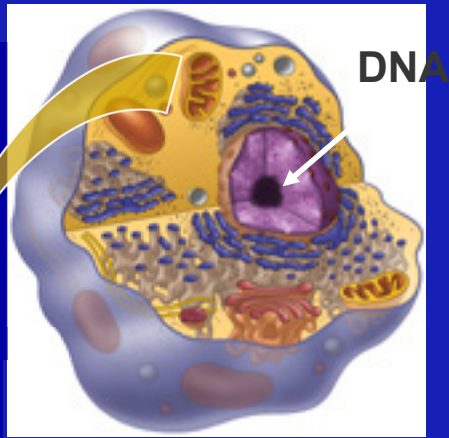


DNA Barcodes Today

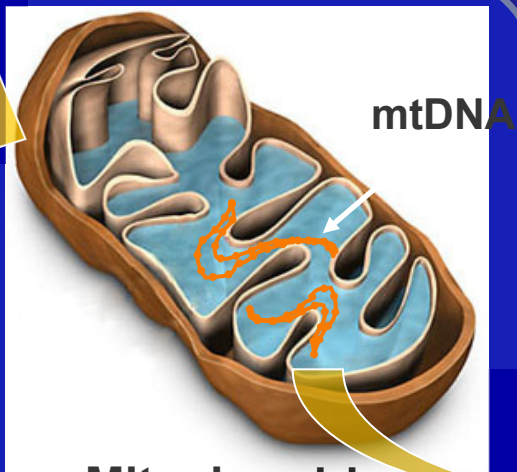
576K specimens, 55K species



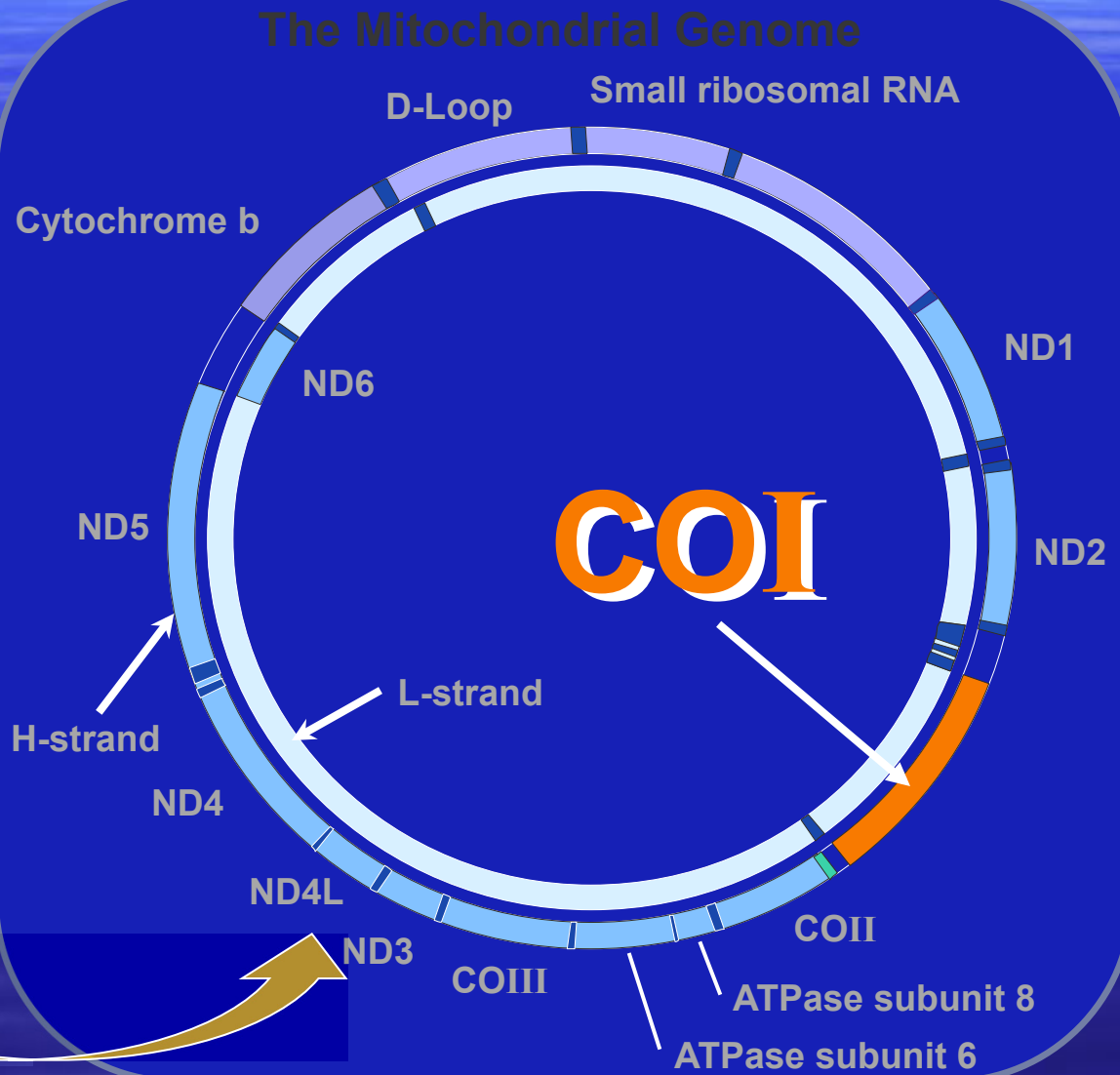
An Internal ID System for All Animals



Typical Animal Cell

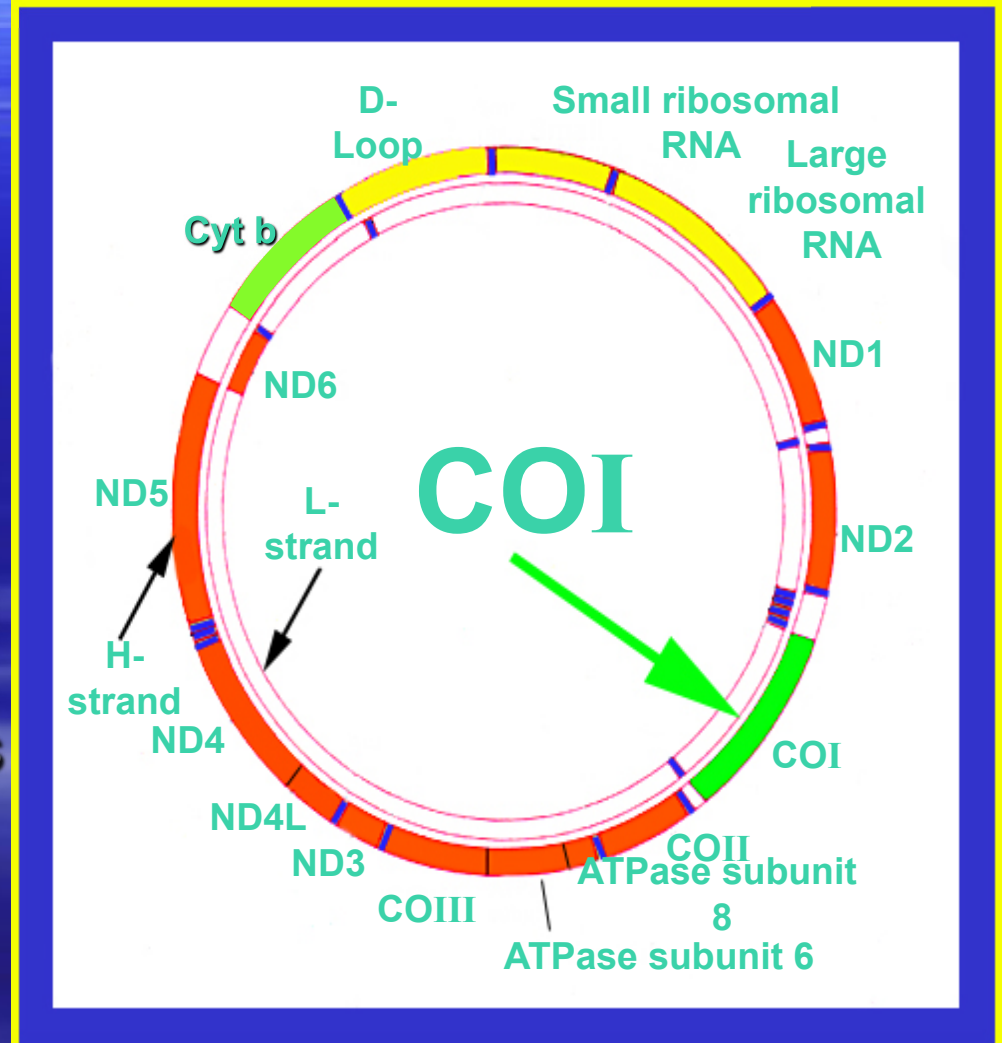


Mitochondrion



The Mitochondrial Genome: Cytochrome c oxidase subunit 1

- No pseudogenes
- Present in all animals
- COI appears to possess enough sequence divergence to regularly allow differentiation between closely related species
- The universal primers for this gene are robust, enabling the recovery of its 5' end from representatives of most animal phyla



Food-Related Applications of DNA Barcodes

- Verifying species of origin
 - Economic fraud (species substitution)
 - Harvesting endangered species
- Border inspection to control
 - invasive species
 - agricultural pests

Barcoding Project

Goals and Objectives

- Develop proficient protocols to extract, amplify, and sequence DNA to develop barcodes
- To create DNA barcodes of targeted species of FDA's interest

The Barcoding Process



Specimen or Tissue Sample



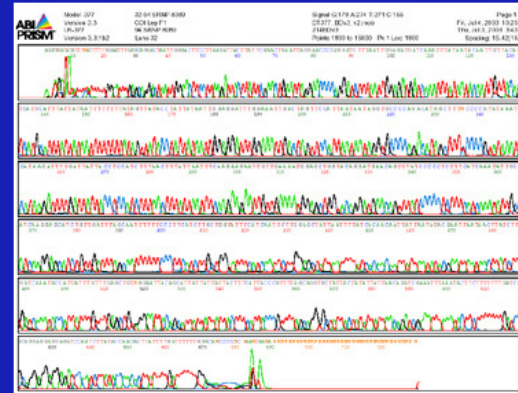
Extract DNA



PCR COI



Sequence COI



DNA Barcode

Barcode of Life Database (BoLD)

- www.barcodinglife.com
- Biodiversity Institute of Ontario, University of Guelph
- Unlimited users
- Password protected
- Worldwide collaboration

How to manage BoLD?

