

## **FDA SOP for Generating DNA Barcodes Suitable for Species Identification of an Unknown Fish Tissue Sample:**

The process of generating DNA barcodes suitable for species identification from an unknown fish tissue sample can be broken down into the following steps: 1. Tissue Sampling, 2. Tissue Lysis and DNA Extraction, 3. Polymerase Chain Reaction – COI Amplification, 4. PCR Cleanup, 5. Cycle Sequencing Reaction, 6. Sequencing Reaction Cleanup, 7. Sequencing, 8. Post Sequencing Analysis. Individual labs will likely modify this SOP based on available equipment and reagents. In addition, for steps such as tissue lysis and DNA extraction, several commercial kits are available that may be acceptable. Any deviations from the described SOP will need to be proven to provide acceptable results. Performance criteria for individual steps are provided to aid in this process.

A detailed SOP based on the single laboratory validation study detailed in Handy et al. (2011), with some modifications, is provided below.

Modifications from Handy et al. (2011) are noted.

### **1. TISSUE SAMPLING (Note: Only necessary if a subset is needed for long term storage, e.g. reserve samples. If no reserve sample is required, proceed to Step 2.)**

#### **1A. Reagents and Consumables**

- a. Reagent Alcohol, Histological (EtOH 96%); Fisher Scientific Catalog No. A962-4
- b. Nalgene 2.0ml Cryogenic vial: Fisher Scientific Catalog No. 033377D

#### **1B. Protocol for Tissue Sampling**

**Goal:** Sample unknown fish tissues in a manner that prevents cross-contamination between samples, or the introduction of foreign DNA.

**Criteria for Success:** After sequencing, have a chromatogram that gives single peaks for each base as evidenced by the AB1 file.

Several tissues are suitable for DNA extraction from fishes. These include:

- Musculature (preferred): remove one or more cubes (5-7 mm) of lateral muscle (skin removed).
- Fin clips: remove fin rays and membrane from the pectoral or pelvic fin. Use these tissues only if the fish needs to be kept alive.

- Eye: remove the right eye from extremely small specimens such as larvae.

Each fish or fish fillet sampled should be bagged and labeled separately. Tissue samples should be removed with a scalpel and forceps that are flame sterilized after each sample is processed. New (sterile) tubes should be used for each sample. Tissues should be frozen at -20°C or preserved in fresh 95% ethanol and stored in a cool place, preferably in a freezer, until DNA is extracted. For long term storage of critical samples (e.g. reference standards, investigation reserve samples, etc.), tissues should be stored at -80°C.

## **2. TISSUE LYSIS AND DNA EXTRACTION**

**Goal:** Extract DNA acceptable for use in PCR.

**Criteria for Success:** Obtain  $\geq 5$  ng/ $\mu$ L of DNA for all samples (Note: this was the lowest quantity used for the SLV, lower quantities may work) measured on a Nanodrop ND 1000 Spectrophotometer (Thermo-scientific, Wilmington, DE) with a 260 nm/280 nm ratio of  $\sim 1.8$ . In addition, a negative control with no added DNA should give a reading of  $\sim 0$  ng/ $\mu$ L.

**DNeasy Blood & Tissue Kit Qiagen Catalog No. 69504 (50) or 69506 (250)**

### **2A. Reagents and Consumables**

- a. Collection Tubes (2 ml)
- b. Buffer ATL
- c. Buffer AL
- d. Buffer AW1 (concentrate)
- e. Buffer AW2 (concentrate)
- f. Buffer AE 22 ml 2 x 60 ml
- g. Proteinase K 1.25 ml 6 ml
- h. Pipets and pipet tips (see below for recommendations)
- i. Vortexer
- j. Ethanol (96-100%)
- k. Microcentrifuge tubes (1.5 ml)
- l. Microcentrifuge with rotor for 1.5 ml and 2.0 ml tubes
- m. Thermomixer, shaking water bath, rocking platform or dry block heater for heating microcentrifuge tubes at 56°C

### **2B. Protocol for Tissue Sampling**

#### **Preparation of Buffer AW1, AW2, and AL**

- A. Buffer AW1 and Buffer AW2 are supplied as concentrates. Before using for the first

time, add the appropriate volume of ethanol (96–100%) as indicated on the bottle and shake thoroughly. Buffer AW1 and Buffer AW2 are stable for at least 1 year after the addition of ethanol when stored closed at room temperature (15– 25°C).

