Monitoring Allis and Twaite Shad: quality assurance and species identification using molecular techniques

Dr David Stone
Centre for Environment, Fisheries & Aquaculture Science

NRW Evidence Report No 53
About Natural Resources Wales

Natural Resources Wales is the organisation responsible for the work carried out by the three former organisations, the Countryside Council for Wales, Environment Agency Wales and Forestry Commission Wales. It is also responsible for some functions previously undertaken by Welsh Government.

Our purpose is to ensure that the natural resources of Wales are sustainably maintained, used and enhanced, now and in the future.

We work for the communities of Wales to protect people and their homes as much as possible from environmental incidents like flooding and pollution. We provide opportunities for people to learn, use and benefit from Wales' natural resources.

We work to support Wales’ economy by enabling the sustainable use of natural resources to support jobs and enterprise. We help businesses and developers to understand and consider environmental limits when they make important decisions.

We work to maintain and improve the quality of the environment for everyone and we work towards making the environment and our natural resources more resilient to climate change and other pressures.
Evidence at Natural Resources Wales

Natural Resources Wales is an evidence based organisation. We seek to ensure that our strategy, decisions, operations and advice to Welsh Government and others are underpinned by sound and quality-assured evidence. We recognise that it is critically important to have a good understanding of our changing environment.

We will realise this vision by:
- Maintaining and developing the technical specialist skills of our staff;
- Securing our data and information;
- Having a well resourced proactive programme of evidence work;
- Continuing to review and add to our evidence to ensure it is fit for the challenges facing us; and
- Communicating our evidence in an open and transparent way.

This Evidence Report series serves as a record of work carried out or commissioned by Natural Resources Wales. It also helps us to share and promote use of our evidence by others and develop future collaborations. However, the views and recommendations presented in this report are not necessarily those of NRW and should, therefore, not be attributed to NRW.
Report series: NRW Evidence Report
Report number: 53
Publication date: February 2015
Contract number: R007608
Contractor: Centre for Environment, Fisheries & Aquaculture Science
Contract Manager: Hatton-Ellis TW
Title: Monitoring Allis and Twaite Shad: quality assurance and species identification using molecular techniques
Author(s): Stone, David M
Technical Editor: Hatton-Ellis, Tristan W
Approved By: Jones, Peter S
Restrictions: None

Distribution List (core)
NRW Library, Bangor 2
National Library of Wales 1
British Library 1
Welsh Government Library 1
Scottish Natural Heritage Library 1
Natural England Library (Electronic Only) 1

Distribution List (others)
Leila Thornton, NRW
Heather Garrett, NRW
Alex Harding, NRW
Rhian Thomas, NRW
Liz Howe, NRW
Miran Aprahamian, Environment Agency
Kerry Walsh, Environment Agency
Colin Bean, SNH
Chris Mainstone, Natural England

Recommended citation for this volume:
Contents (This is automatic)

1. Crynodeb Gweithredol ............................................................................................................ 7
2. Executive Summary ................................................................................................................... 8
3. Introduction .............................................................................................................................. 9
   3.1. Background ............................................................................................................................ 9
   3.2. Project objectives .................................................................................................................... 10
       The key objectives of the current study were: .............................................................................. 10
4. Materials and Methods .......................................................................................................... 11
   4.1. Sample collection and sample processing ........................................................................... 11
   4.2. Primer design ......................................................................................................................... 11
   4.3. DNA extraction and amplification ....................................................................................... 11
   4.4. Sequencing ............................................................................................................................ 12
5. Results ....................................................................................................................................... 13
   5.1. Egg Collection Locations .................................................................................................... 13
   5.2. *Alosa*-specific Cyt b gene amplification ............................................................................ 13
   5.3. COI gene amplification ......................................................................................................... 13
   5.4. Sequence analysis .................................................................................................................. 17
6. Discussion .................................................................................................................................. 21
7. Conclusions ............................................................................................................................. 22
8. Acknowledgements ................................................................................................................ 22
9. References .................................................................................................................................. 22
10. Appendices ............................................................................................................................. 24
    10.1. Appendix A: Sampling procedures for Shad genetic work ................................................. 24
    10.1.1. Egg collection and storage for genetic work analysis ...................................................... 24
    10.1.2. CCW Shad Egg Monitoring Protocol ......................................................................... 25
    10.2. Appendix B: Cyt Bb gene sequences ............................................................................... 27
    10.2.1. Input file for the partial Cytb sequences obtained for the shad eggs collected on the River Tywi in 2014 ......................................................................................................................... 27
    10.3. Appendix C: BLAST results .............................................................................................. 31
    10.3.1. Blast results for partial COI gene sequence for the egg sample 2a/13 ............................ 31
    10.3.2. BLAST results for partial COI gene sequence for egg sample 2a/11 ............................. 40
Data Archive Appendix .................................................................................................................. 49
List of Figures

Figure 1: *Alosa* specific PCR amplicons generated for eggs samples collected from the River Tywi.. ...........................................................................................................................................15

Figure 2: COI generic PCR amplicons generated for eggs samples collected from the river Tywi....................................................................................................................................................16

Figure 3: Alignment of the partial COI gene sequence generated for 15 putative Shad eggs..................................................................................................................................................................19

Figure 4: Phylogenetic relationships between the partial COI gene sequences obtained for *Alosa fallax* and *Alosa alosa* eggs from the riverTywi .................................................................................................................................20
List of Tables

Table 1: Putative Alosa spp. eggs sampled from the River Tywi.........................................................14
1. Crynodeb Gweithredol

Mae gwangod (Alosa spp.) yn bysgod sy’n ymdebygu i benwaig ac yn cael eu diogelu gan y Gyfarwyddeb Gynfinoedd a’u monitro drwy ddefnyddio cic-samplu ar gyfer eu hwyau yn ystod y týmor sílio. Fodd bynnag, mae wyau rhywogaethau pysgod eraill sydd heb fod yn darged yn cael eu dal yn y ffordd hon ambell waith yn ogystal. Prif nod hwn oedd sicrhau ansawdd y rhaglen samplu wyau gwangod ym masnau afonydd Cymru (Gwy, Wysg a Thywi) drwy ddefnyddio technegau genetig er mwyn penderfynu ai gwangod oedd sampl cynrychioladol o wyau a gasglwyd.

Casglwyd 162 o wyau gwangod tybiedig o Afon Tywi, Sir Gaerfyrddin gan dîm maes a benodwyd gan Cyfoeth Naturiol Cymru (CNC). Roedd y samplau wyau wedi cael eu sefydlogi a’u storio mewn ethanol 95% a’u cludo i Labordy Cefas yn Weymouth. Oherwydd tywydd gwael, ni chasglwyd unrhyw wyau o Afonydd Gwy ac Wysg.

Tynnwyd DNA o wyau unigol a’u sgrinio drwy adwaith cadwyn polymeras (ACP) drwy ddefnyddio set ysgogi benodol ar gyfer Alosa spp. sy’n targedu’r genyn mitocondriaidd, Sytocrom b, er mwyn penderfynu a oeddyn nhw’n rhywogaethau gwangod. Dau sampl yn unig (1.2%) a oeddyn nhw’n nefyddol ar gyfer dilyniant targed Sytb sy’n benodol i Alosa drwy ddefnyddio ACP. Roedd yr un samplau yn nefyddol wrth ddefnyddio prawf ACP generig ar gyfer y genyn Sytocrom ocsidas I (SOI), ac yn dangos bod methiant y prawf penodol yn fwyaf tebygol yn ganlyniad i brinder o’r DNA targed yn y sampl. Cynhyrchodd y 160 o wyau a oedd yn weddill signal penodol ar gyfer Alosa spp.

Gwnaed dadansoddiad ymiant ar gynhyrchion mwyhau a gynhyrchwyd yn gyfochrog drwy ddefnyddio coctel ysgogi generig SOI er mwyn cadarnhau tarddiad y wyau a oeddyn nhw’n cynhyrchu arwyddion ACP gwannach sy’n benodol i Alosa. O’r 15 o wyau a brofwyd, cafodd dau ddilyniant gwahanol eu hadnabod ac roeddyn nhw’n rhannu 99% o unfathiant y niwcleotid gyda’i gilydd. Roedd y ddau ddilyniant yn tarddu o wangod, yn seiliedig ar unfathiant y niwcleotid. Roedd y dilyniantau genyn mitocondriaidd (SOI) ar gyfer y genyn Sytocrom o’r 12 wy a oedd yn weddill signal penodol o gyfochrog. Cyfrwyd un o’r 9 safle a gafoedd eu harchwilio ar Afon Tywi yn 2014 wedi cael eu cofnodi’i anghywir fel eu bod yn cefnogi silio gwangod drwy gael eu hadnabod yn y maes yn unig. Cadarnhawyd silio gwangod yn Nantgaredig, lle profodd samplau a gasglwyd yn 2013 i fod yn rhai pilcod. Mae’r canlyniadau hyn yn cymharol â phrincipiwl gyda’r wraith blanerol ac maen nhw’n cynyrrchol a gyfrann o lwyddiant.
2. Executive Summary

Shads (*Alosa* spp.) are herring-like fish that are protected by the Habitats Directive and monitored using kick sampling for their eggs during the spawning season. However, eggs of other non-target fish species are also sometimes caught in this way. The primary aim of this project was to quality assure the shad egg sampling programme on Welsh river basins (Wye, Usk and Tywi) by using genetic techniques to determine whether a representative sample of eggs collected were shad.