- Obtain a username and sponsor
- Create a project
- Submit specimen
- Upload Sequences
- Creation of Barcode

Build an Informatics Platform

BARCODE OF LIFE DATA SYSTEMS

Advancing species identification and discovery through the analysis of short, standardized gene regions

 [About BOLD](#) [Contact Us](#)

The Barcode of Life Data Systems (BOLD) is an online workbench that aids collection, management, analysis, and use of DNA barcodes. It consists of 3 components (MAS, IDS, and ECS) that each address the needs of various groups in the barcoding community.



MANAGEMENT & ANALYSIS



BOLD-MAS provides a repository for barcode records coupled with analytical tools. It serves as an online workbench for the DNA barcode community.

Username
Password
[Request a new user account](#)

- ▶ [Introductory tutorial](#)
- ▶ [View published projects](#)
- ▶ [Documentation](#)
- ▶ [Request an account](#)
- ▶ [Citation](#)

DATABASE STATISTICS

Species Barcoded	19,212
Total Barcodes	118,569
Source	Breakdown
GenBank	26,398
Canadian Centre	89,309
Others	2,862

BARCODING CAMPAIGNS



Lepidoptera: This initiative will assemble barcodes for 25,000 species of Lepidoptera with a focus on the faunas of Australia, Canada, Costa Rica and United States.



Fishes: The FISH-BOL campaign will gather barcodes for at least 15,000 fish species with an emphasis on marine species.



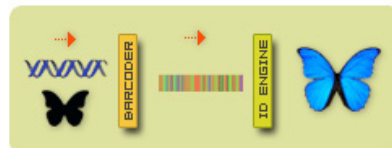
Birds: The All-Birds Barcode Initiative (ABBI) plans to assemble DNA barcodes for all 10,000 bird species within 5 years.



IDENTIFICATION ENGINE



BOLD-IDS provides a species identification tool that accepts DNA sequences from the barcode region and returns a taxonomic assignment to the species level when possible.



- ▶ [Request identification](#)
- ▶ [Citation](#)

BARCODING WEBSITES



Consortium for the Barcode of Life: The coalition of organizations advancing DNA barcode assembly.



Canadian Centre for DNA Barcoding: The hub of the Canadian Barcode Network, barcoding protocols and background.



Rockefeller University: Background on DNA barcoding assembled by the Program for the Human Environment.



EXTERNAL CONNECTIVITY



BOLD-ECS provides web developers and bioinformaticians the ability to build tools and workflows that can be integrated with the BOLD framework. We welcome the addition of new analytical modules.



- ▶ [NCBI](#)
- ▶ [EMBL](#)
- ▶ [DDBJ](#)
- ▶ [Citation](#)

Build an Informatics Platform

Management and Analysis

MAIN MENU | LOGOUT

BOLDSYSTEMS

Management & Analysis



General All Birds of North America [KGBNA]

Options

- List All Projects
- Summary View
- Move Records to another Project

Downloads

- Sequences
- Data Spreadsheet

Sequence Analysis

- Taxon ID Tree
- Distance Summary
- Sequence Composition
- Nearest Neighbor Summary

Specimen Aggregates

- Distribution Map
- Image Comparison

Project Data : <input type="text" value="Select"/>		Sequences/Specimens : 2573 / 3691				
Identification ▼	Sample ID ▼	ProcessID ▼	Length ▼	Contains	Extra Info ▼	Set
<input type="checkbox"/> Accipiter cooperii	1B-3244	TZBNA241-03	692		Accipiter cooperii	
<input type="checkbox"/> Accipiter cooperii	1B-3685	TZBNA288-03	652		Accipiter cooperii	
<input type="checkbox"/> Accipiter gentilis	1B-3677	TZBNA293-03	694		Accipiter gentilis	
<input type="checkbox"/> Accipiter gentilis	1B-3306	TZBNA297-03	668		Accipiter gentilis	
<input type="checkbox"/> Accipiter striatus	1B-3242	TZBNA238-03	693		Accipiter striatus	
<input type="checkbox"/> Acridotheres tristis	SVN 850	TZBNA018-03	692		Acridotheres tristis	
<input type="checkbox"/> Acridotheres tristis	SVN 844	TZBNA027-03	691		Acridotheres tristis	
<input type="checkbox"/> Actitis macularia	JGS 1881	TZBNA131-03	692		Actitis macularia	
<input type="checkbox"/> Actitis macularia	JAD 7476	TZBNA140-03	692		Actitis macularia	
<input type="checkbox"/> Aechmophorus occidentalis	1B-2755	HCBR175-03	448		Aechmophorus occidentalis	
<input type="checkbox"/> Aegolius acadicus	1B-3650	TZBNA294-03	694		Aegolius acadicus	
<input type="checkbox"/> Aegolius acadicus	1B-3352	TZBNA302-03	694		Aegolius acadicus	
<input type="checkbox"/> Aegolius funereus	1B-3234	TZBNA242-03	666		Aegolius funereus	
<input type="checkbox"/> Aegolius funereus	1B-3688	TZBNA289-03	668		Aegolius funereus	
<input type="checkbox"/> Aegolius funereus	1B-3674	TZBNA290-03	512		Aegolius funereus	
<input type="checkbox"/> Aethia cristatella	JP 2221	TZBNA220-03	694		Aethia cristatella	
<input type="checkbox"/> Aethia psittacula	JP 2351	TZBNA224-03	694		Aethia psittacula	
<input type="checkbox"/> Aethia psittacula	JP 2366	TZBNA225-03	694		Aethia psittacula	
<input type="checkbox"/> Aethia pusilla	JP 3139	TZBNA221-03	694		Aethia pusilla	
<input type="checkbox"/> Aethia pygmaea	JP 3386	TZBNA222-03	694		Aethia pygmaea	
<input type="checkbox"/> Agelaius phoeniceus	1B-2901	HCBR173-03	629			
<input type="checkbox"/> Agelaius phoeniceus	1B-2918	HCBR174-03	662			
<input type="checkbox"/> Aimophila ruficeps	JCB 5476	TZBNA054-03	694		Aimophila ruficeps	
<input type="checkbox"/> Aimophila ruficeps	JCB 5475	TZBNA063-03	694		Aimophila ruficeps	
<input type="checkbox"/> Aix sponsa	1B-3146	TZBNA257-03	693		Aix sponsa	
<input type="checkbox"/> Aix sponsa	1B-3143	TZBNA268-03	694		Aix sponsa	
<input type="checkbox"/> Aix sponsa	1B-3345	TZBNA295-03	661		Aix sponsa	
<input type="checkbox"/> Alca torda	GC 001	HCBR171-03	658			
<input type="checkbox"/> Alca torda	GC 002	HCBR172-03	658			

PROJECT MANAGEMENT - Bees of Nova Scotia

Options

- Submit Specimens
- Upload Sequences
- Modify Project Properties

Analysis (selected items)

- Sequence Composition
- Distance Summary (Fast)
- Distance Summary (Full)
- Taxon ID Tree
- Amino Acid Tree
- Sequence Recovery Stats
- Taxon Congruence (tree)
- Taxon Congruence (dist)
- Compare Images
- Distribution Map

Project Data Start Date : 2003-08-26 Sequences/Specimens : 317/390

<input checked="" type="checkbox"/>	Identification ▼	Specimen ID ▼	Sequence ID ▼	Sequence Length ▼	Properties	Specimen label	B#
<input type="checkbox"/>	<i>Agapostemon virescens</i>	Bee198	HCBNS209-03	638		Halictinae	UG1
<input type="checkbox"/>	<i>Agapostemon virescens</i>	Bee205	HCBNS210-03	624		Halictinae	UG1
<input type="checkbox"/>	<i>Agapostemon virescens</i>	SUP 7-2	HCBNS143-03	639		Halictinae	UG1
<input type="checkbox"/>	<i>Agapostemon virescens</i>	SUP 7-1	HCBNS142-03	627		Halictinae	UG1
<input type="checkbox"/>	<i>Andrena canadensis</i>	Bee3-1	HCBNS017-03	626		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena canadensis</i>	sheff-c22	HDBNS022-03	391		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena canadensis</i>	Bee3-2	HCBNS018-03	639		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena carlini</i>	Bee188	HCBNS208-03	385		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena carlini</i>	Bee191	HCBNS230-03	401		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena crataegi</i>	sheff-c36	HDBNS036-03	0		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena crataegi</i>	sheff-c40	HDBNS040-03	390		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena crataegi</i>	sheff-c37	HDBNS037-03	0		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena crataegi</i>	sheff-c39	HDBNS039-03	384		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena crataegi</i>	sheff-c38	HDBNS038-03	0		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena dunningi</i>	Bee3-11	HCBNS027-03	638		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena dunningi</i>	Bee3-12	HCBNS028-03	639		Andreninae	UG1
<input type="checkbox"/>	<i>Andrena dunningi</i>	Bee3-13	HCBNS029-03	639		Andreninae	UG1

Tools

Identification

Record Links

Details

Custom Labels

Build an Informatics Platform

BOLDSYSTEMS | Management & Analysis PRINT

BOLDSYSTEMS | Management & Analysis PRINT

Spingidae of the ACG 1 [MHASA]

Specimen Identifiers

Sample ID :	02-SRNP-9473	Catalog Number :	02-SRNP-9473
Isolate / Field Num:		Collection Code :	
Donated By :		Vouchered at :	2

Specimen Identifiers
Sample ID :
Isolate / Field Num:
Donated By :

Taxonomy


Identifier :	
phylum :	
class :	
order :	
family :	
subfamily :	
genus :	
species :	

Collection


Collectors :	
Date Collected :	
Country :	
State/Province :	
Region/County :	
Sector :	
Exact Site :	
Latitude :	
Longitude :	
Coord. Source :	
Elevation/Depth :	

Photographs

Dorsal View




Ventral View



Build an Informatics Platform

MAIN MENU | LOGOUT
PRINT

BOLDSYS **BOLDSYSTEMS** | Management & Analysis



Hesperiidae of the ACG 1 [CSCR]