- B. Buffer AL and ethanol (96–100%) are added in the same step. Buffer AL and ethanol can be premixed and added together in one step to save time when processing multiple samples.
- C. Add 90 ml ethanol (96–100%) to the bottle containing 86 ml Buffer AL or 260 ml ethanol to the bottle containing 247 ml Buffer AL and shake thoroughly. Mark the bottle to indicate that ethanol has been added. (Please note that, for purification of DNA from animal blood, Buffer AL must be used without ethanol. Buffer AL can be purchased separately if the same kit will be used for purification of DNA from animal blood.) Buffer AL is stable for 1 year after the addition of ethanol when stored closed at room temperature.

#### **Important points before starting**

- If using the DNeasy Blood & Tissue Kit for the first time, read “Important Notes” (Handbook, page 15 [available at: <http://www1.qiagen.com/literature/handbooks/literature.aspx?id=1000641>])
- All centrifugation steps are carried out at room temperature (15–25°C) in a microcentrifuge.
- Vortexing should be performed by pulse-vortexing for 5–10 s.

#### **Things to do before starting**

- Buffer ATL and Buffer AL may form precipitates upon storage. If necessary, warm to 56°C until the precipitates have fully dissolved.
- Buffer AW1 and Buffer AW2 are supplied as concentrates. Before using for the first time, add the appropriate amount of ethanol (96–100%) as indicated on the bottle to obtain a working solution.
- Preheat a thermomixer, shaking water bath, rocking platform, or dry block heater to 56°C for use in step 2.

#### **Protocols for Tissue Lysis and DNA Extraction**

- A. Cut ~10 mg\* tissue, and place in a 1.5 ml sterile microcentrifuge tube. Include a negative extraction control with each set of extractions. The negative extraction control is treated exactly the same as sample tubes except that it does not contain

any fish tissue. **Note\***: Tissue does not need to be accurately weighted but it is essential to only take a small amount of tissue with this extraction kit to obtain the optimal DNA yields for PCR.

- B. Add 50  $\mu$ l of buffer ATL and 5.56  $\mu$ l proteinase K. Mix thoroughly by vortexing, and incubate at 56°C until the tissue is completely lysed. Vortex occasionally during incubation to disperse the sample, or place in a thermomixer, shaking water bath, or on a rocking platform. Lysis time varies depending on the type of tissue processed. Lysis is usually complete in 1–3 h. If it is more convenient, samples can be lysed overnight; this will not affect them adversely. After incubation the lysate may appear viscous, but should not be gelatinous as it may clog the DNeasy Mini spin column. If the lysate appears very gelatinous, see the “Troubleshooting Guide”, page 47, for recommendations.
- C. Vortex for 15 s. Add 55.6  $\mu$ l Buffer AL to the sample, and mix thoroughly by vortexing. Then add 55.6  $\mu$ l ethanol (96–100%), and mix again thoroughly by vortexing. It is essential that the sample, Buffer AL, and ethanol are mixed immediately and thoroughly by vortexing or pipetting to yield a homogeneous solution. Buffer AL and ethanol can be premixed and added together in one step to save time when processing multiple samples. **Note**: A white precipitate may form on addition of Buffer AL and ethanol. This precipitate does not interfere with the DNeasy procedure.
- D. Pipet the mixture from step 3 (including any precipitate) into the DNeasy Mini spin column placed in a 2 ml collection tube (provided). Centrifuge at 6000  $\times$  g (8000 rpm) for 1 min. Discard flow-through and collection tube. Place the DNeasy Mini spin column in a new 2 ml collection tube (provided), add 140  $\mu$ l Buffer AW1, and centrifuge for 1 min at 6000  $\times$  g (8000 rpm). Discard flow-through and collection tube.
- E. Place the DNeasy Mini spin column in a new 2 ml collection tube (provided), add 140  $\mu$ l Buffer AW2, and centrifuge for 3 min at 20,000  $\times$  g (14,000 rpm) to dry the DNeasy membrane. Discard flow-through and collection tube. **Note**: It is important to dry the membrane of the DNeasy Mini spin column, since residual ethanol may interfere with subsequent reactions. This centrifugation step ensures that no residual ethanol will be carried over during the following elution. Following the centrifugation step, remove the DNeasy Mini spin column carefully so that the column does not come into contact with the flow-through, since this will result in carryover of ethanol. If carryover of ethanol occurs, empty the collection tube, then reuse it in another centrifugation for 1 min at 20,000  $\times$  g (14,000 rpm). Conversely, samples can be dried at 37°C for 30 minutes to remove excess alcohol and also to facilitate DNA elution.

- F. Place the DNeasy Mini spin column in a clean 1.5 ml microcentrifuge tube (not provided), and pipet 50 µl Buffer AE (warmed to 37°C) directly onto the DNeasy membrane. Incubate at room temperature for 1 min, and then centrifuge for 1 min at 6000 x g (8000 rpm) to elute.