162 putative shad eggs were collected from the River Tywi, Carmarthenshire by the Natural Resources Wales (NRW) appointed field team. Egg samples were fixed and stored in 95% ethanol and forwarded to the Cefas Weymouth Laboratory. Due to bad weather no eggs were collected from the River Wye and Usk.

DNA was extracted from individual eggs and screened by polymerase chain reaction (PCR) using an *Alosa* spp.-specific primer set targeting the mitochondrial gene, Cytochrome b, to determine if they were shad species. Only two samples (1.2%) were negative for the *Alosa*-specific Cytb target sequence by PCR. The same samples were also negative when using a generic PCR assay for the Cytochrome oxidase I (COI) gene indicating that the failure of the specific assay was most likely the result of insufficient target DNA in the sample. All the remaining 160 eggs produced an *Alosa* spp. specific signal.

Sequence analysis was undertaken on amplification products generated in parallel using the COI generic primer cocktail to confirm the origin of eggs producing the weaker *Alosa*-specific PCR signals. Of the 15 eggs tested, two distinct sequences were identified sharing 99% nucleotide identity with each other. Both sequences were shad in origin based on the nucleotide identities. The mitochondrial (COI) gene sequences for three eggs from Habitat 4 at Penddaulwyn, Habitat 8 at White Mill, and Nantgaredig Bridge were *Alosa fallax* in origin, and the sequences obtained for the remaining 12 eggs from Habitat 4 at Penddaulwyn, Glantowylan, Habitat 8 at White Mill, Llandeilo Bridge and at the Tidal limit were *Alosa alosa* in origin.

Overall, the sequence data indicates that 100% of the eggs that could be identified were correctly identified as shad in origin, with the identity of two eggs being uncertain due to poor DNA quality. None of the 9 sites surveyed on the R.Tywi in 2014 would have been erroneously recorded as supporting shad spawning using field identification alone. Shad spawning was confirmed at Nantgaredig, where samples collected in 2013 had proved to be minnow. These results compare favourably with previous work and represent a very high success rate.
3. Introduction

3.1. Background

The twaite shad *Alosa fallax*, and the allis shad *Alosa alosa*, are clupeid fish once found in a large number of rivers in the south of England and Wales (Aprahamian & Aprahamian 1990). Although they were once common in the rivers such as the Thames, the populations have been reduced primarily through pollution and barriers to migration such as dams and weirs (Aprahamian *et al.* 2003, Maitland & Hatton-Ellis 2003). As a result, shad populations have declined to such an extent that they are protected under Annexe II of the Habitats Directive. Today, the principal strongholds are limited to the rivers of south-west Britain, including the Rivers Wye, Usk and Tywi in south Wales. All three rivers are designated as Special Areas of Conservation (SAC) for both species.

Monitoring shad in a cost-effective manner is challenging, and various approaches have been tried including catch records from anglers and netsmen, hydroacoustic fish counters and seine netting of juveniles (Aprahamian *et al.* 2003; Hillman 2003; Hillman *et al.* 2003; Noble *et al.* 2007). The Countryside Council for Wales (CCW) and the Environment Agency in Wales have carried out kick sampling for shad eggs during the spawning season, which is a simple and cost-effective technique that has provided good semi-quantitative information on shad spawning activity and distribution (Thomas & Dyson 2012a,b, Garrett *et al.* 2013). However, a weakness of this method is that fish eggs are not easily identified, and some eggs being sampled by the may be from non-target species. Variability in egg size has been reported, suggesting either that two shad species are involved, or that eggs of non-target taxa are being sampled.

In 2013, eggs from 12 sampling sites from the Wye, Usk and Tywi were analysed using molecular techniques. A total of 226 eggs were successfully genotyped. 85% of eggs sampled were shad in origin and the remaining non-*Alosa* eggs were identified as belonging to minnow (*Phoxinus phoxinus*) and chub (*Squalius cephalus*) (Hardouin *et al.* 2013). This suggests that although a high proportion of eggs are correctly identified, there is nevertheless the risk that some material is misidentified. In some cases, there were potential implications for the findings of the survey (Hardouin *et al.* 2013). Consequently, regular quality assurance of the monitoring using genetic screening is beneficial, especially where the results may be used to influence significant decisions.
3.2 Project objectives
The key objectives of the current study were:

- To quality assure the taxonomic identification of eggs collected by the kick sampling procedure; confirming that a subsample were either twaite shad *Alosa fallax*, or allis shad *Alosa alosa* in origin by analysis of the Cyt b polymerase chain reaction (PCR) assay (Alexandrino *et al.* 2006)

- To confirm the origin of any non-shad eggs by analysis of the cytochrome oxidase subunit 1 (COI) mitochondrial gene using methods described by Ivanova *et al.* (2007)
4. Materials and Methods

4.1. Sample collection and sample processing
During the spawning season (late May and early June 2014), NRW staff collected individual eggs by kick sampling from 15 sampling sites in the Tywi (Table 1) using a standard protocol (Appendix A). Due to bad weather and consequent high flows, no egg samples were obtained from any of the sampling sites on the Wye or the Usk.

Sampling staff were issued with clear instructions (Appendix A) and standard field equipment including pre-labelled 1.5ml Eppendorf tubes containing 95% ethanol. Briefly, all suspected shad eggs collected from each sampling site were placed carefully in the appropriate pots using a clean pair of forceps, taking care not to burst the egg. Eggs that appeared to be close to hatching were also not collected as these were likely to hatch in the alcohol and thereby posed an increased risk of cross contamination. New pots were used for each sampling site, and no than 30 eggs were placed in a single pot. Multiple pots for the same sample site were used if required. When sampling was completed, the lids of the individual sample tubes were sealed using the parafilm to reduce the risk of ethanol leakage and/or evaporation. Samples were maintained at 4°C and sent by overnight courier to Cefas in a cool box.

4.2. Primer design
The primers used in analysis are those already published (Alexandrino et al. 2006, Ivanova et al. 2007).

The *Alosa* genus-specific primers, alocytf1 (CCTTCTAACATTTGACTGATG) and alocytr1 (AGGATTGTGGCTCTGCAATTAC) were used to amplify a partial fragment of the mtDNA cytochrome b gene (Alexandrino et al. 2006).

A cocktail of 4 primers (C_FishF1t1/C_FishR1t1), VF2_t1 (TGTTAAACGAGGCAGTCAACACAACCACAAAGACACATTGCAC), FishF2_t1 (TGTTAAGACGGTGCCATGCGACTAATCATACAAAGATATCGGCAC), FishR2_t1 (CAGGAAACAGCTATGACACTTTCAAGGTGACCAGAATTCAGAA) and FR1d_t1 (CAGGAAACAGCTATGACACTCTCAGGGTGTCGGCACAARAAAYCARAA) was used to amplify a partial COI sequence for confirmation of the species by sequence analysis (Ivanova et al. 2007).

4.3. DNA extraction and amplification
Individual eggs were examined using a Leica M125 dissecting microscope for the presence of an embryo. Individual eggs containing an embryo were then digested overnight in 500 μl of ATL buffer (Qiagen) containing proteinase K and the DNA extracted from 200 μl of the digest using the DNA Investigator Kit and the Universal BioRobot (Qiagen) following the QIAamp DNA tissue UNIV rcv31 extraction protocol. DNA from individual eggs was eluted in a 50 μl volume in a 96 well format.

Amplifications were performed in a 96 well format using the Cyt b (alocytf1 / alocytr1) and CoxI primer cocktail (C_FishF1t1 / C_FishR1t1). PCR reactions for the Cyt b assay were performed in a 50 μl reaction volume consisting of 1x GoTaq flexi
buffer (Promega, UK), 2.5 mM MgCl₂, 1 mM dNTP mix, 50 pmol of the forward and reverse primers, 1.25 units of GoTaq DNA Polymerase (Promega, UK) and 2.5 µl of the purified DNA template. The reaction mix was overlaid with mineral oil and after an initial denaturing step (5 min at 95°C), was subjected to 40 temperature cycles (1 min at 95°C, 1 min at 60°C and 1 min at 72°C) in a Peltier PTC-225 thermal cycler followed by a final extension step of 10 min at 72°C. A negative control extraction and amplification was included for every 10 eggs processed. Conditions used for the COI gene assay were the same as above, with the exception that the annealing temperature was reduced from 60°C to 48°C according to the published protocol.

To determine which of the eggs were from shad, 15 µl of the reaction products generated using the genus-specific and universal COI primer sets were visualised on 1.5% agarose gels stained with ethidium bromide. By resolving the reaction product from both the Cyt b and CoxI assays it was possible to identify samples that failed to generate products from both assays due to problems with the integrity of the DNA sample.

4.4. Sequencing
COI gene sequence analysis was applied to those samples that produced a product with COI assay only, together with samples that produced only weak products when using the Alosa spp.-specific Cyt b assay. The latter were selected to rule out the possibility of cross reactivity between the primers used in the Alosa spp.-specific Cyt b assay and non-target fish species.

PCR products generated using the COI primers cocktail were extracted and purified by ethanol precipitation. Both DNA strands of the amplicon were sequenced using the ABI PRISM BigDye terminator cycle sequencing system (Life Technologies) and the M13 primers corresponding to the tag sequences on the COI primers used in the initial amplification. Sequencing reactions were analysed on an ABI 3130 genetic analyser. A consensus sequence (with primer derived sequences removed) was determined using Sequencer software (Gene Codes Corporation, Ann Arbor, MI) and the origin of the amplicon sequence identified using the Basic Local Alignment Search Tool (BLAST) facility available at the National Centre for Biotechnology Information (NCBI).