Barcode Identifiers

Barcode ID :	MHAHC145-05	Sample ID :	04-SRNP-30754
Gene :	COX1	GenBank Accession :	
Last Updated :		Translation Matrix :	

Sequencing Runs

Run Date	Run Site	Direction	Trace File	PCR primers	Seq Primer
2005-02-22 16:28:14	Biodiversity Institute of Ontario	Reverse	MHAHC145-05R_E07.ab1	LepF1/LepR1	LepR1
2005-02-22 11:44:37	Biodiversity Institute of Ontario	Forward	MHAHC145-05F_E07.ab1	LepF1/LepR1	LepF1

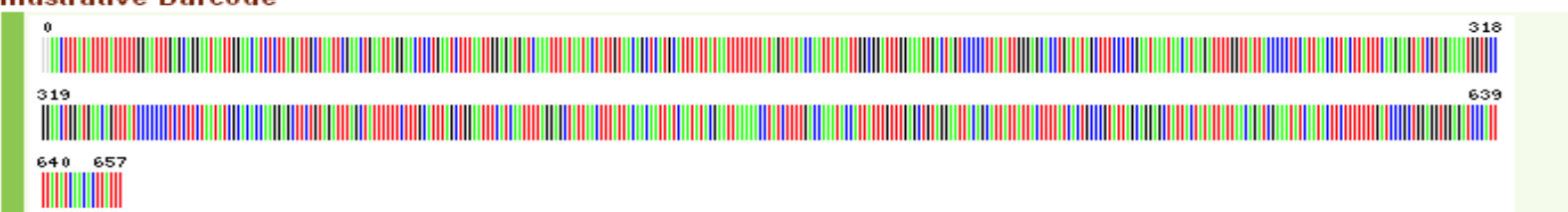
Nucleotide Sequence

Length :	658	NNAAC TTTATATTTTATTTTGGAA TTTGAGCAGGAATAA TTGGAACATCTCTTAGATTGCTAATTCGAACTGAA
Comp. A :	201	TTAGGAACTCCTGAATCTTTAATTGGAGATGATCAAATTTATAA TACTATTGTAACAGCTCATGCATTTATTATA
Comp. G :	94	ATTTTTTTTATAGTTATACCAATTATAA TTGGCGGATTTGGAAAT TGACTAGTCCCCCTTATATTGGGAGCACCT
Comp. C :	114	GATATAGCTTTCCCTCGAATAAATAACATAAGATTTTGGTTATTACCCCTTCATTAACCTTACTTTATTTCAAGA
Comp. T :	249	AGTATCGTAGAAAA TGGTGC CGGAAC TGGATGAACAGTTTATCCCCCTCTCTTCTAATATCGCACACCAAGGA
Updated :	2005-08-26	GCTTCTGTAGATTTAGCTATTTTTCTTTGCA TTTAGCTGGAATTTTCATCAATTTTAGGAGCTATTAAC TTTATT
		ACAACAATTATCAATATACGAATTA AAAACCTATCTTTTGACCAAATACCATTATTTGTTTGAGCTGTAGGAATT
		ACAGCATTATTACTTTTATCACTGCCCGTATTAGCAGGAGCTATTACTATATTATTAACAGATCGAAATATC
		AATACTTCTTTTTTGGATCCCGCTGGAGGTGGAGATCCCATTTTATATCAACACTTATTT

Amino Acid Sequence

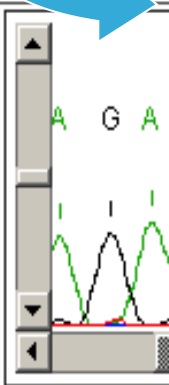
Length :	220	XTLYFIFGIWAGMIGTSL SLLIRTELGTPESLIGDDQIYNTIVTAHAFIMIFFMVMPIMIGGFGNWLVP LMLGAP
		DMAFPRMNMNSFWLLP PSLTLLLISSIVENGAGTGWTVYPP LSSNIAHQGASVDLAI FSLHLGAISSILGAINFI
		TTIINMRIK NLSFDQMP LFWAVGITALLLLSLPVLGAI TMLLDRNINTSFFDPAGGDP ILYQHLF

Illustrative Barcode



This viewer requires Java

[Click here to download](#)



Identification of a specimen

Build an Informatics Platform

Barcode of Life Data Systems

BARCODE OF LIFE DATA SYSTEMS

Advancing species identification and discovery through the analysis of short, standardized gene regions

SEARCH

About BOLD Contact Us



The Barcode of Life Data Systems (BOLD) is an online workbench that aids collection, management, analysis, and use of DNA barcodes. It consists of 3 components (MAS, IDS, and ECS) that each address the needs of various groups in the barcoding community.

MANAGEMENT & ANALYSIS



BOLD-MAS provides a repository for barcode records coupled with analytical tools. It serves as an online workbench for the DNA barcode community.

Username
Password

[Request a new user account](#)

- ▶ [Introductory tutorial](#)
- ▶ [View published projects](#)
- ▶ [Documentation](#)
- ▶ [Request an account](#)
- ▶ [Citation](#)

IDENTIFICATION ENGINE



BOLD-IDS provides a species identification tool that accepts DNA sequences from the barcode region and returns a taxonomic assignment to the species level when possible.



- ▶ [Request identification](#)
- ▶ [Citation](#)

EXTERNAL CONNECTIVITY



BOLD-ECS provides web developers and bioinformaticians the ability to build tools and workflows that can be integrated with the BOLD framework. We welcome the addition of new analytical modules.



- ▶ [NCBI](#)
- ▶ [EMBL](#)
- ▶ [DDBJ](#)
- ▶ [Citation](#)

DATABASE STATISTICS

Species Barcoded	19,212
Total Barcodes	118,569
Source	Breakdown
GenBank	26,398
Canadian Centre	89,309
Others	2,862

BARCODING CAMPAIGNS



Lepidoptera: This initiative will assemble barcodes for 25,000 species of Lepidoptera with a focus on the faunas of Australia, Canada, Costa Rica and United States.



Fishes: The FISH-BOL campaign will gather barcodes for at least 15,000 fish species with an emphasis on marine species.

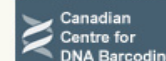


Birds: The All-Birds Barcode Initiative (ABBI) plans to assemble DNA barcodes for all 10,000 bird species within 5 years.

BARCODING WEBSITES



Consortium for the Barcode of Life: The coalition of organizations advancing DNA barcode assembly.



Canadian Centre for DNA Barcoding: The hub of the Canadian Barcode Network, barcoding protocols and background.



Rockefeller University: Background on DNA barcoding assembled by the Program for the Human Environment.



Cytochrome Oxidase Subunit 1 [COX1]

Internal Transcribed Spacer Region
[ITS]

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

The reference database of validated records is used by default and is recommended for all identification purposes.

Search Databases:

- All Barcode Records on BOLD (542,193 sequences)**
Every COI barcode record on BOLD with a minimum sequence length of 500bp (warning: unvalidated database 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 (489,261 sequences/37,849 Species)**
Every COI barcode record with a species level identification and a minimum sequence length of 500bp (warning: unvalidated dataset). This includes many species represented by only one or two specimens as well as all species with interim taxonomy.
- Reference Barcode Database (234,592 Sequences/13,774 Species)**
Validated subset of the full database with a minimum sequence length of 500bp and containing only those species represented by three or more individuals showing less than 2% sequence divergence

```
CCTCTATTTAGTATTTGGTGCCTGAGCCGGGATAGTAGGCACCGCCCTGAGTCTACTGATTCGGGCGGA  
ACTAAG  
CCAGCCGGGCGCTCTTCTGGGGGATGACCAAATCTATAACGTGATCGTCACAGCCCATGCCTTCGTTAT  
GATTTT  
CTTTATAGTCATGCCAATTATAATCGGGGGCTTTGGAAACTGATTAATTCCTTAATAATCGGAGCCCCT  
GATAT  
GGCATTCCCTCGAATAAATAACATAAGCTTCTGACTCCTTCTCCATCCTTTCTCCTCCTCCTGTCTTCAT  
CAGG  
AGTTGAAGCCGGCGCGGGTACTGGATGAACAGTATACCCCCCTCTAGCCGGCAACCTCGCCCACGCAGG  
AGCCTC  
TGTTGATTTAACTATCTTCTCCCTTCATTTAGCTGGAATCTCCTCAATTTTAGGAGCCATTAATTTTATTA
```

Submit

Cancel



Specimen Identification Request

Search Request:

Type : Reference Database Search

Search Result:

Identification Summary :

Taxonomic Level	Taxon Assignment	Probability of Placement (%)
Phylum	Chordata	100
Class	Actinopterygii	100
Order	Salmoniformes	100
Family	Salmonidae	100
Genus	<i>Oncorhynchus</i>	100
Species	<i>Oncorhynchus gorbuscha</i>	100

Distance Summary :

Similarity scores of the top 100 matches

A species level match has been made. This identification is solid unless there is a very closely allied congeneric species that has not yet been analyzed. Such cases are rare.