### 3. POLYMERASE CHAIN REACTION – COI AMPLIFICATION

**Goal:** To amplify ~700 bases starting near the 5' end of the COI mitochondrial gene suitable for sequencing.

**Criteria for Success:** A sample that produces a single band of ~700 bases in size as visualized on a pre-cast agarose gel. Each set of reactions also requires two negative controls, one which consists of the PCR cocktail with no additional DNA template added and another with an addition of the extraction negative control. These samples have to be negative on the agarose gel (no band produced).

#### 3A. Consumables & Equipment

- a. Molecular grade water (dd H<sub>2</sub>O); e.g., Invitrogen Catalog No. 10977023
- b. D – (+) – Trehalose dihydrate; e.g., Sigma Catalog No. 90210-50g (BioChemika)
- c. 10X PCR Buffer, minus Mg; e.g., Invitrogen Catalog No. 10966-034
- d. 50 mM Magnesium Chloride; e.g., Invitrogen Catalog No. 10966-034
- e. Deoxynucleotide Solution Mix; New England Biolabs Catalog No. N0447L
- f. Oligonucleotide Primers; Integrated DNA Technologies, USA.
- g. Platinum Taq DNA Polymerase; Invitrogen Catalog No. 10966-034
- h. Pipetter 2-20 µl volume; e.g. Rainin classic pipet; Catalog No. PR-20 (tips- GP-20F)
- i. Pipetter 20-200 µl volume; e.g., Rainin classic pipet; Catalog No. PR-200 (tips- GP-200F)
- j. Pipetter 0.1-2 µl volume; e.g., Rainin classic pipet; Catalog No. PR-2 (tips- GP-10F)
- k. 8-Channel pipetter 0.5-10 µl volume; e.g., Rainin classic pipet; Catalog No. L8-10 (tips- GP-10F)
- l. 8-Channel pipetter 20-200 µl volume; e.g., Rainin classic pipet; Catalog No. L8-200 (tips- GP-10F)
- m. Thermocycler; e.g., Eppendorf Mastercycler® ep gradient S Thermocycler; Eppendorf Catalog No. 950010045
- n. Adhesive PCR film; e.g., Thermo-scientific AB0558
- o. 8 well strip tubes; e.g., Applied Biosystems N801-0838
- p. 96- well PCR plate; e.g., Catalog No. MPS-499 from Phenix Research
- q. Gel Doc gel documentation system; Bio-Rad

- r. Pre-cast agarose gel; e.g., 2% E-Gel® 96 Invitrogen Catalog No. G7008-02
- s. E-Base® Integrated power supply; Invitrogen Catalog No. EB-M03
- t. AirClean® Systems Ductless PCR Workstation Fisher Scientific Catalog No. 209 36 099 3859
- u. Positive Control fish DNA\* (if required) **Note\***: Any purified fish DNA from previous analyses can be used as a positive control. An SOP for producing positive control DNA will be provided separately.

### 3B. Polymerase Chain Reaction Recipe

Tailed M13 COI primers (modified from Baldwin et al. 2009):

#### FISHCOILBC\_ts:

CACGACGTTGTAAAACGACTCAACYAATCAYAAAGATATYGGCAC

#### FISHCOIHBC\_ts:

GGATAACAATTTACACAGGACTTCYGGGTGRCCRAARAATCA

#### PCR REAGENTS

	Each well (µl)	96-well plate (µl)
1. 10% trehalose	6.250	625.0
2. dd H <sub>2</sub> O	2.000	200.0
3. 10X buffer	1.250	125.0
4. 50 nM MgCl <sub>2</sub>	0.625	62.5
5. 10 µM primer A	0.125	12.5
6. 10 µM primer B	0.125	12.5
7. 10 mM dNTPs	0.062	6.2
8. Platinum Taq (5 U/µl)	0.060	6.0
<b>Total</b>	<b>10.50</b>	<b>1050.00</b>

### 3C. Protocols for Polymerase Chain Reaction (example for a 96 well plate)

- A. Add all 8 reagents to a single 1.5-2 ml microcentrifuge tube in the volumes listed under the “96 well plate”.
- B. This forms the “cocktail” of reagents for the PCR reaction. Vortex the cocktail vigorously or mix well by pipetting (Note: vortexing will cause liquid to be trapped on the cap of the tube. Before proceeding to the next step make sure that this liquid is returned to the rest of the cocktail by a 15 second spin in a mini centrifuge).