Multiple sequence alignments and phylogenetic analysis were performed using a 242 nucleotide partial COI gene sequence obtained for the 15 of the eggs that produced weak amplification products when using the Alosa spp.-specific primer set. A partial COI gene sequence from the American shad, Alosa sapidissima (KC015147) was used as an outgroup. Multiple alignments were performed using Clustal W (Thompson et al. 1997) with the following Clustal parameters: a gap opening penalty of 15 and gap extension penalty of 6.66. Phylogenetic analyses were conducted using MEGA version 4 (Tamura et al. 2007). The neighbour-joining tree was constructed using a maximum composite likelihood model, and the robustness of the tree was tested using 1000 bootstrap replicates.
5. Results

5.1. Egg Collection Locations
A total of 162 putative *Alosa* spp. eggs were collected from the River Tywi for genetic analysis (Table 1). Eggs were obtained from 9 sites, Penddauwyn, Penddauwyn Habitat 4 and Habitat 5, Glantowylan, White Mill Habitat 8, Nantgaredig Bridge, Cothi confluence, Llandeilo Bridge and the Tidal limit. No eggs were obtained from Cothi Bridge, Dryslwyn, Cilsan Bridge, Manordeilo, Llanwrda and Llwynjack, Llanegwad, Llandovery and Dolauhirion were not sampled. Only three sites (Nantgaredig Bridge, White Mill and Cothi Confluence) were sampled in 2013 (Hardouin et al. 2013); all of these were resampled in 2014.

5.2. *Alosa*-specific Cyt b gene amplification
DNA was extracted from individual eggs and screened using the *Alosa* spp.-specific primers as described above, and the PCR products were resolved by agarose gel electrophoresis (Figure 1). In the majority of cases a strong PCR product was obtained indicating that the eggs were *Alosa* spp. in origin. However, two eggs (0.61%) did not produce an *Alosa*.-specific product; sample 2A/15 from Habitat 4 at Penddauwyn and sample 10/3 from Llandeilo Bridge suggesting these egg samples represent a different fish species. A further 19 eggs produced weak products suggesting that while they were *Alosa* spp. in origin, the DNA yields and/or the quality of the DNA was poor.

5.3. COI gene amplification
Duplicate DNA samples for all 162 eggs were screened using the COI primer cocktail (C_FishF1t1 / C_FishR1t1). The quality of the amplification was poor relative to the *Alosa* spp.-specific primer set, possibly due to the length of the primers and PCR conditions used but nonetheless, in most cases samples generated a product (Figure 2). Importantly, the samples from Habitat 4 at Penddauwyn (2A/15) and from Llandeilo Bridge (10/3), that failed to yield an amplification product when using the *Alosa* spp.-specific primer set also failed to produce a product when using the COI primer cocktail. Samples 4a/12 to 4a/18 also produced a weak product with the COI primers (Figure 2) but since a strong product was obtained when using the *Alosa* spp.-specific primer set it can be concluded that they are *Alosa* spp. in origin.
Table 1: Summary of the results of DNA analysis of eggs sampled from different locations in the River Tywi. Details of the number of eggs collected at each of the sampling points and the results PCR tests undertaken using the *Alosa*-specific primers set taken from Alexandrino *et al* (2006). The results of COI sequencing are also shown. NS= not sampled

<table>
<thead>
<tr>
<th>Site No.</th>
<th>NGR</th>
<th>Site Name</th>
<th>Total No. Eggs</th>
<th><em>Alosa</em> specific PCR</th>
<th>COI sequence data</th>
<th>Allis shad A. alosa</th>
<th>Twaites shad A. fallax</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>strong positives</td>
<td>negative</td>
<td>weak positives</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>SN 44780 20495</td>
<td>Tidal limit</td>
<td>22</td>
<td>16</td>
<td>0</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>SN 46231 20402</td>
<td>Penddauwyn</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2a</td>
<td>SN 46363 20537</td>
<td>Habitat 4</td>
<td>25</td>
<td>19</td>
<td>1</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>2b</td>
<td>SN 46812 20700</td>
<td>Habitat 5</td>
<td>25</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>SN 469 210</td>
<td>Glantowylan</td>
<td>32</td>
<td>31</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>SN 46740 21494</td>
<td>White Mill</td>
<td>NS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4a</td>
<td>SN 47146 21161</td>
<td>Habitat 8</td>
<td>24</td>
<td>20</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>SN 493 203</td>
<td>Nantgaredig Bridge</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>SN 49902 20108</td>
<td>Cothi confluence</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6a</td>
<td>SN 50504 20250</td>
<td>Cothi Bridge</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>SN 55031 20345</td>
<td>Dryslwyn</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>SN 59168 21463</td>
<td>Cilsan Bridge</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>SN 62653 21991</td>
<td>Llandeilo Bridge</td>
<td>17</td>
<td>12</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>SN 68759 26802</td>
<td>Manordeilo</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>SN 71809 31006</td>
<td>Llanwrda</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>SN75488 33138</td>
<td>Llwynjack</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>162</td>
<td>140</td>
<td>2</td>
<td>20</td>
<td>15</td>
</tr>
</tbody>
</table>
Figure 1: *Alosa*-specific PCR amplicons generated for eggs samples collected from the River Tywi. Individual eggs are numbered based on the site number followed by the egg number. The solid arrow indicates the product expected when using the alocytbf1/ alocybr1 primer set and the open arrow indicates the primer dimers * indicates the eggs where no *Alosa*-specific signal was not obtained.
Figure 2: COI generic PCR amplicons generated for eggs samples collected from the River Tywi. Individual eggs are numbered based on the site number followed by the egg number, and are the same as those given in Figure 1. The solid arrow indicates the product expected when using the C_FishF1t1 / C_FishR1t1 primer cocktail and the open arrow indicates the primer dimers * indicates the eggs where a PCR signal was not obtained.
5.4. Sequence analysis

Sequence analysis was undertaken on 15 of the sequences generated using the COI generic primer cocktail to determine the origin of the PCR signal. The samples chosen were those that produced a weak signal when using the Alosa-specific primer set, but also generated a signal when using the generic primer sets. In addition, sample 5/7, an egg considered too small to be shad in origin based on taxonomic examination by the sampling team, was also sequenced.

The intensity of the products generated by the PCR assays were generally weak which made sequencing difficult, however, it was possible to obtain a consensus sequence for each of the products of between 277 and 652 nucleotides in length (Appendix B).