Tree Based Identification

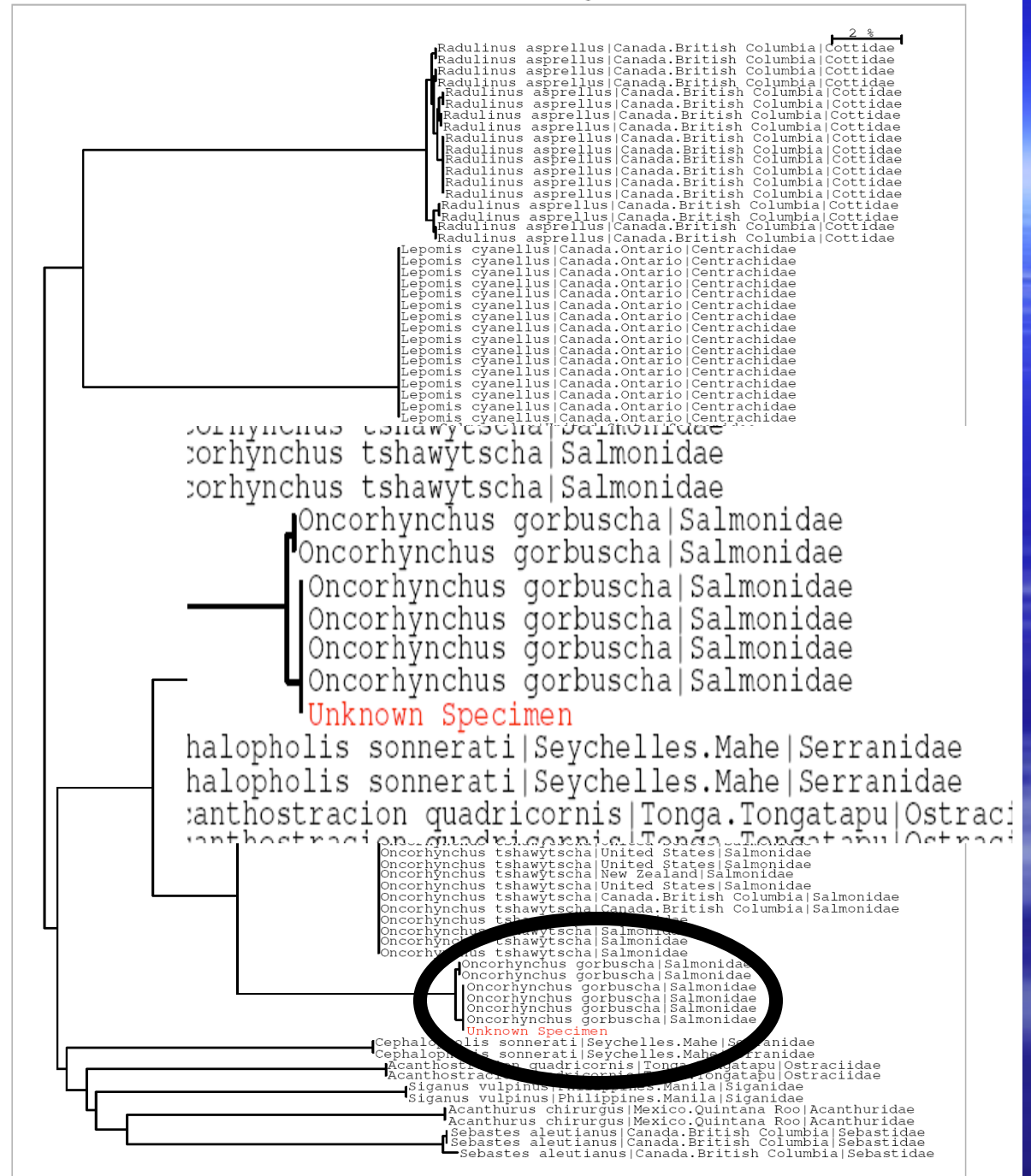
Species Page

TOP 20 Matches :

Display option: default

Phylum	Class	Order	Family	Genus	Species	Specimen Similarity (%)
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>gorbuscha</i>	100
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>gorbuscha</i>	100
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>gorbuscha</i>	100
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>gorbuscha</i>	100
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>gorbuscha</i>	99.69
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>gorbuscha</i>	99.69
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>gorbuscha</i>	99.69
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	89.37
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	89.37
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	89.37
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	89.37
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	89.37
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	89.37
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	89.37
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	89.37
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	89.37
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	89.37
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	89.37
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	88.89
Chordata	Actinopterygii	Salmoniformes	Salmonidae	Oncorhynchus	<i>tshawytscha</i>	88.89

Tree Based Identification



Species page

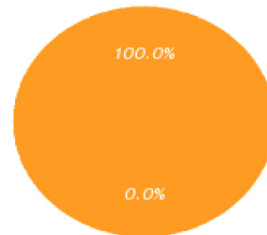
Oncorhynchus gorbuscha [species]

Lineage : Chordata; Actinopterygii; Salmoniformes; Salmonidae; Oncorhynchus;

Specimens : 3

Identifiers :

Deposited in :



Remaining: 0 Barcoded: 3

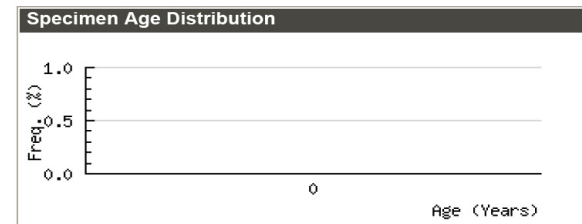
External Species Pages : [FishBase](#)



Collectors: Collected in :

Warning: array_unique(): The argument should be an array in /var/www/html/pm3/views/speciesrecord.php on line 261

Warning: Invalid argument supplied for foreach() in /var/www/html/pm3/views/speciesrecord.php on line 263



Lateral



MPF: Fillet & Steak

2. Speed Analysis

Costs and Time

	Fresh/Frozen	Time
	\$0.41	10
	\$0.34	10
	\$0.24	20
	\$0.35	5
	\$1.04	30
	\$0.32	5
	\$0.40	35
<u>Total:</u>	\$3.10	115 min

To Do List (Fall 2005)

- Obtain and barcode the remaining samples of the Regulatory Fish Encyclopedia (RFE) (Seafood Products Research Center Pacific Regional Laboratory – Northwest)
- Complete the Barcode of Life Database (BoLD) pages for each fish on the RFE
- Quality Assurance/Quality Check of protocols and data
 - The FBI Laboratory's Combined DNA Index System (CODIS)
- Either replace or link DNA barcode to the RFE

Regulatory Fish Encyclopedia (RFE)


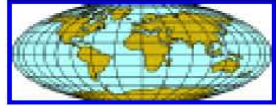
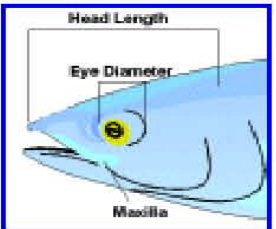
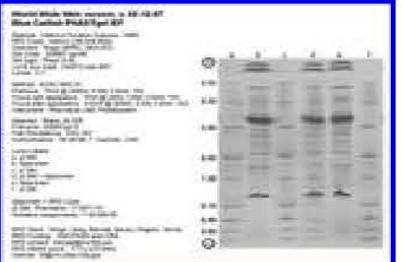
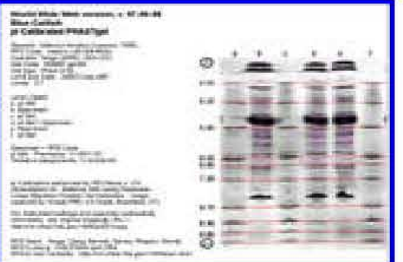
- The Regulatory Fish Encyclopedia (RFE) is a compilation of data in several formats that assists with the accurate identification of fish species
- It was developed by FDA's RFE Team to help federal, state, and local officials and purchasers of seafood to identify species substitution and economic deception in the marketplace
- <http://www.cfsan.fda.gov>





Regulatory Fish Encyclopedia (RFE)

- Includes 94 commercially relevant fish for sale in the U.S. market
 - Visual comparison
 - Scanned digital images (whole fish and their market form)
 - Biochemical comparison
 - Isoelectric focusing (IEF)

	<p>U.S. Food & Drug Administration Seafood Products Research Center Center for Food Safety & Applied Nutrition Regulatory Fish Encyclopedia</p>	
--	--	---

RFE Page 0 - Top Page for Blue Catfish (Catfish; *Ictalurus furcatus*)

				
<p><u>RFE PAGE 1</u> <i>High Resolution Photographic Images</i></p>	<p><u>RFE PAGE 2</u> <i>Hazard, Market, Geographic and Nomenclature Information</i></p>	<p><u>RFE PAGE 3</u> <i>Taxonomic Information</i></p>	<p><u>RFE PAGE 4</u> <i>IEF Protein Patterns</i></p>	<p><u>RFE PAGE 5</u> <i>IEF Protein Pattern Analysis Toolkit</i></p>

 {Data not available yet}	 {Data not available yet}	 {Data not available yet}	 {Data not available yet}
<p>RFE PAGE 6 <i>DNA Fragment Patterns</i></p>	<p>RFE PAGE 7 <i>DNA Fragment Pattern Analysis Toolkit</i></p>	<p>RFE PAGE 8 <i>DNA Sequence Information</i></p>	<p>RFE PAGE 9 <i>Species-Related and General References</i></p>

This [DISCLAIMER](#) pertains to the use of the RFE and to the use of images and data therein.

Isoelectric Focusing (IEF) Protein Fingerprint

World Wide Web version, v. 02-12-97 Channel Catfish PHASTgel IEF

Species: *Ictalurus punctatus* (Rafinesque, 1818)
RFE Code: ictapunc K95-004 #04p
Operator: Tenge (SPRC, SEA-DO)
Gel Code: 020697 gel #8
Gel type: Phast (3-9)
Lot & Exp Date: 243072 exp 4/97
Lanes: 2-7

Method: AOAC 980.16
Prefocus: 75Vh @ 2000V, 3.5W, 2.5mA, 15C
Focus with applicators: 15Vh @ 200V, 3.5W, 2.5mA, 15C
Focus after applicators: 410Vh @ 2000V, 3.5W, 2.5mA, 15C
Instrument: Pharmacia LKB, PhastSystem

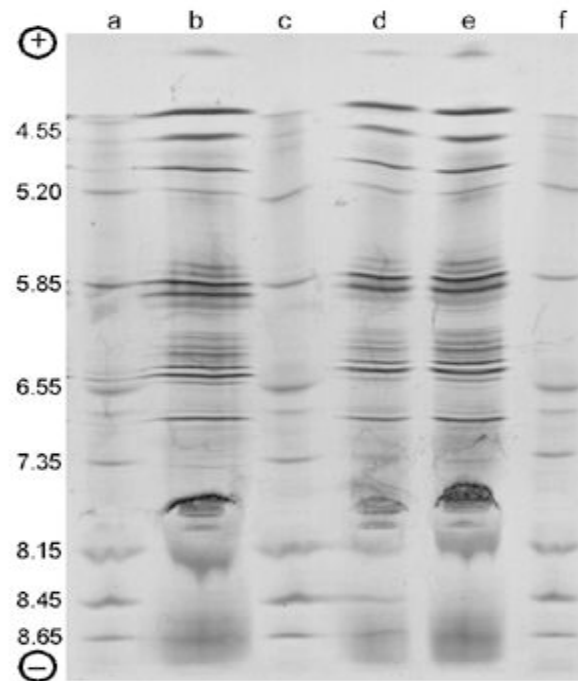
Scanner: Sharp JX-325
Filename: 020697g8.tif
Fish Provided by: NOL-DO
Authentication: 09-08-98, T. Iwamoto, CAS