- C. Aliquot 130  $\mu$ l of cocktail into each well of an 8 well strip tube using the single-channel 200  $\mu$ l pipette. Use the same tip for each transfer in this step.
- D. Transfer 10.5  $\mu$ l of cocktail from the last row to the empty wells in each row of the 96-well plate using the 8-channel 200  $\mu$ l pipette. Use the same tips for each transfer in this step.
- E. Aliquot 1  $\mu$ l of appropriate DNA template (or positive control DNA) to each well of the 96-well plate.
- F. Place the plate in a thermocycler block with a plastic mat on top of the plate and close the lid of the thermocycler. **Note:** the plastic mat is only required by some thermocyclers. Start the program: Thermal conditions are 94°C for 2 min, 35 cycles of 94°C for 30 sec, 55°C for 40 sec, and 72°C for 1 min, with a final extension at 72°C for 10 min.

### 3D. PCR Product Check

**NOTE:** This section is written for an E-Base<sup>®</sup> 48 or 96 well system. If labs have gel systems other than the one listed, follow manufacturer's instructions. Samples should be run for 5-6 min in a 1-2% agarose gel.)

- A. Plug the Mother E-Base into an electrical outlet. Press and release the 'pwg/prg' (power/program) button on the base to select program EG. Select a run time of 6 min by pressing the 'time' button.
- B. Remove gel from the package and remove plastic comb from the gel. Slide gel into the two electrode connections on the Mother E-Base.
- C. Load 16  $\mu$ l of dd H<sub>2</sub>O into each well
- D. Load 4  $\mu$ l of each PCR product into a different well
- E. To begin electrophoresis, press the 'pwd/prg' button. The red light should change to green.
- F. At the end of run (signaled with a flashing red light and rapid beeping), press and release the power button.
- G. Remove the gel cassette from the base and capture a digital image of the gel with the Bio-Rad Gel Doc system.
- H. Arrange lanes and manipulate image as necessary using the Invitrogen E-Editor

software (available at <http://www.invitrogen.com/egels>).

## 4. PCR Cleanup

**Goal:** To produce an amplicon free of extra dNTPs and excess primers that might interfere with the sequencing reaction.

**Criteria for Success:** Examine the PCR thermocycler to ensure that the run has completed at the correct temperatures and there are no error messages.

### 4A. Consumables & Equipment

- a. ExoSAP-IT- USB Corporation; Cleveland, OH Cat. No. 78201
- b. Thermocycler; e.g., Eppendorf Mastercycler® ep gradient S Thermocycler; Eppendorf Catalog No. 950010045

### 4B. Protocols for PCR Cleanup

- A. Add 2 µl of ExoSAP-IT to 5 µl of PCR product
- B. Incubate this mixture at 37°C for 15 min followed by 80°C for 15 min in a thermocycler.

**Note:** products can be measured on a Nanodrop ND 1000 Spectrophotometer to quantify DNA at this step. Negative controls should read ~0 ng/ul

## 5. CYCLE SEQUENCING REACTION

**Goal:** To attach fluorescently labeled ddNTPs using PCR-amplified template DNA.

**Criteria for success:** Examine the thermocycler to ensure that the run has completed at the correct temperatures and there are no error messages.

### 5A. Consumables & Equipment

- a. Molecular grade water (dd H<sub>2</sub>O); e.g., Invitrogen Catalog No. 10977023
- b. D – (+) – Trehalose dihydrate; e.g., Sigma Catalog No. T9531-100g
- c. BigDye® Terminator v3.1 Cycle Sequencing Kit; Applied Biosystems Catalog No. 4337457
- d. 5X Sequencing Buffer (400 mM Tris-HCl pH 9.0 + 10 mM MgCl<sub>2</sub>); included in



- BigDye Terminator v.3.1 Cycle Sequencing Kit
- e. Oligonucleotide Primer; Integrated DNA technologies, USA
  - f. Pipetter 2-20 µl volume; e.g., Rainin classic pipet; Catalog No. PR-20 (tips- GP-20F)
  - g. Pipetter 20-200 µl volume; e.g., Rainin classic pipet; Catalog No. PR-200 (tips- GP-200F)
  - h. Pipetter 0.1-2 µl volume; e.g., Rainin classic pipet; Catalog No. PR-2 (tips- GP-10F)
  - i. 8-Channel pipetter 05-10 µl volume; e.g., Rainin classic pipet; Catalog No. L8-10 (tips- GP-10F)
  - j. 8-Channel pipetter 20-200 µl volume; e.g., Rainin classic pipet; Catalog No. L8-200 (tips- GP-10F)
  - k. Semi skirted 96 well sequencing plate; e.g., Phenix Research MPS-3590
  - l. AirClean® Systems Ductless PCR Workstation; Fisher Scientific Catalog No. 36 099 3859
  - m. 8 well strip tubes; e.g., Applied Biosystems N801-0838
  - n. Thermocycler; e.g., Eppendorf Mastercycler® ep gradient S Thermocycler; Eppendorf Catalog No. 950010045