Alignment of the sequences revealed two distinct sequences that shared 98.56% nucleotide identity (Figure 3), and phylogenetic analysis confirmed the two distinct lineages (Figure 4). BLAST analysis showed that 20% of the COI gene sequences (2a/13 4a/16 and 5/7) were A. fallax in origin and the remaining sequences (2a/10, 2a/11 and 18; 3/31; 4a/9. 4a/14 and 4a/22; 10/15, 10/16 and 10/17; 16/7 and 16/8) were A. alosa in origin (Appendix C).
<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>#KJ204646_Alosa_fallax</td>
<td>CTTCCTGGAG ATGATCAGAT CTATAACGTC ATCGTTACGG CGCACGCCTT CGTAATAATC TTCTTCATAG</td>
</tr>
<tr>
<td>#KJ128407_Alosa_alosa</td>
<td></td>
</tr>
<tr>
<td>#2a/10</td>
<td></td>
</tr>
<tr>
<td>#2a/11</td>
<td></td>
</tr>
<tr>
<td>#2a/13</td>
<td></td>
</tr>
<tr>
<td>#2a/18</td>
<td></td>
</tr>
<tr>
<td>#3/31</td>
<td></td>
</tr>
<tr>
<td>#4/9</td>
<td></td>
</tr>
<tr>
<td>#4/14</td>
<td></td>
</tr>
<tr>
<td>#4/16</td>
<td></td>
</tr>
<tr>
<td>#4/22</td>
<td></td>
</tr>
<tr>
<td>#5/7</td>
<td></td>
</tr>
<tr>
<td>#10/15</td>
<td></td>
</tr>
<tr>
<td>#10/16</td>
<td></td>
</tr>
<tr>
<td>#10/17</td>
<td></td>
</tr>
<tr>
<td>#16/7</td>
<td></td>
</tr>
<tr>
<td>#16/8</td>
<td></td>
</tr>
<tr>
<td>#KC015147_Alosa_sapidissima</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>#KJ204646_Alosa_fallax</td>
<td>TAATGCCAAT TCTAATTGGC GGCTTTGGGA ATTGACTAGT CCCCCTTATG ATCGGGGCAC CAGACATGGC</td>
</tr>
<tr>
<td>#KJ128407_Alosa_alosa</td>
<td></td>
</tr>
<tr>
<td>#2a/10</td>
<td></td>
</tr>
<tr>
<td>#2a/11</td>
<td></td>
</tr>
<tr>
<td>#2a/13</td>
<td></td>
</tr>
<tr>
<td>#2a/18</td>
<td></td>
</tr>
<tr>
<td>#3/31</td>
<td></td>
</tr>
<tr>
<td>#4/9</td>
<td></td>
</tr>
<tr>
<td>#4/14</td>
<td></td>
</tr>
<tr>
<td>#4/16</td>
<td></td>
</tr>
<tr>
<td>#4/22</td>
<td></td>
</tr>
<tr>
<td>#5/7</td>
<td></td>
</tr>
<tr>
<td>#10/15</td>
<td></td>
</tr>
<tr>
<td>#10/16</td>
<td></td>
</tr>
<tr>
<td>#10/17</td>
<td></td>
</tr>
<tr>
<td>#16/7</td>
<td></td>
</tr>
<tr>
<td>#16/8</td>
<td></td>
</tr>
<tr>
<td>#KC015147_Alosa_sapidissima</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>#KJ204646_Alosa_fallax</td>
<td>ATTCCCACGA ATGAACAAACA TGAGCTTCTG GCTACTTCCA CCCTCATTCC TCCTGCTCCT TGCCCTCCTCC</td>
</tr>
<tr>
<td>#KJ128407_Alosa_alosa</td>
<td></td>
</tr>
<tr>
<td>#2a/10</td>
<td></td>
</tr>
<tr>
<td>#2a/11</td>
<td></td>
</tr>
<tr>
<td>#2a/13</td>
<td></td>
</tr>
<tr>
<td>#2a/18</td>
<td></td>
</tr>
<tr>
<td>#3/31</td>
<td></td>
</tr>
<tr>
<td>#4/9</td>
<td></td>
</tr>
<tr>
<td>#4/14</td>
<td></td>
</tr>
<tr>
<td>#4/16</td>
<td></td>
</tr>
<tr>
<td>#4/22</td>
<td></td>
</tr>
<tr>
<td>#5/7</td>
<td></td>
</tr>
<tr>
<td>#10/15</td>
<td></td>
</tr>
<tr>
<td>#10/16</td>
<td></td>
</tr>
<tr>
<td>#10/17</td>
<td></td>
</tr>
<tr>
<td>#16/7</td>
<td></td>
</tr>
<tr>
<td>#16/8</td>
<td></td>
</tr>
<tr>
<td>#KC015147_Alosa_sapidissima</td>
<td></td>
</tr>
<tr>
<td>Sequence</td>
<td>Alignment</td>
</tr>
<tr>
<td>-------------------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>#2a/11</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#2a/13</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#2a/18</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#3/31</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#4/9</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#4/14</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#4/16</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#4/22</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#5/7</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#10/15</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#10/16</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#10/17</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#16/7</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#16/8</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#KC015147_Alosa_sapidissima</td>
<td>.............................................. A........... .................. ..........</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>#KJ204464_Alosa_fallax</td>
<td>GGGGTTGAAG CCGGGCGAGG GACCGGTGA AC ..........</td>
</tr>
<tr>
<td>#KJ128407_Alosa_alosa</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#2a/10</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#2a/11</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#2a/13</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#3/31</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#4/9</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#4/14</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#4/16</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#4/22</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#5/7</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#10/15</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#10/16</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#10/17</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#16/7</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#16/8</td>
<td>..................................................................</td>
</tr>
<tr>
<td>#KC015147_Alosa_sapidissima</td>
<td>.............................................. A........... .................. ..........</td>
</tr>
</tbody>
</table>

Figure 3: Alignment of the partial COI gene sequence (242 nucleotides) generated for putative shad eggs from the River Tywi with sequences from *Alosa alosa*, *Alosa fallax* and the American shad, *Alosa sapidissima*. Eggs chosen for COI sequencing had produced weak products when screening using the *Alosa*-specific PCR assay. Positions of nucleotide variation are indicated using the appropriate IUB code.
Figure 4: Phylogenetic relationships between the partial COI gene sequences obtained for the putative *Alosa* spp. eggs collected on the River Tywi. The tree was generated using neighbour-joining distance methods and bootstrap values >70% are shown at the branch points. Partial COI sequence from the American shad *Alosa sapidissima* was used as an out-group. The scale is the number of nucleotide substitutions per nucleotide.
6. Discussion

This study was undertaken as part of a regular quality assurance of the shad egg sampling programme in Welsh rivers (Wye, Usk and Tywi), using a genetic approach to determine if a representative sample of eggs were correctly identified as shad. Similar studies suggest that although a high proportion of eggs are correctly identified, but there is nevertheless a risk that some material is misidentified (Hardouin et al. 2013).

In this current study, none of the eggs collected from the river Tywi were identified as a non-target species. A high proportion (98.8%) eggs were verified as being shad, and the remainder did not yield DNA of sufficient quality to identify using molecular methods. This result is substantially better than the 85% accuracy achieved in 2013, and shows a significant increase in the reliability of the taxonomic identification in the field. Based on this level of accuracy for the taxonomic identification, none of the 9 sites surveyed on the R.Tywi would have been erroneously recorded as supporting the spawning of shad, if using field identification alone. Shad spawning was confirmed at Nantgaredig Bridge, where all samples collected in 2013 had proved to be minnow (Hardouin et al. 2013). The confirmation of successful spawning at Llandeilo Bridge sample is potentially significant as although there are numerous records of shad spawning upstream of this location (Garrett et al. 2013), this represents the upstream record of shad spawning that has been confirmed using molecular methods.

Some of the eggs identified as shad based on the use of Alosa-spp.-specific Cyt b primers set (Alexandrino et al. 2006) produced relative low amplification signals. To eliminate the possibility of poor primer specificity and potential cross reactivity of the primers with the mitochondrial (Cyt b) gene target sequence from other unrelated species, these eggs were also confirmed as shad by the amplification and sequence analysis of a second mitochondrial gene (COI) sequence generated using a universal primer set (Ivanova et al. 2007). Two eggs that could not be identified genetically using the Alosa spp.-specific PCR assay also failed to generate an amplification product with the COI primer set, and therefore, failure to generate products was most likely due to a poor DNA yield or DNA quality rather than the egg originating from a different species.

Of the 15 eggs selected for COI sequencing all were shown to be shad in origin, and 20% of the eggs were assigned to A. fallax and 80% assigned to A. alosa based on the sequence alignments. These results are consistent with previous studies which estimated the percentage of A. alosa mtDNA in the Tywi shad populations as 71% and 72% respectively (Alexandrino & Faria 2004; Hardouin et al. 2013). Analysis of nuclear DNA sequences would be required to determine if these eggs represent the progeny of a hybridisation between Alosa fallax and Alosa alosa to resolve this issue.

Overall, identification of the species origin of the eggs sampled from the R. Tywi using a combination of Cyt b and COI primers (Alexandrino et al. 2006, Ivanova et al. 2007) was successful. However, the amplification signals generated when using the M13 tagged primers taken from Ivanova et al. (2007) were generally poor, with a very obvious primer dimer formation (Figure 2). If this type of quality assurance is done.
regularly it would benefit significantly from a re-evaluation of the suitability of this primer set. The poor quality of the amplification data was most likely due to a combination of the high melting temperatures (Tm) for these primers (Tm = 74.0°C - 85.9°C) and the low annealing temperature (48°C) used in the PCR assay. It the future, it would be advised that the Tm of the primers is reduced by removing the M13 tags, and/or the annealing temperature in the assay is increased. Since the 3’ termini of the primers start at the same position on the COI gene the same cocktail of truncated primers could also be used to sequence the products without adversely affecting the quality of the sequence data.

7. Conclusions

- 162 eggs were collected from 19 known and potential shad spawning sites on the Tywi. Of these, 160 were confirmed as shad eggs and two yielded poor quality DNA that could not be identified. These results indicate that samplers were correctly identifying shad eggs.
- Shad spawning was confirmed from Nantgaredig.

8. Acknowledgements

I am extremely grateful to Dr. Tristan Hatton-Ellis for helping to improve the report. Heather Garrett, Leila Thornton and Alex Harding co-ordinated and led fieldwork. Other samplers were Mark Bishop, Jill Howells, Hilary Foster, Meryl Tandy, Iestyn Evans, Julie Gething, Chloe Jennings, James Moon, Richard West, Emma Keenan, Ali Baird, Nicola Broadbrige, Paul Hyatt and Kerry Rogers.

9. References


10. Appendices

10.1. Appendix A: Sampling procedures for Shad genetic work

10.1.1 Egg collection and storage for genetic work analysis

Cefas will provide clean, 15 prelabelled sample pots, parafilm and marker pens for each river and instructions for the dispatch of the egg samples to the laboratory. Sample pots will contain 95% ethanol to enable samples to be fixed and stored. In addition to normal kick sampling equipment, sampling teams will need to bring the following:
- Forceps
- Disposable PVC gloves
- Coolbag with icepacks or other means of keeping samples cool.
- PPE required for handling 95% ethanol as determined by your risk assessment.

All suspected shad eggs collected from each site should be placed carefully in the pots using a clean pair of forceps, taking care not to burst the egg. Transfer of excessive quantities of water with the eggs should also be avoided as this will dilute the fixative. Eggs that appear to be close to hatching should not be collected as these are likely to hatch in the alcohol and thereby pose an increased risk of contamination; a new pot should be used for each sampling site.

Each pot should contain eggs from one site only. It is recommended to collect no more than 30 eggs in a single pot, and use multiple pots for the same sample site if required. Cefas will separate individual eggs in the laboratory.

When sampling is completed, the lids of the individual sample tubes should be sealed using the parafilm provided. This will reduce the risk of ethanol leakage and/or evaporation.

Gloves should be worn when handling eggs in order to avoid cross-contamination. The usual health and safety measures associated with handling ethanol in the field should be observed. There is no need to clean forceps between eggs unless they are obviously dirty, in which case they may be rinsed with alcohol.