Lane Labels:

a pl Std
b Specimen
c pl Std
d pl Std + Specimen
e Specimen
f pl Std

Specimen = RFE Code
pl Std: Pharmacia, 17-0471-01
Tentative assignments, 11-B-045-05

RFE Team: Tenge, Dang, Barnett, Savary, Rogers, Gerrity
RFE Funding: OS/CFSAN and ORA
RFE contact: btenge@ora.fda.par
RFE WWW coord.: F. Fry (CFSAN)
Internet: frf@vm.cfsan.fda.gov



Isoelectric focusing protein patterns for the pH range 3-10 for fresh, frozen, uncooked seafood based on AOAC 980.16

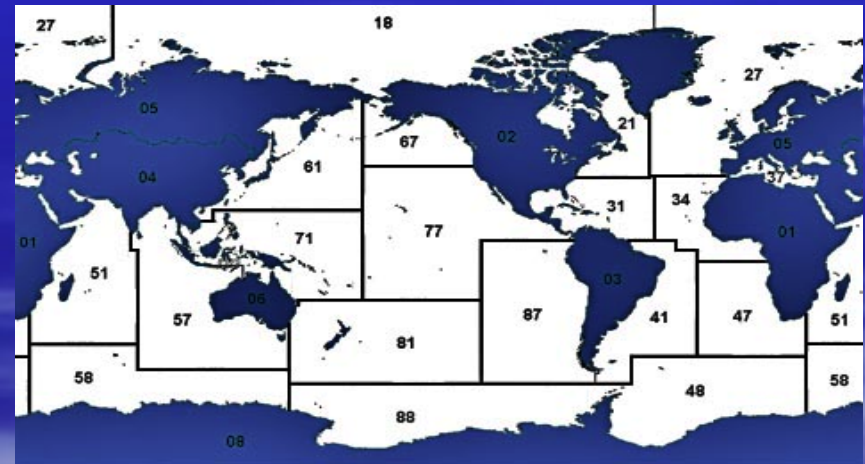
Problems with IEF Identification

- Proteins
 - Sample condition
 - Will not work on denatured samples (cooked/neglected)
 - Sample preparation
 - Uses Association of Official Analytical Chemists (AOAC) validated method, but control samples are becoming exhausted
- Different platforms
 - Inconsistent gel patterns

FISH-BOL

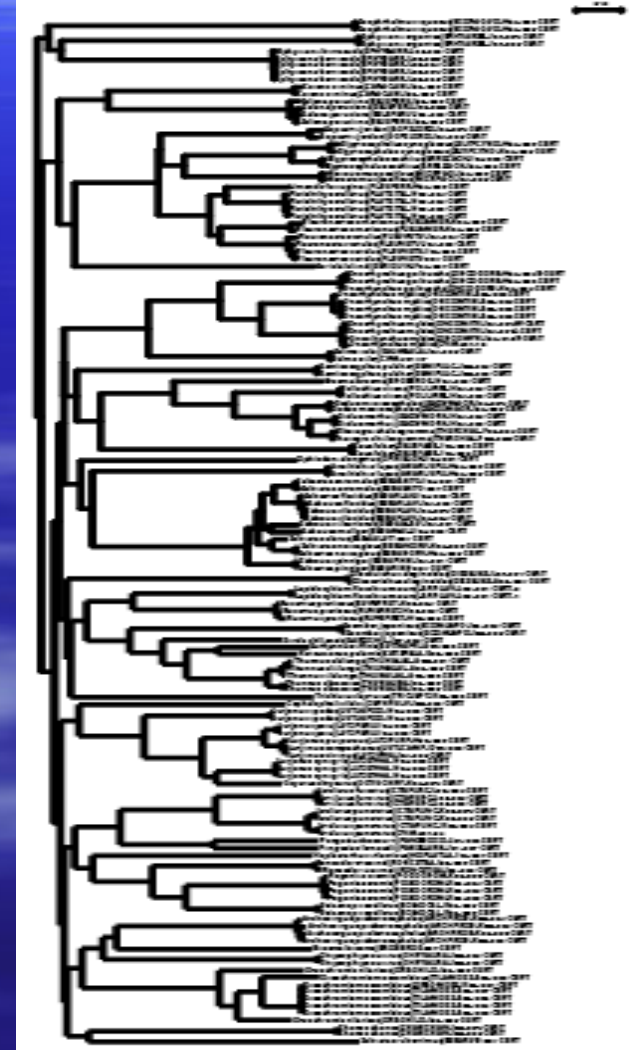


- Dr. Paul Hebert and Dr. Bob Hanner Biodiversity Institute of Ontario, University of Guelph
- Established June 2005
- 5,876 species barcoded
- 34,908 barcodes





- FDA RFE project
 - 179 specimens
 - 172 barcodes (sequences)
 - 73 species
- Blind Sample Trail
 - 60 samples were tested
 - 60/60 were identified correctly



Research Note

Potential Use of DNA Barcodes in Regulatory Science: Applications of the *Regulatory Fish Encyclopedia*

HAILE F. YANCY,^{1*} TYLER S. ZEMLAK,² JACQUILINE A. MASON,³ JEWELL D. WASHINGTON,¹
BRADLEY J. TENGE,⁴ NGOC-LAN T. NGUYEN,^{4†} JAMES D. BARNETT,^{4†} WARREN E. SAVARY,⁴
WALTER E. HILL,^{4†} MICHELLE M. MOORE,⁴ FREDERICK S. FRY,⁵ SPRING C. RANDOLPH,⁵
PATRICIA L. ROGERS,⁵ AND PAUL D. N. HEBERT²

¹U.S. Food and Drug Administration, Center for Veterinary Medicine, Office of Research, 8401 Muirkirk Road, Laurel, Maryland 20708, USA; ²Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, Canada N1G 2W1; ³Department of Microbiology, College of Medicine, Howard University, Washington, D.C. 20059, USA; ⁴U.S. Food and Drug Administration, Office of Regulatory Affairs, Seafood Products Research Center, Bothell, Washington 98021, USA; and ⁵U.S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, College Park, Maryland 20740, USA

MS 07-141: Received 16 March 2007/Accepted 17 August 2007

ABSTRACT

The use of a DNA-based identification system (DNA barcoding) founded on the mitochondrial gene cytochrome *c* oxidase subunit I (COI) was investigated for updating the U.S. Food and Drug Administration *Regulatory Fish Encyclopedia* (RFE; <http://www.cfsan.fda.gov/~frf/rfe0.html>). The RFE is a compilation of data used to identify fish species. It was compiled to help regulators identify species substitution that could result in potential adverse health consequences or could be a source of economic fraud. For each of many aquatic species commonly sold in the United States, the RFE includes high-resolution photographs of whole fish and their marketed product forms and species-specific biochemical patterns for authenticated fish species. These patterns currently include data from isoelectric focusing studies. In this study, we describe the generation of DNA barcodes for 172 individual authenticated fish representing 72 species from 27 families contained in the RFE. These barcode sequences can be used as an additional identification resource. In a blind study, 60 unknown fish muscle samples were barcoded, and the results were compared to the RFE barcode reference library. All 60 samples were correctly identified to species based on the barcoding data. Our study indicates that DNA barcoding can be a powerful tool for species identification and has broad potential applications.

All aquatic animals harvested, processed, distributed, and sold in the United States must be safe, wholesome, and properly labeled. Under the Federal Food, Drug, and Cosmetic Act, the Fair Packaging and Labeling Act, and the Public Health Service Act, the U.S. Food and Drug Administration (FDA) carries out a program that includes inspection, sampling, analysis, research, and education concerning seafood issues, safety, and labeling. The FDA also has oversight over economic fraud and food safety. Cases of consumer deception include the misbranding or improper labeling of a product and the substitution of an inferior product for a superior product. Seafood has garnered increasing attention because of potential health-related risks associated with misbranding. The major areas of concern and examples of species-specific hazards are listed in the hazard analysis and critical control point guide (28) (Table 1).

The Food Allergen and Protection Act requires unambiguous identity labeling of a food that is or contains an ingredient that is a major food allergen. Some of the aquatic species that may cause allergenic reactions are haddock,

cod, hake, halibut, mackerel, tuna, salmon, orange roughy, shrimp, and crab. The act also identifies species that may not present a particular health concern, such as catfish and basa, but that are covered by laws or regulations that require their identity to be monitored because of trade and tariff restrictions.

The FDA has been dealing with the problem of misbranding for many years. For example, the increase in seafood consumption and species substitution led the FDA and the National Marine Fisheries Service to recognize the need for a single source of market names to facilitate responsible trade in the marketplace and reduce confusion among consumers. In 1988, the FDA published the Fish List. Initially, the list contained only those fish species sold as part of interstate commerce, but it was revised in 1993 to include additional fish species and invertebrates and was renamed the Seafood List. Currently, this list is available as an updated searchable database on the FDA Center for Food Safety and Applied Nutrition web site (30).

One challenge faced by both consumers and regulators is the detection of seafood substitution in the marketplace, a practice where low-value species (or species with potential toxins) are mislabeled and/or substituted in whole or in

* Author for correspondence. Tel: 301-210-4096; Fax: 301-210-4685; E-mail: haile.yancy@fda.hhs.gov.

† Retired.