### 5B. Cycle Sequencing Reaction Recipe

Reagents	Each well (µl)	96-well plate (µl)
1. 10 µM Primer M13F (or M13 R)	1	104
2. 5X SEQ buffer*	1.875	195
3. BigDye	0.25	26
4. dd H2O	0.875	91
5. 10% Trehalose	5	520
<b>Subtotal</b>	9	936
DNA template	1	n/a
<b>Total</b>	10	n/a

**\*Note:** 5X Sequencing buffer is: 400 mM Tris-HCl pH 9.0, 10 mM MgCl<sub>2</sub> or 5X ABI sequencing buffer.

### M13 Primers (Steffens et al. 1993)

M13F (-29) CACGACGTTGTAACGAC

M13R GGATAACAATTCACACAGG

### 5C. Protocols for Cycle Sequencing

- A. Add reagents 1-5 to a single 1.5 ml microcentrifuge tube in the volumes listed under the “96 well plate”. This forms the “cocktail” of reagents for the PCR

reaction Repeat this step using the other primer. The end result is 2 cocktails, one containing the forward primer and the other containing the reverse primer. **Note:** BigDye is light sensitive and should remain in the freezer except when in use. Never leave it out for longer than a few minutes.

- B. Vortex the cocktail vigorously. **Note:** Vortexing will cause liquid to be trapped on the cap of the tube. Before proceeding to the next step make sure that the liquid has been returned to the rest of the cocktail by quick spin in a mini-centrifuge.
- C. Aliquot 130  $\mu$ l of cocktail into each well of an 8 well strip tube using the single-channel 200  $\mu$ l pipette. Use the same tip for each transfer in this step. Use the same tip only for those transfers involving a single cocktail (i.e. before transferring the second cocktail you must change the tip).
- D. Transfer 9.0  $\mu$ l of cocktail from the last row into each row of the 96-well plate using the 8-channel 10  $\mu$ l pipette for both cocktails. Use the same tip only for those transfers involving a single cocktail (i.e. before transferring the second cocktail you must change the tip).
- E. Aliquot 1-1.5  $\mu$ l of PCR product into each well of the 96-well plate. Use the 10  $\mu$ l 8- channel pipette to do this row by row. **Note:** change the tips after each row.
- F. Seal the 96-well plate.
- G. Place the plate in a thermocycler block.
- H. Start the program Seq3.1: 96°C for 2 min, 30 cycles of 96°C for 30 sec, 55°C for 15 sec and 60°C for 4 min, followed by a 4°C hold.
- I. After the program has finished, stop the thermocycler and store the plates at 4°C in a dark box to avoid degradation of the light-sensitive sequencing products.

## 6. SEQUENCING REACTION CLEANUP

**Goal:** To remove unincorporated dye terminators and salts from sequencing reactions so that they will not interfere with the base pair determination of the fragment.

**Criteria for success:** A 96-well sequencing plate containing 10-15  $\mu$ l of cleaned sequencing reaction product after centrifugation. Success of unincorporated dye removal\* cannot be assessed until Step 8: Post-Sequencing Analysis.

**\*Note:** Incomplete unincorporated dye removal will cause “dye blobs” to appear in the

raw ABI file chromatograms that will interfere with proper base assignments, and therefore reduce the quality of the processed DNA sequence. Several websites exist that are devoted to the troubleshooting of DNA sequencing. A helpful example is <http://www.etonbio.com/troubleshooting.php>.

#### **6A. Consumables & Equipment:**

- a. Edge Bio PERFORMA DTR V3 96-well short plate kit; Edge Bio Catalog No. 89939. **Note:** for lower sample numbers, Edge Bio PERFORMA DTR gel filtration cartridges (Edge Bio Catalog No. 42453) can be used in combination with a microcentrifuge capable of 850 x g. Instructions for both provided.
- b. Centrifuge capable of spinning 96-well plates at 850 x g; e.g., Sorvall Evolution RC, Thermo-scientific
- c. Semi skirted 96 well sequencing plate; e.g., Phenix Research MPS-3590
- d. 8-Channel pipetter 0.5-10 µl; e.g., Fisher Scientific Catalog No. 13-688-507
- e. 1-10 µl filter tips; e.g., Fisher Scientific Catalog No. CS004807
- f. Flat bottom waste plate; Costar Catalog No. 9017 (NOTE: Included in the Edge Bio kit No. 42453)
- g. Hi-Di™ Formamide; Applied Biosystems Catalog No. 4311320