Once placed in the sample tubes the eggs will need to be placed in a cool bag and transferred to a refrigerator at approximately 5°C as soon as possible after collection in the field. The sample tubes will then need to be shipped to the Cefas Weymouth laboratory by a reliable courier with the temperature maintained at approximately 5°C. NRW will pay for the costs of the courier from point of collection to point of delivery.

Sample containers should be sent to:-

Dr David Stone
Cefas Weymouth Laboratory,
Barrack Road, The Nothe, Weymouth, Dorset. DT4 8UB
10.1.2 CCW Shad Egg Monitoring Protocol

The shad monitoring protocol is slightly different to that in the LIFE+ monitoring methodology publication. Only shad eggs should be recorded. Any ‘suspect’ eggs can be noted but recorded as unknown eggs.

The protocol is as follows;

1. A field sheet should be filled in for each site with NGR and data collected as cited below.

2. 10 kick samples should be undertaken and the number of shad eggs in each kick counted and recorded on the field sheet. Each kick sample lasts for 15 seconds.

3. If more than 10 shad eggs have been recorded at the end of 10 kicks then the survey is complete.

4. However, if 10 or less shad eggs are found then kick surveys should continue (with number of shad eggs recorded for each kick on the survey form) until over 10 shad eggs are recorded up to a maximum of 25 kick samples.

5. If 10 or less shad eggs have been recorded the survey should be terminated after 25 kick samples.

6. Presence of other SAC species (e.g. bullheads or lampreys) and invasive species (e.g. knotweed) should be recorded on the survey form.

7. Photos should be taken of up and downstream at each site

8. All survey forms (electronic or paper copy) should be sent to Alex Harding. If a paper copy is sent in the post then a copy should also be retained with the surveyor in case it gets lost in the post.

In addition;
- Photos should be taken of any eggs for which there is uncertainty over whether they are shad eggs. This uncertainty should be recorded on the survey form.
- Biosecurity – equipment and wellington boots/waders should be completely air dried or dipped in Virkon S Aquatic when moving between catchments.

The protocol is shown in flow chart below:
Figure 5: shad egg survey protocol

Biosecurity – ensure equipment and footwear disinfected before going out on site.

10 kick surveys undertaken (15 seconds each). Number of shad eggs in each kick counted and recorded on field sheet.

More than 10 eggs found

Continue kick sampling until more than 10 eggs found. Record number of eggs in each kick sample on survey form

More than 10 eggs found

10 or less eggs found

10 or less eggs found

Survey terminated. Field sheets sent to Rhian.

Biosecurity – ensure equipment and footwear disinfected before going to different catchment.

Fill in spreadsheet on ffynnon to record which sites been to and whether shad eggs were found OR e-mail Alex Harding with info.

Field sheets & photos sent to Alex Harding

Terminate survey after a maximum of 25 kicks
10.2. Appendix B: Cyt Bb gene sequences

10.2.1 Input file for the partial Cytb sequences obtained for the shad eggs collected on the River Tywi in 2014

>KKC015147_Alosa sapidissima
CTTTTACCTAGTTATTTTTGCTGGCTGAGCAGGAACAGTAGTAGGCACTGCTTATATCCGGAGCCCG
AACTGAGCACAACCGGGGCGCTTCTCGGAGATGATCAGATCTACAACGTCATCGTTACGGCGCACGC
CTTCGTAATAATCTTCTTTATATGGAAGACTATGACACTACAAGTGCTTATCTCTACGCTTCCTCCTC
TATGATCGGGGCACCAGACATGGCATCTCCCCAGAATGAACACATGAGCTTCTTTCCTCTCCTACTCCT
CATGGCTAGCTGCTGGGATTACAATGCTCCTAACAGACCGAAATCTAAATACAACCTTCTTTGACCC
GGCAGGGGGGAGGGACCCCAATTTTATATCAACACCTA

>KJK204646_Alosa fallax
CTTTTACCTAGTTATTTTTGCTGGCTGAGCAGGAACAGTAGTAGGCACTGCTTATATCCGGAGCCCG
AACTGAGCACAACCGGGGCGCTTCTCGGAGATGATCAGATCTATAACGTCATCGTTACGGCGCACGC
CTTCGTAATAATCTTCTTTATATGGAAGACTATGACACTACAAGTGCTTATCTCTACGCTTCCTCCTC
TATGATCGGGGCACCAGACATGGCATCTCCCCAGAATGAACACATGAGCTTCTTTCCTCTCCTACTCCT
CATGGCTAGCTGCTGGGATTACAATGCTCCTAACAGACCGAAATCTAAATACAACCTTCTTTGACCC
GGCAGGGGGGAGGGACCCCAATTTTATATCAACACCTA

>KJ128407_Alosa alosa
CTAGTATTTTTGCTGGCTGAGCAGGAACAGTAGTAGGCACTGCTTATATCCGGAGCCCG
AACTGAGCACAACCGGGGCGCTTCTCGGAGATGATCAGATCTATAACGTCATCGTTACGGCGCACGC
CTTCGTAATAATCTTCTTTATATGGAAGACTATGACACTACAAGTGCTTATCTCTACGCTTCCTCCTC
TATGATCGGGGCACCAGACATGGCATCTCCCCAGAATGAACACATGAGCTTCTTTCCTCTCCTACTCCT
CATGGCTAGCTGCTGGGATTACAATGCTCCTAACAGACCGAAATCTAAATACAACCTTCTTTGACCC
GGCAGGGGGGAGGGACCCCAATTTTATATCAACACCTA

>2a/10
10.3. Appendix C: BLAST results

10.3.1 Blast results for partial COI gene sequence for the egg sample 2a/13

BLASTN 2.2.30+


RID: 6IMXSX5U013

Database: Nucleotide collection (nt)
29,089,105 sequences; 82,820,087,835 total letters
Query=2a/13
Length=625

Sequences producing significant alignments:

<table>
<thead>
<tr>
<th>Reference</th>
<th>Description</th>
<th>Identities</th>
<th>Gaps</th>
<th>Score</th>
<th>E Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>gb</td>
<td>KJ204646.1</td>
<td>Alosa fallax voucher MT01883 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial</td>
<td>625/625 (100%)</td>
<td>0/625 (0%)</td>
<td>1155</td>
</tr>
<tr>
<td>gb</td>
<td>KJ128409.1</td>
<td>Alosa fallax voucher NR:57055 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial</td>
<td>625/625 (100%)</td>
<td>0/625 (0%)</td>
<td>1155</td>
</tr>
<tr>
<td>gb</td>
<td>KJ128408.1</td>
<td>Alosa fallax voucher NR:52513 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial</td>
<td>625/625 (100%)</td>
<td>0/625 (0%)</td>
<td>1155</td>
</tr>
<tr>
<td>gb</td>
<td>KC500181.1</td>
<td>Alosa alosa voucher TR212EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial</td>
<td>625/625 (100%)</td>
<td>0/625 (0%)</td>
<td>1149</td>
</tr>
<tr>
<td>gb</td>
<td>KC500180.1</td>
<td>Alosa alosa voucher TR211EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial</td>
<td>625/625 (100%)</td>
<td>0/625 (0%)</td>
<td>1149</td>
</tr>
<tr>
<td>gb</td>
<td>KJ552649.1</td>
<td>Alosa agone isolate Ex04E10 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial</td>
<td>625/625 (100%)</td>
<td>0/625 (0%)</td>
<td>1144</td>
</tr>
<tr>
<td>gb</td>
<td>KJ552379.1</td>
<td>Alosa agone isolate Ex04F2 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial</td>
<td>625/625 (100%)</td>
<td>0/625 (0%)</td>
<td>1144</td>
</tr>
<tr>
<td>gb</td>
<td>KC500174.1</td>
<td>Alosa alosa voucher TR202EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial</td>
<td>625/625 (100%)</td>
<td>0/625 (0%)</td>
<td>1144</td>
</tr>
<tr>
<td>gb</td>
<td>KC500173.1</td>
<td>Alosa alosa voucher TR201EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial</td>
<td>625/625 (100%)</td>
<td>0/625 (0%)</td>
<td>1144</td>
</tr>
</tbody>
</table>

Alignments:

```
Query  1    AGGGATAGTAGGCACTGCCTTAAGTCTCTTAATCCGAGCCGAACTGAGCCAACCGGGGC  60
Sbjct  28   AGGGATAGTAGGCACTGCCTTAAGTCTCTTAATCCGAGCCGAACTGAGCCAACCGGGGC  87
```

```
Query  61   ACTTCTCGGAGATGATCAGATCTATAACGTCATCGTTACGGCGCACGCCTTCGTAATAAT  120
Sbjct  88   ACTTCTCGGAGATGATCAGATCTATAACGTCATCGTTACGGCGCACGCCTTCGTAATAAT  147
```

```
Query  121  CTTCTTCAGATAGAATGCCAATTCTAATTGGCGGCTTTGGGAATTGACTAGTCCCCCTAT  180
Sbjct  148  CTTCTTCAGATAGAATGCCAATTCTAATTGGCGGCTTTGGGAATTGACTAGTCCCCCTAT  207
```

```
Query  181  GATCGGGGCACCAGACATGGCATTCCCACGAATGAACAACATGAGCTTCTGGCTACTTCC  240
Sbjct  285  GATCGGGGCACCAGACATGGCATTCCCACGAATGAACAACATGAGCTTCTGGCTACTTCC  317
```