Laboratory Information Bulletin

**A Protocol for Validation of DNA-Barcoding for the Species
Identification of Fish for FDA Regulatory Compliance**

Haile F. Yancy¹, Frederick S. Fry², Spring C. Randolph², Jonathan Deeds², Natalia V. Ivanova³, Christopher M. Grainger³, Robert Hanner³, Lee A. Weigt⁴, Amy Driskell⁴, Jeffrey Hunt⁴, Andrea Ormos⁴ and Paul D.N. Hebert³

¹U. S. Food and Drug Administration, Center for Veterinary Medicine, Office of Research, 8401 Muirkirk Road, Laurel, Maryland 20708, USA

²U. S. Food and Drug Administration, Center for Food Safety and Applied Nutrition, College Park, MD 20740, USA

³Canadian Centre for DNA Barcoding, Biodiversity Institute of Ontario, University of Guelph, 579 Gordon Street, Guelph, ON, Canada, N1G 2W1

⁴Laboratories of Analytical Biology, National Museum of Natural History, Smithsonian Institution, MRC 534, Washington DC 20013-7012

The Laboratory Information Bulletin is a communication from the Division of Field Science, Office of Regulatory Affairs, U.S. Food and Drug Administration for the rapid dissemination of laboratory methods (or scientific regulatory information) which appear to solve a problem or improve an existing problem. In many cases, however, the report may not represent completed analytical work. The reader must assure, by appropriate validation procedures, that the reported methods or techniques are reliable and accurate for use as a regulatory method. Reference to any commercial materials, equipment, or process does not, in any way, constitute approval, endorsement, or recommendation by the U.S. Food and Drug Administration.

Assay Development Using Barcode Database



PCR product
~650 bp



PCR product
~350-250 bp

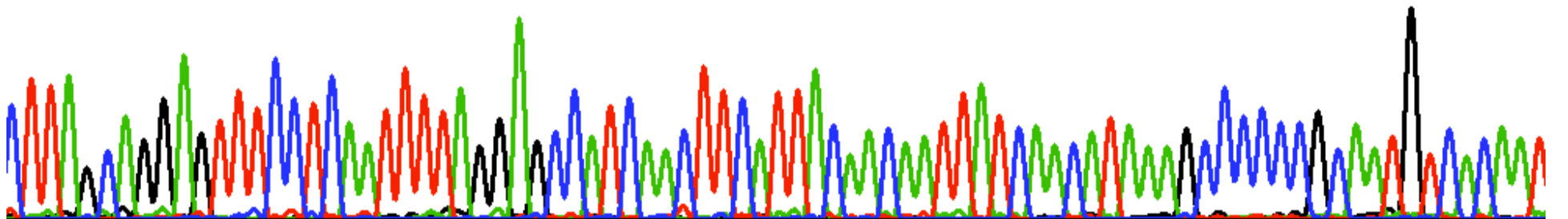


PCR product
~250- 50 bp



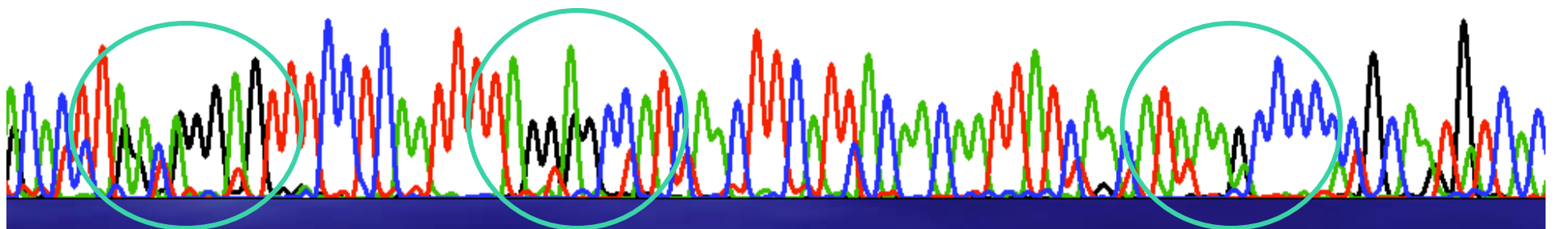
Tuna

410 420 430 440 450 460 470 480
C T T A G C A G G A G T T T C C T C A A T T T T A G G A G C C A T C A A C T T C A T T A C A A C A A T T A T C A A C A T A A A G C C C C C G C A A T G T C A C A A T



Mixed Sample (tuna/catfish)

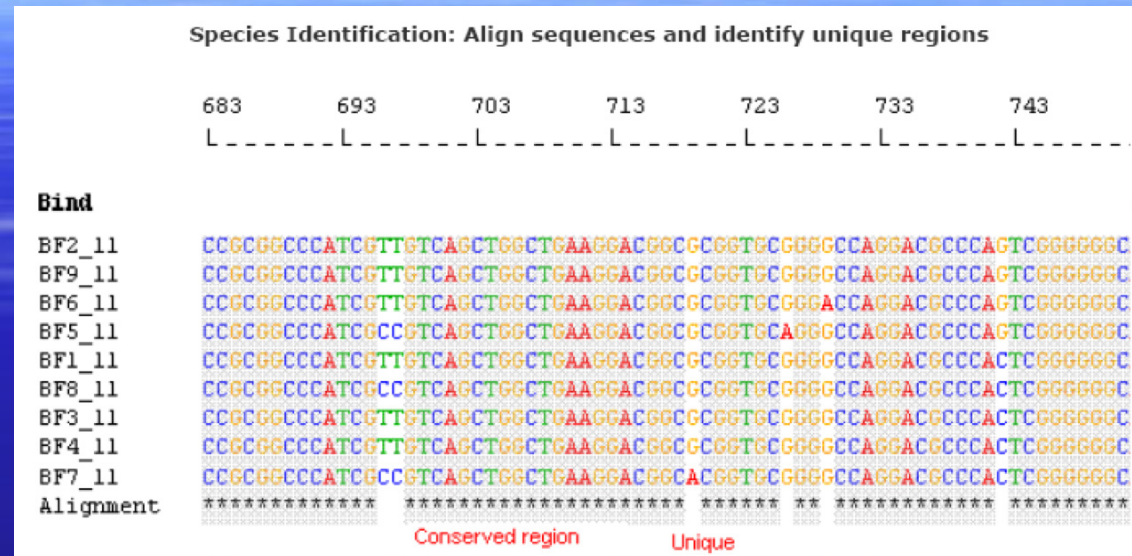
410 420 430 440 450 460 470 480
N N N C T N N N N N G A G T T T C C T C A A T T T T A G N A N C C A T C A A C T T C A T T A C A A C A A T T A T C A A C A T A N N G C C C C N C G C A A T G N C A C