#### **6B. Protocols for Sequencing Cleanup (96 well plate format):**

- A. Prep an Edge Bio PERFORMA DTR V3 96-well short plate by first removing the bottom adhesive tape, followed by the top. Take care to keep the plate horizontal to avoid losing any gel.
- B. Stack the V3 96-well short plate on top of a 96 well waste plate.
- C. Place the stacked plates (2) on a cushioned centrifuge carrier.
- D. Centrifuge for 3 minutes at 850 × g.
- E. Remove from centrifuge and discard excess water from waste plate.
- F. Carefully Transfer products directly into the prepped Edge Bio PERFORMA DTR V3 96-well short plate. **Note:** Pipet slowly, and do not touch the sides of the wells.
- G. Place V3 96-well short plate on top of a 96-well semi-skirted capillary plate containing 10 µl of Hi-Di™ Formamide in each well, and then back into a waste plate.
- H. Place the stacked plates (3) on a cushioned centrifuge carrier, taking care that they will spin without hindrance.

- I. Centrifuge for 2 minutes at 850 x g. **Note:** Instead of 5 minutes as the manufacturer suggests, because of specific instructions from Applied Biosystems.
- J. Retain eluate in 96-well semi-skirted plate and discard the V3 96-well short plate. **Note:** Some residual liquid (3-5 µL) will come through even in the empty wells according to the manufacturer so that the final volume in the wells will be 10-20µL.

#### **6B alternate: Protocols for Sequencing Cleanup (single tube format):**

- A. Centrifuge the Performa Gel Filtration Cartridge for 3 minutes at 850 x g in a microcentrifuge.
- B. Transfer the cartridge to the provided 1.5-ml microcentrifuge tube and add the sample to the packed column. Be sure the fluid runs into the gel.
- C. Close the cap and centrifuge for 3 minutes at 850 x g. Retain eluate.
- D. Transfer 10 µl of eluate to a 96-well semi-skirted capillary plate containing 10 µl of Hi-Di™ Formamide in each well. Spin plate down in a centrifuge before loading on sequencer.

## **7. SEQUENCING**

**Goal:** Determine the accurate base pair composition of the COI amplicon.

**Criteria for Success:** Generation of an AB1 file suitable for post sequencing analysis, though determination of the accurate base pair composition cannot be assessed until successful completion of Step 8: Post-Sequencing Analysis.

#### **7A. Consumables & Equipment:**

- a. Pop-7™ Polymer\* for 3730xl DNA Analyzers; Applied Biosystems Catalog No. 4335611
- b. 3730xl DNA Analyzer Capillary Array, 50 cm; Applied Biosystems Catalog No. 4331246
- c. Running Buffer, 10X for 3730xl DNA Analyzers; Applied Biosystems Catalog No. 4335613

**\*Note:** Pop-6™ Polymer; Applied Biosystems Catalog No. 402837 was also found to provide acceptable results for barcode analysis (Handy et al. 2011).

#### **7B. Protocols for Sequencing:**

- A. Samples can now be put directly on the ABI 3730 sequencer. **Note:** Other types of sequencers can be used successfully but this was the instrument used in the SLV of the barcoding method. Follow instructions for particular sequencer in your individual laboratory.
- B. Import the plate record, and run the instrument using protocol default\_50cm and the analysis protocol 3730BDTv3-KB-DeNovo\_v5.2

## 8. POST SEQUENCING ANALYSIS

**Goal:** To process the sequence from the AB1 file into a usable unit for comparison with sequences from a database of authenticated standards.

**Criteria for Success:** Bi-directional sequences, of at least 500 bp in length, with <2% ambiguous bases in the contig, or a single read with a 98% HQ (98% high quality bases).

### 8A. Consumables & Equipment:

- a. Geneious Pro; Biomatters Ltd., Auckland, New Zealand
- b. Degenerate bony fish barcoding sequence, FISHREF08a:  
CCTTTATCTAGTATTTGGTGCCTGAGCCGGAATAGTAGGCACAGCCCTAAGCCTACTA  
ATTCGAGCTGAACTAAGCCAACCTGGCGCCCTTCTAGGGGACGACCAAATTTATAAT  
GTAATCGTAACTGCCCACGCCTTTGTAATAATTTCTTTATAGTAATACCAATTATGAT  
TGGAGGCTTTGGAACTGACTAATCCCCCTAATGATTGGGGCCCCCGACATGGCCTT  
CCCTCGAATAAACAAACATAAGCTTTTGACTTCTTCCCCCTTCTTCTTCTTCTCCTAGC  
ATCCTCTGGAGTTGAAGCCGGGGCCGGAACAGGATGAACAGTTTACCCCCCTCTAGC  
AGGAAACCTAGCCCACGCAGGAGCCTCTGTAGACCTAACAAATTTCTCCCTTCATCTA  
GCAGGAATYTCCTCAATTCTAGGGGCAATTAACCTTTATTACAACAATTCTTAACATGA  
AACCCCCAGCCATTTACAATACCAACACCCCTATTTGTTTGAGCTGTCCTAATTACA  
GCCGTCCTTCTTCTATCCCTCCCAGTCCTTGCTGCTGGCATTACAATGCTTCTCAC  
AGACCGAAACCTAACACAACCTTCTTTGACCCTGCAGGAGGAGGAGACCCATTCT  
GTACCAACACCTATTC