Score = 1155 bits (625), Expect = 0.0
Identities = 625/625 (100%), Gaps = 0/625 (0%)
Strand=Plus/Plus
>gb|KC500180.1| Alosa alosa voucher TR211EK cytchrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
Length=654

Score =  1149 bits (622),  Expect = 0.0
Identities = 624/625 (99%), Gaps = 0/625 (0%)
Strand=Plus/Plus

Query  1    AGGGATAGTAGGCACTGCCTTAAGTCTCTTAATCCGAGCCGAACTGAGCCAACCCGGGGC
Sbjct  27   AGGGATAGTAGGCACTGCCTTAAGTCTCTTAATCCGAGCCGAACTGAGCCAACCCGGGGC

Query  61   ACTTCTCGGAGATGATCAGATCTATAACGTCATCGTTACGGCGCACGCCTTCGTAATAAT
Sbjct  87   ACTTCTCGGAGATGATCAGATCTATAACGTCATCGTTACGGCGCACGCCTTCGTAATAAT

Query  121  CTTCTTCTATAGTAATGCCAATTCTAATTGGCGGCTTTGGGAATTGACTAGTCCCCCTTAT
Sbjct  147  CTTCTTCTATAGTAATGCCAATTCTAATTGGCGGCTTTGGGAATTGACTAGTCCCCCTTAT

Query  181  GATCGGGGCACCAGACATGGCATTCCCACGAATGAACAACATGAGCTTCTGGCTACTTCC
Sbjct  207  GATCGGGGCACCAGACATGGCATTCCCACGAATGAACAACATGAGCTTCTGGCTACTTCC

Query  241  ACCCTCATTCCTCCTCCTCCTTGCCTCCTCCGGGGTTGAAGCCGGGGCAGGGACCGGGTG
Sbjct  267  ACCCTCATTCCTCCTCCTCCTTGCCTCCTCCGGGGTTGAAGCCGGGGCAGGGACCGGGTG

Query  301  AACAGTCTACCCGCCTTGGCAGGCAATCTTGCCCACGCCGGGGCATCCGTCGACCTGAC
Sbjct  327  AACAGTCTACCCGCCTTGGCAGGCAATCTTGCCCACGCCGGGGCATCCGTCGACCTGAC

Query  361  TATCTTCTCTCTTCTCTCACTTAGGATTATCTCAATCGGATTCTCGGGCCATTATTTTATAC
Sbjct  387  TATCTTCTCTCTTCTCTCACTTAGGATTATCTCAATCGGATTCTCGGGCCATTATTTTATAC

Query  421  CACAAATCAATTAATGAAACCCCCCTGCAATCTCACAATATCAAAACCCCTATTATTATG
Sbjct  447  CACAAATCAATTAATGAAACCCCCCTGCAATCTCACAATATCAAAACCCCTATTATTATG

Query  481  ATCCGGTCCTGTAAACCCGCTCTCTCTCTCTCCTCTCCTCCTCTCCCTGCTAGCTGCTGGGAT
Sbjct  507  ATCCGGTCCTGTAAACCCGCTCTCTCTCTCTCCTCTCCTCCTCTCCCTGCTAGCTGCTGGGAT

Query  541  TACAAATGCTCTAACCAGCGGGAAAACTACAAATACAACATCAATCGGATTCTCGGGCCATTATTTTATAC
Sbjct  567  TACAAATGCTCTAACCAGCGGGAAAACTACAAATACAACATCAATCGGATTCTCGGGCCATTATTTTATAC

Query  601  GGACCCAATTTTATACCAACACCTA  625
Sbjct  627  GGACCCAATTTTATACCAACACCTA  651
>gb|KJ552649.1| Alosa agone isolate Ex04E10 cytochrome oxidase subunit I gene, partial cds; mitochondrial
Length=652

Score = 1144 bits (619), Expect = 0.0
Identities = 623/625 (99%), Gaps = 0/625 (0%)
Strand=Plus/Plus

Query 1
AGGGATAGTAGCAGCCTTAACTCCTCTATATAATGCGGGAACCTAGGCCAACCCCGGGGC 60

Sbjct 28
AGGGATAGTAGCAGCCTTAACTCCTCTATATAATGCGGGAACCTAGGCCAACCCCGGGGC 87

Query 61
ACTTCTCGAGATGATACGATCTATACGTACGTCTTTACGGGACGCCCTTGTAATAAT 120

Sbjct 88
ACTTCTCGAGATGATACGATCTATACGTACGTCTTTACGGGACGCCCTTGTAATAAT 147

Query 121
CTTCTTCATAGTAATGCCAATTCTAATTGGCGGCTTTGGGAATTGACTAGTCCCCCTTAT 180

Sbjct 148
CTTCTTCATAGTAATGCCAATTCTAATTGGCGGCTTTGGGAATTGACTAGTCCCCCTTAT 207

Query 181
GATCGGGGCACCAGACATGGCATTCCCACGAATGAACAACATGAGCTTCTGGCTACTTCC 240

Sbjct 208
GATCGGGGCACCAGACATGGCATTCCCACGAATGAACAACATGAGCTTCTGGCTACTTCC 267

Query 241
AACAGTCTACCCGCCTTTGGCAGGCAATCTTGCCCACGCCGGGGCATCCGTCGACCTGAC 300

Sbjct 328
AACAGTCTACCCGCCTTTGGCAGGCAATCTTGCCCACGCCGGGGCATCCGTCGACCTGAC 387

Query 301
TATCTTCTCTCTCTCTCTAATCTAGCAGGTATTTCATCGATTCTTGGGGCCATTAATTTTATTAC 360

Sbjct 388
TATCTTCTCTCTCTCTCTAATCTAGCAGGTATTTCATCGATTCTTGGGGCCATTAATTTTATTAC 447

Query 361
CACAAATCATTAATATGAAACCCCCTGCAATCTCACAATATCAAACACCCCTATTTGTGTG 420

Sbjct 448
CACAAATCATTAATATGAAACCCCCTGCAATCTCACAATATCAAACACCCCTATTTGTGTG 507

Query 421
ATCCGTGCTTGTAACGGCCGTTCTCCTTCTCTCACTCCCTGTGCTAGCTGCTGGGAT 480

Sbjct 508
ATCCGTGCTTGTAACGGCCGTTCTCCTTCTCTCACTCCCTGTGCTAGCTGCTGGGAT 567

Query 481
TACAATGCTCCTAACAGACCGAAATCTAAATACAACCTTCTTTGACCCGGCAGGGGGAGG 540

Sbjct 568
TACAATGCTCCTAACAGACCGAAATCTAAATACAACCTTCTTTGACCCGGCAGGGGGAGG 627

Query 601
GGACCCAATTTTATACCAACACCTA 625

Sbjct 628
GGACCCAATTTTATACCAACACCTA 652

>gb|KJ552379.1| Alosa agone isolate Ex04F2 cytochrome oxidase subunit I gene, partial cds; mitochondrial
gb|KJ552455.1| Alosa agone isolate Ex04E11 cytochrome oxidase subunit I gene, partial cds; mitochondrial
gb|KJ552682.1| Alosa agone isolate Ex04E8 cytochrome oxidase subunit I gene, partial cds; mitochondrial
gb|KJ552733.1| Alosa agone isolate Ex04E12 cytochrome oxidase subunit I gene, partial cds; mitochondrial
Length=652

Score = 1144 bits (619), Expect = 0.0
Identities = 623/625 (99%), Gaps = 0/625 (0%)
Strand=Plus/Plus

Query 1
AGGGATAGTAGCAGCCTTAACTCCTCTATATAATGCGGGAACCTAGGCCAACCCCGGGGC 60

Sbjct 28
AGGGATAGTAGCAGCCTTAACTCCTCTATATAATGCGGGAACCTAGGCCAACCCCGGGGC 87

Query 61
ACTTCTCGAGATGATACGATCTATACGTACGTCTTTACGGGACGCCCTTGTAATAAT 120

Sbjct 88
ACTTCTCGAGATGATACGATCTATACGTACGTCTTTACGGGACGCCCTTGTAATAAT 147

Query 121
CTTCTTCATAGTAATGCCAATTCTAATTGGCGGCTTTGGGAATTGACTAGTCCCCCTTAT 180

Sbjct 148
CTTCTTCATAGTAATGCCAATTCTAATTGGCGGCTTTGGGAATTGACTAGTCCCCCTTAT 207

Query 181
GATCGGGGCACCAGACATGGCATTCCCACGAATGAACAACATGAGCTTCTGGCTACTTCC 240

Sbjct 208
GATCGGGGCACCAGACATGGCATTCCCACGAATGAACAACATGAGCTTCTGGCTACTTCC 267

Query 241
AACAGTCTACCCGCCTTTGGCAGGCAATCTTGCCCACGCCGGGGCATCCGTCGACCTGAC 300

Sbjct 328
AACAGTCTACCCGCCTTTGGCAGGCAATCTTGCCCACGCCGGGGCATCCGTCGACCTGAC 387

Query 301
TATCTTCTCTCTCTCTCTAATCTAGCAGGTATTTCATCGATTCTTGGGGCCATTAATTTTATTAC 360

Sbjct 388
TATCTTCTCTCTCTCTCTAATCTAGCAGGTATTTCATCGATTCTTGGGGCCATTAATTTTATTAC 447

Query 421
CACAAATCATTAATATGAAACCCCCTGCAATCTCACAATATCAAACACCCCTATTTGTGTG 420

Sbjct 448
CACAAATCATTAATATGAAACCCCCTGCAATCTCACAATATCAAACACCCCTATTTGTGTG 507

Query 481
ATCCGTGCTTGTAACGGCCGTTCTCCTTCTCTCACTCCCTGTGCTAGCTGCTGGGAT 480

Sbjct 508
ATCCGTGCTTGTAACGGCCGTTCTCCTTCTCTCACTCCCTGTGCTAGCTGCTGGGAT 567

Query 601
GGACCCAATTTTATACCAACACCTA 625

Sbjct 628
GGACCCAATTTTATACCAACACCTA 652
>gb|KC500174.1| Alosa alosa voucher TR202EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500175.1| Alosa alosa voucher TR206EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500176.1| Alosa alosa voucher TR207EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500177.1| Alosa alosa voucher TR208EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500178.1| Alosa alosa voucher TR209EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500179.1| Alosa alosa voucher TR210EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500187.1| Alosa alosa voucher TR203EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500190.1| Alosa alosa voucher TR204EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500191.1| Alosa alosa voucher TR205EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500192.1| Alosa alosa voucher TR206EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial