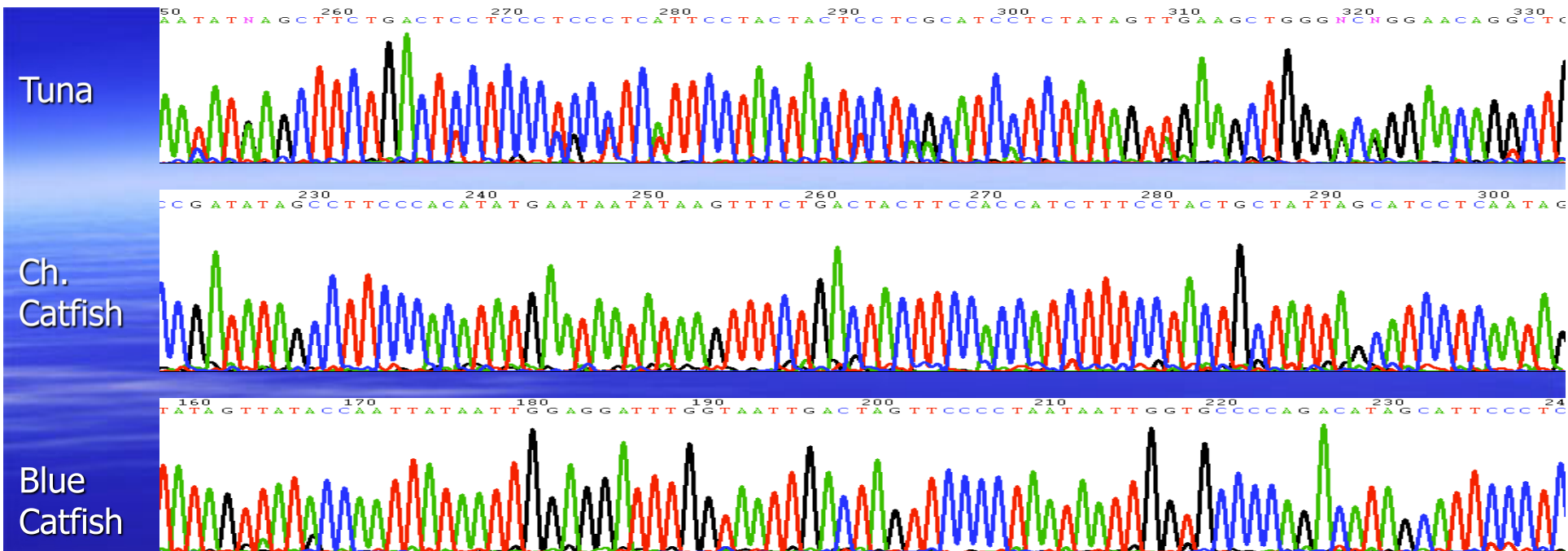
Assay Development

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

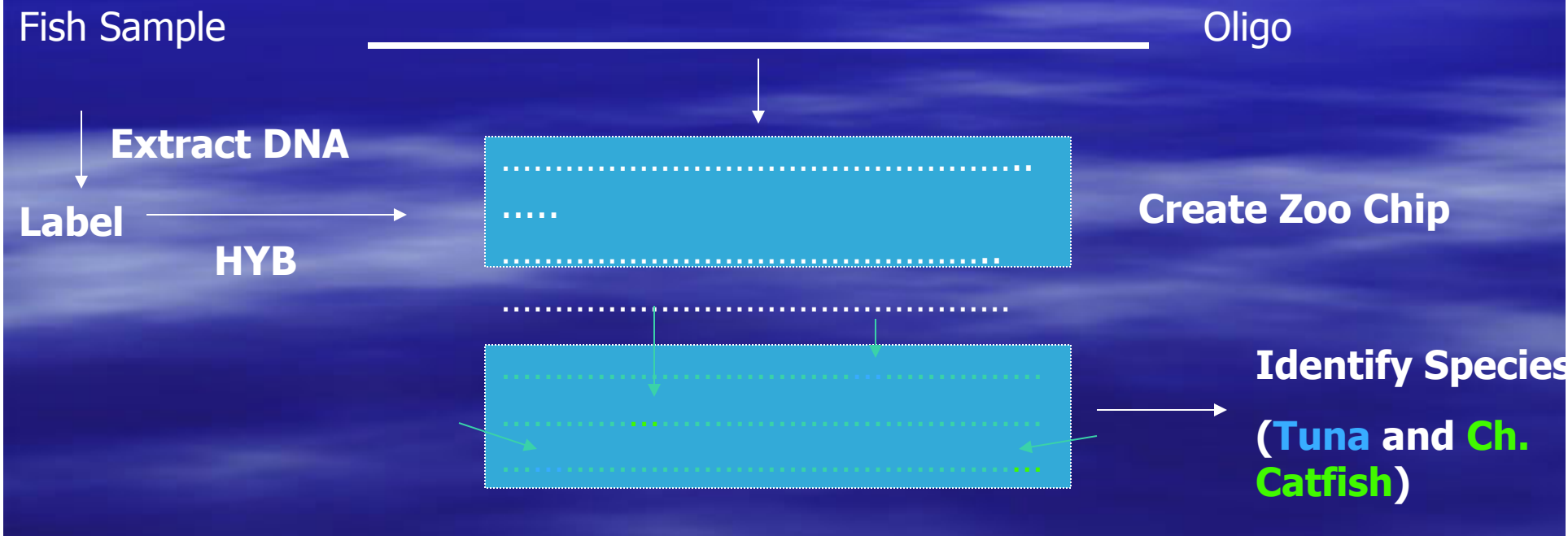
Allele ID



- 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



Allele ID

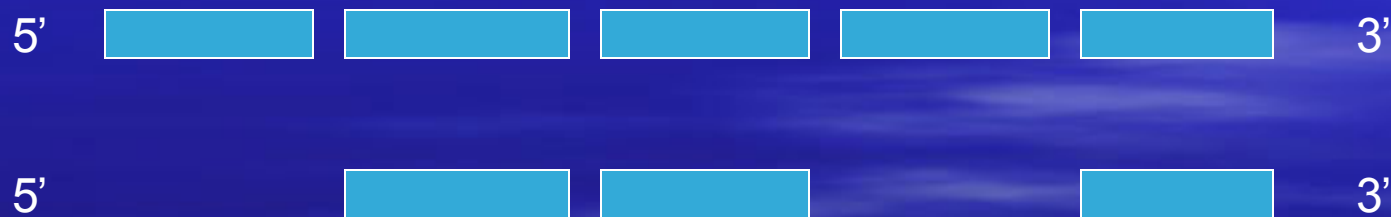


Designing Probes from the COI Gene

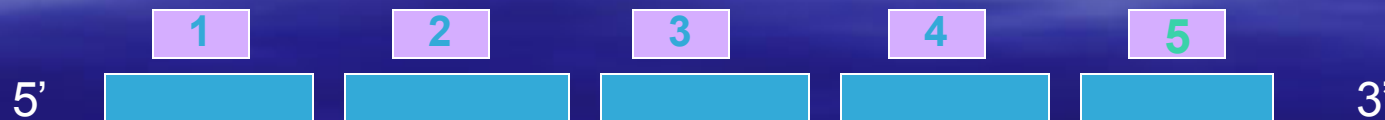
Generally probes are designed at a specific place along a gene sequence



Available DNA of cooked, processed, or degraded sample

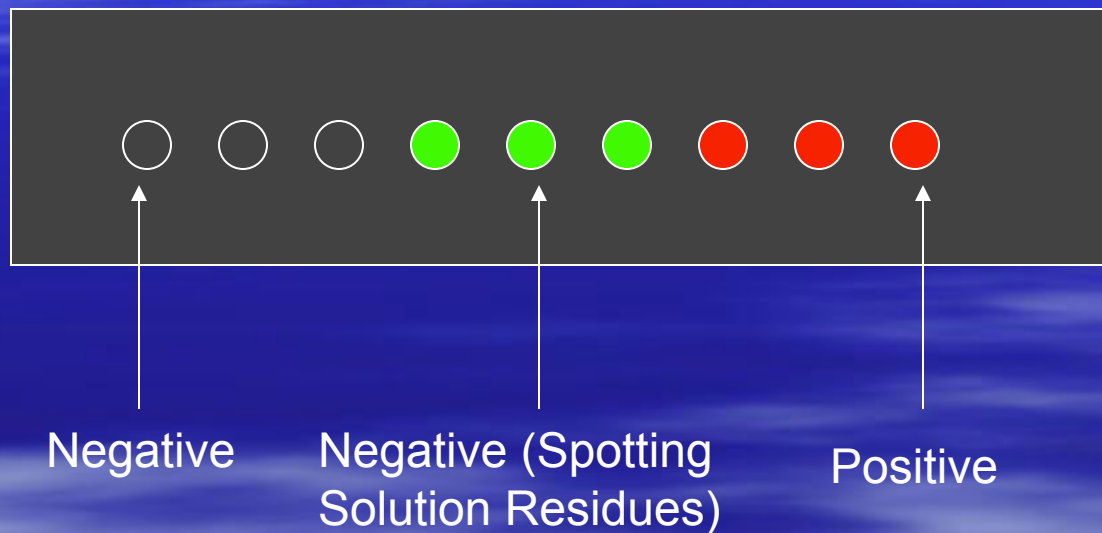


Several Probes Designed Along Span of Gene 45-55 nucleotides in length



Criteria for Analyzing Microarray Slides

1. Color of spots (specificity)



2. Ratio of detection (red) to background (green) fluorescence above 2.0 (sensitivity)

Binding of DNA to Rainbow Trout Probes

Fluorescence of Rainbow Trout DNA

Rainbow Trout Probe	Ratio of red to green fluorescence
Probe 1	5.34
Probe 5	20.34

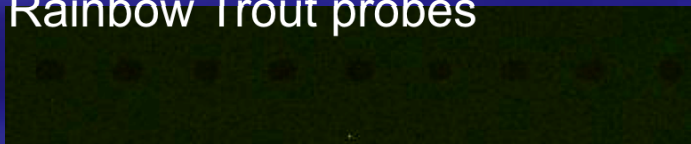
Rainbow Trout DNA with Rainbow Trout probes



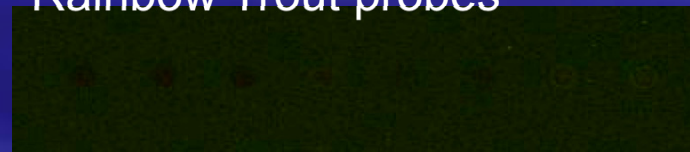
Channel Catfish DNA with Rainbow Trout probes



Great Barracuda DNA with Rainbow Trout probes



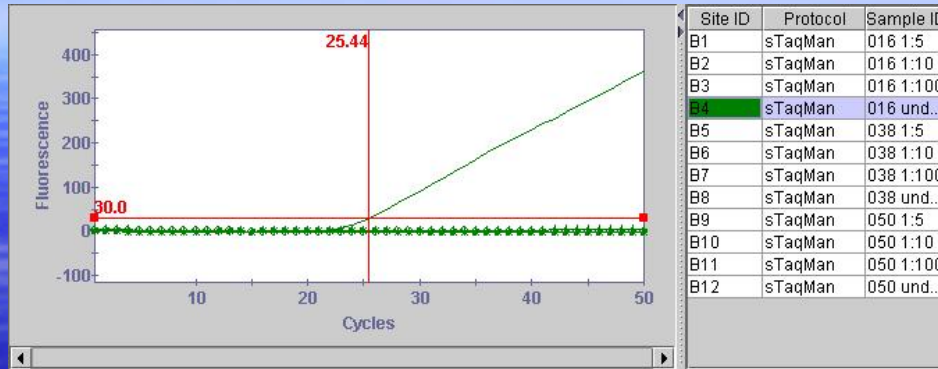
Blue Catfish DNA with Rainbow Trout probes



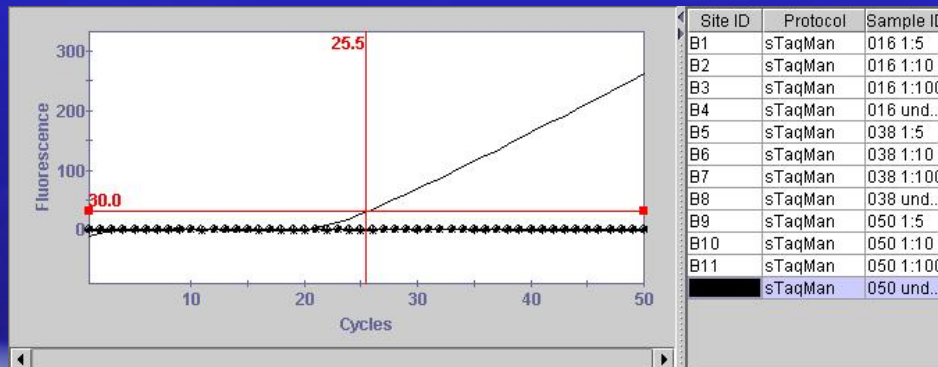
Real- Time PCR Assay

- Allele ID
- Aligned barcode sequences
- Create Taqman[®] probes that detect species of interest