### 8B. Protocols for Post Sequencing Analysis:

- A. Create folder in Geneious Pro\*, and name it with appropriate regulatory (e.g. FACTS) sample number **\*Note:** This is a deviation from the published SLV method where a combination of programs were used for this step. Other types of editing software may also provide acceptable results but in the end an unknown sequence with either >500 bases of bidirectional coverage with <2% ambiguous

bases or single reads >500 bases with >98% high quality must be obtained for use in this analysis.

- B. Import AB1 files generated by the sequencer into the newly created folder.
- C. Import Fish reference sequence FISHREF08a (in fasta format) into the folder.
- D. Highlight all AB1 files and the Fish reference sequence and then click “Assembly”
  - 1. Under Data
    - a. Choose “Assemble to reference” and be sure the reference sequence is in the tab.
    - b. Assemble by a delimiter used in the sample name (e.g., “\_”) that allows the two parts of the bidirectional sequence to be combined.
  - 2. Under Method
    - a. Set the sensitivity to Highest Sensitivity/Slow
    - b. Set the Fine Tuning to Maximum (Slowest)
  - 3. Under Trim Sequences
    - a. Have Trim sequences clicked. (Under options have: “remove new trimmed regions from sequences” clicked)
    - b. Have Trim primers clicked
    - c. Allow 5 mismatches
    - d. Minimum Match Length set to 5
    - e. Error Probability Limit set at 0.05
    - f. Click Trim 5’ End
    - g. Click Trim 3’ End
    - h. Click on Maximum length after trim set to 655
    - i. Click OK
  - 4. Under Results
    - a. Change Assembly Name to the “name of the regulatory sample “ Assembly
    - b. Click Save Assembly report, save in sub-folder, save contigs
- E. In the assembly file, select the Minimum Sequence Length heading to sort sequences by length. If any sequences are under 500 bp and < 98 %HQ, repeat the assembly step on the AB1 files without these sequences.
- F. Highlight all assembled files, right click and choose “Generate Consensus sequence”

1. Threshold set at Highest Quality
  2. Assign Quality set on Total (which is the sum of the quality of contributing bases minus non-contributing bases)
  3. Click Trim to reference sequence.
  4. Click Ok
- G. After an assembly file and consensus file are created for the sample
1. Check the ambiguities by examining the consensus file and looking at the tab for ambiguities (must be less than 2%). **Note:** In order to determine the percent ambiguities, divide the number of ambiguities listed in the tab by the post-trim length of the sequence and then multiply by 100.
  2. For any samples that are left as single reads- be sure the sample has a 98 %HQ (98% high quality bases) or better.
- H. Import the appropriate FDA Reference Standard Sequence Library alignment fasta file\*.

\* The FDA Reference Standard Sequence Library will be provided to FDA Office of Regulatory Affairs (ORA) field analysts through the FDA Intranet CFSAN site. This database will be updated as more standards are added to the library. The database will be dated to distinguish it from previous versions. Analysts should check periodically to assure they are using the current database. In the event of any critical changes, CFSAN will notify ORA analysts directly. These sequences will also be available individually through the public CFSAN website along with FDA approved market names, authentication information, etc. In the event that a sequence from an unknown tissue sample does not match any sequence in the FDA database, the sequence can be screened against other publically available sources such as GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>) and BOLD ([www.boldsystems.org](http://www.boldsystems.org)). These identifications should be considered presumptive unless the sequences were derived from specimens with proper authoritative taxonomic identification. FDA will only make regulatory decisions based on identifications using adequately authenticated standards. Inclusion of sequences from publically available databases should be noted in the FDA analytical worksheet.

1. Click on the small box in the upper left hand corner of the Geneious Pro folder next to the tab "name" to organize the files based on type.

2. Highlight the FDA fish database alignment and the organized consensus sequences
- I. Form an alignment with both the FDA fish database alignment and the consensus sequences of the unknowns by clicking the alignment button.
  - J. Align sequences in Geneious using the “Muscle Alignment” tab with default settings.
  - K. Using this file, construct a UPGMA tree\* (using Geneious treebuilder) with a Jukes-Cantor genetic distance model (no resampling). **Note:** (A neighbor-joining tree would also be appropriate but with variable sequence lengths the branches may be less smooth.)