Length=654

Score =  1144 bits (619),  Expect = 0.0
Identities = 623/625 (99%),  Gaps = 0/625 (0%)
Strand=Plus/Plus

Query 1  AGGGATAGTAGGCGACTGCTTTAATATCCGAGGCCGAGGAGCACTGAGCAGCCCAAACCGGGCC  60

Sbjct 27  AGGGATAGTAGGCGACTGCTTTAATATCCGAGGCCGAGGAGCACTGAGCAGCCCAAACCGGGCC  86

Query 61  ACTTCTCGGAGATGATCAGATCTATAACGTCATCGTTACGGCGCACGCCTTCGTAATAAT  120

Sbjct 87  ACTTCTCGGAGATGATCAGATCTATAACGTCATCGTTACGGCGCACGCCTTCGTAATAAT  146

Query 121  CTTCTTCATAGTAATGCCAATTCTAATTGGCGGCTTTGGGAATTGACTAGTCCCCCTTAT  180

Sbjct 147  CTTCTTCATAGTAATGCCAATTCTAATTGGCGGCTTTGGGAATTGACTAGTCCCCCTTAT  207

Query 181  GATCGGGGACCAGACATGGCATTCCCACGAATGAACAACATGAGCTTCTGGCTACTTCC  240

Sbjct 207  GATCGGGGACCAGACATGGCATTCCCACGAATGAACAACATGAGCTTCTGGCTACTTCC  266
<table>
<thead>
<tr>
<th>Query</th>
<th>Sbjct</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>241</td>
<td>267</td>
<td>ACCCTCTCTCTCTCTGCTCTCTCCGGGTGTTGAGCCGGCGACGGACGGTG</td>
</tr>
<tr>
<td>301</td>
<td>327</td>
<td>AACAGTCTACGCCGACCGTCTTTGCTCTCTCTGCTCTCCGGGGCATCCGTCGACCTGAC</td>
</tr>
<tr>
<td>361</td>
<td>387</td>
<td>TATCTTTCTCTCTCCATCTACAGGTTATTTCACTAGATTCTTGGGGCGATTAAATTATTAC</td>
</tr>
<tr>
<td>421</td>
<td>447</td>
<td>CACAAATCTTATATAATGAAACCCCCTGCAATCTCACAATATCAAACACCCCTATTTGTGTG</td>
</tr>
<tr>
<td>481</td>
<td>507</td>
<td>ATCCGTGCTTGTAACGGCCGTTCTCCTTCTCTCTCTTCTCTGCTCTAGCTGCTGGGAT</td>
</tr>
<tr>
<td>541</td>
<td>567</td>
<td>TACAATGCTCCTAACAGACCGAAATCTAAATACAACCTTCTTTGACCCGGCAGGGGGAGG</td>
</tr>
<tr>
<td>601</td>
<td>627</td>
<td>GGACCCAATTTTATAAACACACTTA</td>
</tr>
<tr>
<td>622</td>
<td>651</td>
<td>GGACCCAATTTTATACCAACACTTA</td>
</tr>
</tbody>
</table>

>gb|KC500173.1| Alosa alosa voucher TR201EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
Length=654

Score = 1144 bits (619), Expect = 0.0
Identities = 623/625 (99%), Gaps = 0/625 (0%)
Strand=Plus/Plus
Query 601 GGACCAATTTTATACCAACACCTA 625

Sbjct 627 GGACCAATTTTATACCAACACCTA 651

Database: Nucleotide collection (nt)
Posted date: Nov 8, 2014 10:16 PM
Number of letters in database: 82,820,087,835
Number of sequences in database: 29,089,105

Lambda      K        H
1.33    0.621     1.12

Gapped Lambda      K        H
1.28    0.460    0.850

Matrix: blastn matrix:1 -2
Gap Penalties: Existence: 0, Extension: 0
Number of Sequences: 29089105
Number of Hits to DB: 0
Number of extensions: 0
Number of successful extensions: 0
Number of sequences better than 10: 40
Number of HSP's better than 10 without gapping: 0
Number of HSP's gapped: 40
Number of HSP's successfully gapped: 40
Length of query: 625
Length of database: 82820087835
Length adjustment: 34
Effective length of query: 591
Effective length of database: 81831058265
Effective search space: 48362155434615
Effective search space used: 48362155434615
A: 0
X1: 13 (25.0 bits)
X2: 32 (59.1 bits)
X3: 54 (99.7 bits)
S1: 13 (25.1 bits)
S2: 23 (43.6 bits)
10.3.2 BLAST results for partial COI gene sequence for egg sample 2a/11

BLASTN 2.2.30+


RID: 61PYBX3A01R

Database: Nucleotide collection (nt)
29,089,105 sequences; 82,820,087,835 total letters
Query=2a/11
Length=652

<table>
<thead>
<tr>
<th>Score</th>
<th>E</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1199 bits</td>
<td>0.0</td>
<td></td>
</tr>
</tbody>
</table>

Sequences producing significant alignments:

<table>
<thead>
<tr>
<th>gb</th>
<th>KJ552458.1</th>
<th>Alosa alosa isolate Ex53B1 cytochrome oxidase ...</th>
<th>1199</th>
<th>0.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>gb</td>
<td>KJ128407.1</td>
<td>Alosa alosa voucher NRM:54753 cytochrome oxidase ...</td>
<td>1181</td>
<td>0.0</td>
</tr>
<tr>
<td>gb</td>
<td>KJ554963.1</td>
<td>Alosa alosa isolate Ex53B3 cytochrome oxidase ...</td>
<td>1177</td>
<td>0.0</td>
</tr>
<tr>
<td>gb</td>
<td>KJ552463.1</td>
<td>Alosa vistonica isolate CIBIOS2 cytochrome ox...</td>
<td>1171</td>
<td>0.0</td>
</tr>
<tr>
<td>gb</td>
<td>KJ552478.1</td>
<td>Alosa alosa isolate Ex53B5 cytochrome oxidase ...</td>
<td>1170</td>
<td>0.0</td>
</tr>
<tr>
<td>gb</td>
<td>KJ768202.1</td>
<td>Alosa fallax voucher KELPI252 cytochrome oxid...</td>
<td>1160</td>
<td>0.0</td>
</tr>
<tr>
<td>gb</td>
<td>KJ204646.1</td>
<td>Alosa alosa voucher MT01883 cytochrome oxid...</td>
<td>1155</td>
<td>0.0</td>
</tr>
<tr>
<td>dbj</td>
<td>AP009131.1</td>
<td>Alosa alosa mitochondrial DNA, complete genome</td>
<td>1155</td>
<td>0.0</td>
</tr>
<tr>
<td>gb</td>
<td>KJ552379.1</td>
<td>Alosa agone isolate Ex04F2 cytochrome oxidase ...</td>
<td>1155</td>
<td>0.0</td>
</tr>
<tr>
<td>gb</td>
<td>KJ552463.1</td>
<td>Alosa vistonica isolate CIBIOS2 cytochrome ox...</td>
<td>1171</td>
<td>0.0</td>
</tr>
<tr>
<td>dbj</td>
<td>AP009131.1</td>
<td>Alosa alosa mitochondrial DNA, complete genome</td>
<td>1155</td>
<td>0.0</td>
</tr>
</tbody>
</table>

ALIGNMENTS

>gb|KJ552458.1| Alosa alosa isolate Ex53B1 cytochrome oxidase subunit I gene, partial cds; mitochondrial
|---|-----------|--------------------------------------------------|------|-----|
| gb|KJ552509.1| Alosa alosa isolate Ex53B6 cytochrome oxidase subunit I gene, partial cds; mitochondrial
| gb|KJ552581.1| Alosa alosa isolate CIBIOS2 cytochrome oxidase subunit I gene, partial cds; mitochondrial
| gb|KJ552691.1| Alosa alosa isolate Ex53B2 cytochrome oxidase subunit I gene, partial cds; mitochondrial
| gb|KJ552695.1| Alosa alosa isolate Ex53B4 cytochrome oxidase subunit I gene, partial cds; mitochondrial

Length=652

Score = 1199 bits (649), Expect = 0.0
Identities = 651/652 (99%), Gaps = 0/652 (0%)
Strand=Plus/Plus