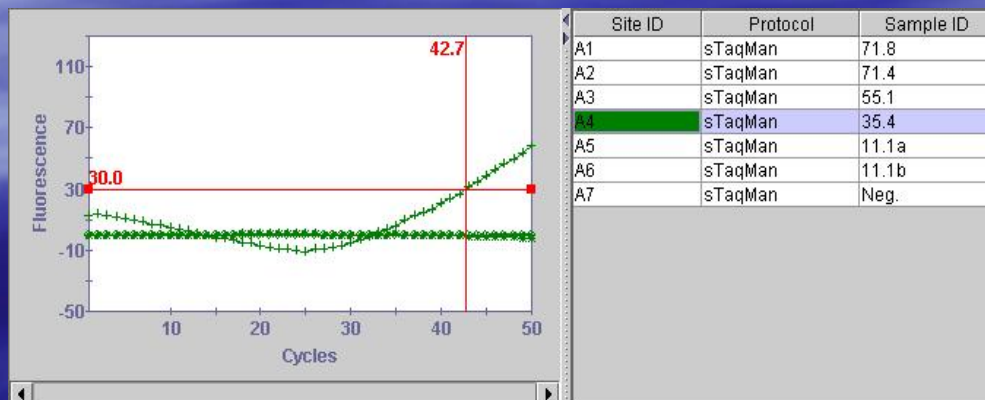
Optimization of Pufferfish Primers



Pufferfish Primer
PU016: Ct=25.44

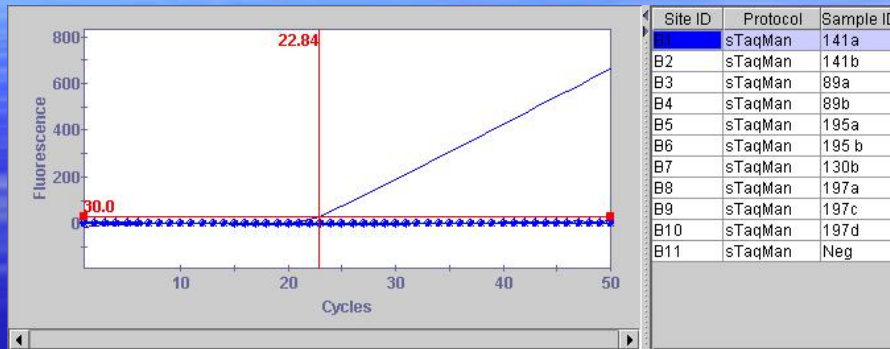


Pufferfish Primer
PU050: Ct=25.5

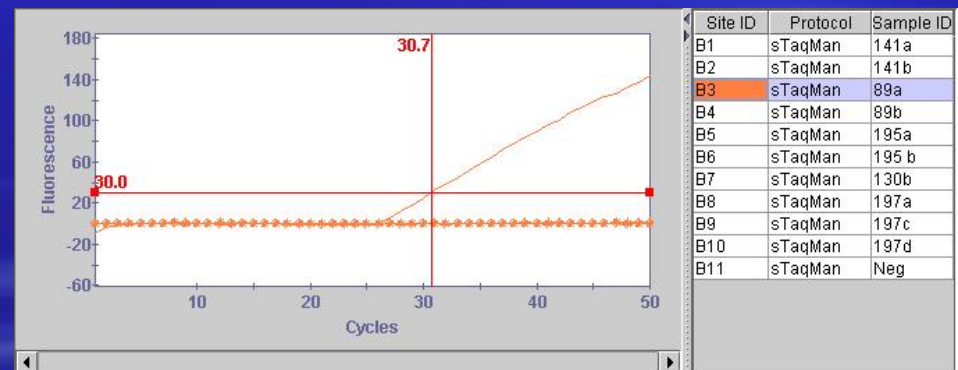


Pufferfish Primer
PU038: Ct=42.7

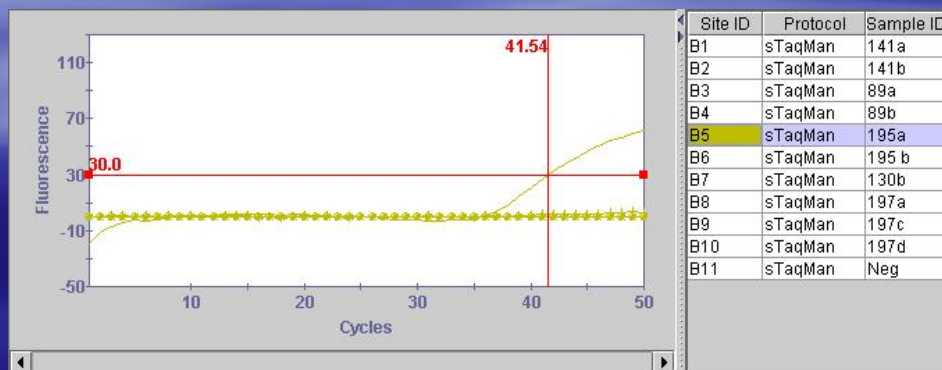
Known Pufferfish Samples with Primer PU016



Left: Florida Southern
Puffer: Ct=22.84



Right: Florida Bandtail
Puffer: Ct=30.7



Left: Florida Checkered
Puffer: Ct=41.54

Pufferfish Primer PU016 Specificity

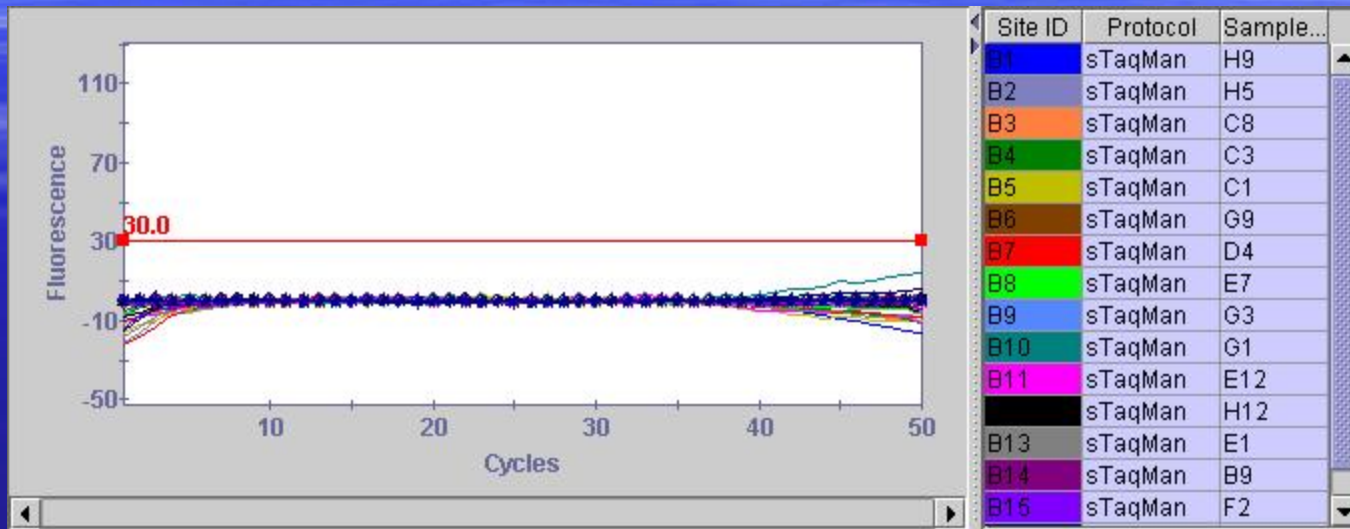


Figure 1. Primer PU016 was tested against 16 other fish species closely related to pufferfish to determine specificity. All samples came up negative

Cooked Fish from Soup

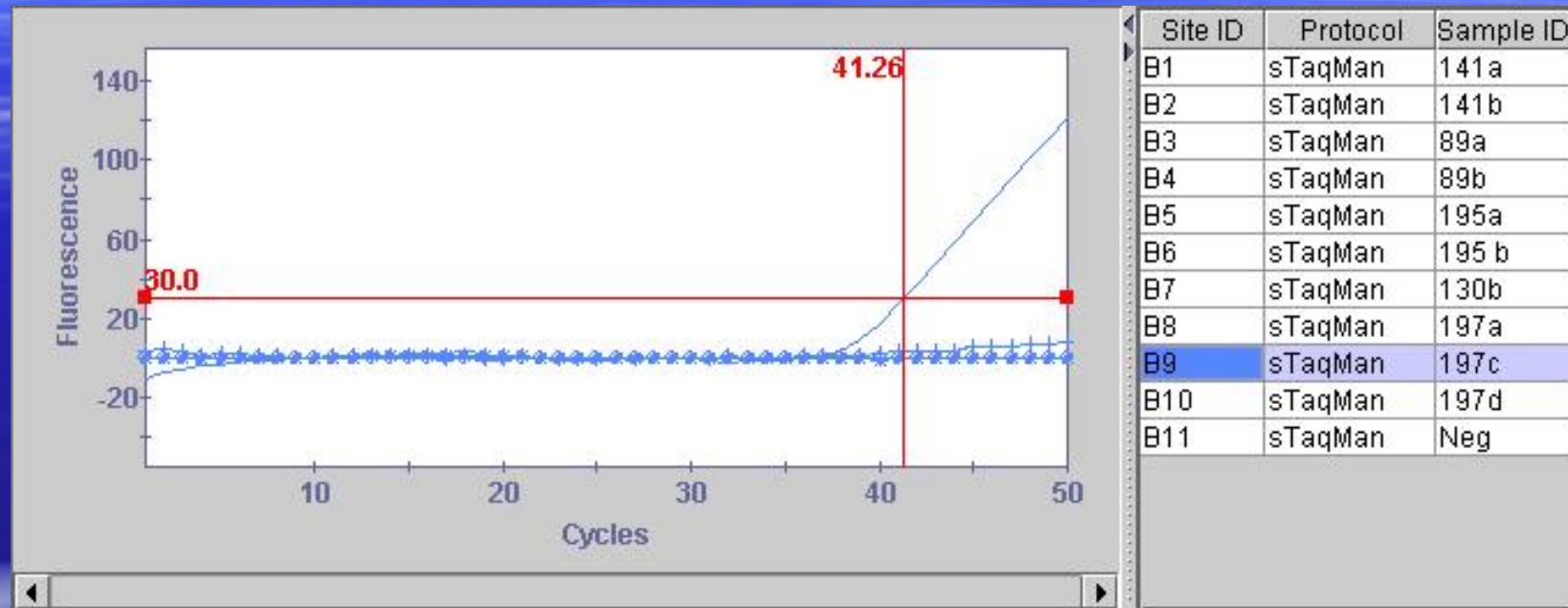


Figure 1. Cooked puffer sample from the soup that caused the illnesses. Primer PU016 was used. Ct=41.26

A Regulatory Tool for the Third Millennium



Acknowledgements

CVM

Dr. Dave White
Dr. Renate Reimschuessel
Dr. Jacqueline Mason
Lauren Callahan
Michelle LeCluyse
Haley Oliver

CFSAN

Dr. Fred Fry
Spring Randolph
Dr. Jon Deeds
Dr. Marleen Wekell

CONSORTIUM FOR THE BARCODE OF LIFE

Dr. Paul Hebert
Tyler Zemplak
Dr. Scott Miller
Dr. David Schindel
Dr. Bob Hanner

Seafood Products Research Center Pacific Regional Laboratory – Northwest

Dr. Rick Long
Dr. Cindy Wu
Jim Barnett
Ngoc-Lan Nguyen
Dr. Michelle Moore
Dr. Bradley Tenge

Thank You

Questions?