**\*Note:** UPGMA or neighbor-joining analysis with distance thresholds, using a generalized cutoff of 2% sequence divergence, can be used in most cases to successfully delimit commercial species. More elaborate statistical methods to define species boundaries are also available (Kerr et al. 2009). For example, a character-based approach was recently used to successfully distinguish some of the closely related commercial species of tuna (Lowenstein et al. 2009). The most common application of DNA Barcode analysis at FDA involves the confirmation of product labeling according to their acceptable market names, found in the FDA Seafood List (FDA 2010). For this application, 2% sequence divergence is typically sufficient to provide resolution. Note, if 2 sequences are >2% divergent and the sequence quality standards provided here are followed, they are likely not the same species, but select examples exist where 2 closely related but distinct species have <2 % sequence divergence in this region of the COI gene. In some cases they can still be distinguished with this method, in other cases they cannot. Few examples exist in the FDA Seafood List where 2 closely related species that cannot be distinguished with this method have different approved market names.

- L. After examination of the tree, select appropriate species to include in a smaller subset (these should include several of the most closely related species, which MUSCLE will put in close proximity to each other, and whatever the product was labeled as if the product was misbranded). Click on the Alignment View tab (to the right of the Tree View tab), select the samples to include in the subset and click on the Extract button. Select the option “Extract region as alignment”.
- M. Using the extracted file, construct a UPGMA or neighbor-joining tree (using Geneious treebuilder) with a Jukes-Cantor genetic distance model (with no resampling). To save and print the tree, make sure the Tree View tab is selected, then go to File, Save as Image File and save as a PDF, under More Options select “Print all labels”. Include this tree\* in the FDA analytical worksheet. Resampling can be done to verify statistical support if desired.



**\*Note:** The trees will only contain the latin scientific name for each standard. Please consult the FDA Seafood List at [http://www.accessdata.fda.gov/Scripts/search\\_seafood/index.cfm?other=complete](http://www.accessdata.fda.gov/Scripts/search_seafood/index.cfm?other=complete) or the FDA Reference Standard Sequence Library page for additional information, such as FDA acceptable market name, common scientific name, etc.

N. In the alignment window, extract the unknown sample and the most closely related species into a new alignment by highlighting the two sequences and clicking on the "Extract" button (or by right-clicking to select "Extract regions"). Choose the "Extract region as alignment" option. The new alignment will appear in your Geneious folder. Scroll to the right to find the "% Pairwise Identity" heading. Include this number in the results section of FDA analytical worksheet as well as the % pairwise identity for the unknown sample and the species it was labeled as (generated in the same way) if this is different than the most closely related species from above.

## Literature Cited

- ABI. (2007). User Bulletin Part number 4362968 Rev. B.
- Baldwin, CC, Mounts JH, Smith DG, Weigt LA. (2009). Genetic identification and color descriptions of early life-history stages of Belizean *Phaeoptyx* and *Astrapogon* (Teleostei: Apogonidae) with comments on identification of adult *Phaeoptyx*. *Zootaxa* 1-22.
- Ewing B, Hillier L, Wendl MC, & Green P. (1998). Base-calling of automated sequencer traces using *Phred*. I. Accuracy assessment. *Genome Research* **8**, 175-185.
- FDA. (2010). The Seafood List: FDA Guide to Acceptable Market Names for Food Fish Sold in Interstate Commerce. Office of Food Safety, Division of Seafood Safety, FDA Center for Food Safety and Applied Nutrition, US Department of Health and Human Services.  
<http://www.fda.gov/Food/GuidanceComplianceRegulatoryInformation/GuidanceDocuments/Seafood/ucm113260.htm>
- Handy, SM, Deeds, JR, Ivanova, NV, Hebert, PDN, Hanner, R., Ormos, A, Weigt, L.A., Moore, M.M., Yancy, H.F. (2011). A single laboratory validated method for the generation of DNA barcodes for the identification of fish for regulatory compliance. *J. AOAC* **94(1)**: 201-210.
- Kerr, K, Birks, S, Kalyakin, M, Red'kin, Y, Koblik, E, Hebert, PDN. (2009). Filling the gap – COI barcode resolution in eastern Palearctic birds. *Frontiers in Zoology* **6**, 29.
- Lowenstein J H, Amato G, Kolokotronis S-O. (2009). The real *maccoyii*: Identifying tuna sushi with DNA barcodes – contrasting characteristic attributes and genetic distances. *PLoS One* **4 (11)**, e7866.
- Steffens DL, Sutter SL, Roemer SC. (1993). An alternate universal forward primer for improved automated DNA sequencing of M13. *Biotechniques* **15(4)**: 580-582.