<table>
<thead>
<tr>
<th>Query</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>241</td>
<td>300</td>
</tr>
</tbody>
</table>
>gb|KJ128407.1| Alosa alosa voucher NRM:54753 cytochrome oxidase subunit I (COI) gene, partial cds; mitochondrial

Length=648  
Score = 1181 bits (639), Expect = 0.0  
Identities = 643/645 (99%), Gaps = 0/645 (0%)  
Strand=Plus/Plus  

 Query 8  CTAGTATTTGGTGCCTGAGCAGGGATAGTAGGCACTGCCTTAAGTCTCTTAATCCGAGCC  67  
 Sbjct 241  GAACAACATGACCTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG  300  

 Query 301  GGTTGAAAGCGGCGGCAAGAACCGGTTAGAAGCTATCCCGCTTTGGCGAGCAATCTG  360  
 Sbjct 301  GGTTGAGCAGCGGCGGGAGAACGGTGGAGAACTATCCCGCTTTGGCGAGCAATCTG  360  

 Query 361  CCACGCCGAGACATCCTGGCACTAACTACCTCTCTCTCTCTCTCTCTGATTATTCATC  420  
 Sbjct 361  CCACGCCGAGACATCCTGGCACTAACTACCTCTCTCTCTCTCTCTCTCTGATTATTCATC  420  

 Query 421  AAATCTTGGGCAATTATTTTATTACCAAACTTAAATAGAAACCCCTGCAACTCC  480  
 Sbjct 421  AAATCTTGGGCAATTATTTTATTACCAAACTTAAATAGAAACCCCTGCAACTCC  480  

 Query 481  ACAATACAAACCCCTATTTGATGACGACTTTGACGATCAGACATCTATACCTATGAGAGCC  540  
 Sbjct 481  ACAATACAAACCCCTATTTGATGACGACTTTGACGATCAGACATCTATACCTATGAGAGCC  540  

 Query 541  CTACCTCCTCTCGATGCTGCTCGGTTATTCAAGCTCCTAAACAGGCAAATCTAAATAC  600  
 Sbjct 541  CTACCTCCTCTCGATGCTGCTCGGTTATTCAAGCTCCTAAACAGGCAAATCTAAATAC  600  

 Query 601  AAACCTTCTTTGACCCGGCAGGGGGAGGGGACCCAATTTTATACCAACACCTA  652  
 Sbjct 601  AAACCTTCTTTGACCCGGCAGGGGGAGGGGACCCAATTTTATACCAACACCTA  652  

 Query 878  AATTGACTAGTCCCCCTTATGATCGGGGCACCAGACATGGCATTCCCACGAAATGAACAAC  247  
 Sbjct 241  ATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGGGGTTGAA  300  

 Query 368  GGAGCATCCGTCGACCTAACTATCTTCTCTCTTCTCTATCTAGCAGGTATTTCATCAATTCTT  427  
 Sbjct 361  GGAGCATCCGTCGACCTAACTATCTTCTCTCTTCTCTATCTAGCAGGTATTTCATCAATTCTT  427  

 Query 428  GGGGCCATTAATTTTATTACCAAACTTAAATAGAAACCCCTGCAACTCC  487  
 Sbjct 421  GGGGCCATTAATTTTATTACCAAACTTAAATAGAAACCCCTGCAACTCC  487  

 Query 488  CAAACGCCCTATTTTATTACCAAACTTAAATAGAAACCCCTGCAACTCC  547  
 Sbjct 481  CAAACGCCCTATTTTATTACCAAACTTAAATAGAAACCCCTGCAACTCC  547  

 Query 548  CCTGTGACTGCTGCTGGGATTACAAATGCTCCTAACAGCGAAATCTAAATACACCTT  607  
 Sbjct 541  CCTGTGACTGCTGCTGGGATTACAAATGCTCCTAACAGCGAAATCTAAATACACCTT  607  

 Query 608  TTTGACCCGGCAGGGGGAGGGGACCCAATTTTATACCAACACCTA  652  
 Sbjct 601  TTTGACCCGGCAGGGGGAGGGGACCCAATTTTATACCAACACCTA  652  

www.naturalresourceswales.gov.uk
<table>
<thead>
<tr>
<th>Query</th>
<th>Sbjct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>60</td>
</tr>
<tr>
<td>76</td>
<td>136</td>
</tr>
<tr>
<td>121</td>
<td>196</td>
</tr>
<tr>
<td>256</td>
<td>301</td>
</tr>
<tr>
<td>241</td>
<td>361</td>
</tr>
<tr>
<td>421</td>
<td>481</td>
</tr>
<tr>
<td>496</td>
<td>556</td>
</tr>
<tr>
<td>541</td>
<td>616</td>
</tr>
<tr>
<td>601</td>
<td>601</td>
</tr>
</tbody>
</table>

**gb|KJ554963.1| Alosa alosa isolate Ex53B3 cytochrome oxidase subunit I gene, partial cds; mitochondrial**

Length=637

Score = 1177 bits (637), Expect = 0.0

Identities = 637/637 (100%), Gaps = 0/637 (0%)

Strand=Plus/Plus

---

**gb|KJ552463.1| Alosa vistonica isolate CIBIOS2 cytochrome oxidase subunit I gene, partial cds; mitochondrial**

Length=652

Score = 1171 bits (634), Expect = 0.0

Identities = 646/652 (99%), Gaps = 0/652 (0%)

Strand=Plus/Plus

---

**gb|KJ552592.1| Alosa immaculata isolate Ex51B2 cytochrome oxidase subunit I gene, partial cds; mitochondrial**

Length=652

Score = 1171 bits (634), Expect = 0.0

Identities = 646/652 (99%), Gaps = 0/652 (0%)

Strand=Plus/Plus

---

**gb|KJ552592.1| Alosa alosa isolate Ex53B3 cytochrome oxidase subunit I gene, partial cds; mitochondrial**

Length=637

Score = 1177 bits (637), Expect = 0.0

Identities = 637/637 (100%), Gaps = 0/637 (0%)

Strand=Plus/Plus
gb|KJ552478.1| Alosa alosa isolate Ex53B5 cytochrome oxidase subunit I gene, partial cds; mitochondrial

Length=635

Score = 1170 bits (633), Expect = 0.0
Identities = 634/635 (99%), Gaps = 0/635 (0%)
Strand=Plus/Plus

Query 17  GGTGCCTGAGCAGGGATAGTAGGCACTGCCTTAAGTCTCTTAATCCGAGCCGAACTGAGC
76

Sbjct 121 CGTTACGCCAGCAGCCTTGTAATAATCTTCTTCTCA TATGCAAATTGCAATTGCGG
180

Query 181  CTTTGGGAATTGACTAGTCCCCCTTATGATCGGGGCACCAGACATGGCATTCCCACGAAT
240

Sbjct 181 CTTTGGGAATTGACTAGTCCCCCTTATGATCGGGGCACCAGACATGGCATTCCCACGAAT
240

Query 241  GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
300

Sbjct 241 GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
300

Query 301  GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
360

Sbjct 301 GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
360

Query 361  GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
420

Sbjct 361 GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
420

Query 421  GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
480

Sbjct 421 GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
480

Query 481  GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
540

Sbjct 481 GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
540

Query 541  GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
600

Sbjct 541 GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
600

Query 601  GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
652

Sbjct 601 GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG
652
>gb|KJ768202.1| Alosa fallax voucher MLFP1252 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
Length=652

Score = 1160 bits (628),  Expect = 0.0
Identities = 644/652 (99%), Gaps = 0/652 (0%)
Strand=Plus/Plus

Query 1  CTTTTACCTAGTATTTGGTGCCTGAGCAGGGATAGTAGGCACTGCCTTAAGTCTCTTAAT  60
Sbjct  1  CTTTTACCTAGTATTTGGTGCCTGAGCAGGGATAGTAGGCACTGCCTTAAGTCTCTTAAT  60

Query 61  CCGAGCCGAACTGAGCCAACCCGGGGCACTTCTCGGAGATGATCAGATCTATAACGTCAT  120
Sbjct  61  CCGAGCCGAACTGAGCCAACCCGGGGCACTTCTCGGAGATGATCAGATCTATAACGTCCT  120

Query 121  CGTTACCGCGCACGCCTTCGTAATAATCTTCTTCATAGTAATGCCAATTCTAATTGGCGG  180
Sbjct  121  CGTTACCGCGCACGCCTTCGTAATAATCTTCTTCATAGTAATGCCAATTCTAATTGGCGG  180

Query 181  CTTTGGGAATTGACTAGTCCCCCTTATGATCGGGGCACCAGACATGGCATTC  240
Sbjct  181  CTTTGGGAATTGACTAGTCCCCCTTATGATCGGGGCACCAGACATGGCATTCGCGG  240

Query 241  GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG  300
Sbjct  241  GAACAACATGAGCTTCTGACTACTTCCGCCCTCATTCCTCCTCCTCCTTGCCTCCTCCGG  300

Query 301  GGTTGAAGCCGGGGCAGGAACCGGGTGAACAGTCTATCCGCCTTTGGCAGGCAATCTTGC  360
Sbjct  301  GGTTGAAGCCGGGGCAGGAACCGGGTGAACAGTCTATCCGCCTTTGGCAGGCAATCTTGC  360

Query 361  CCACGCCGGAGCATCCGTCGACCTAACTATCTTCTCTCTTCATCTAGCAGGTATTTCATC  420
Sbjct  361  CCACGCCGGAGCATCCGTCGACCTAACTATCTTCTCTCTTCATCTAGCAGGTATTTCATC  420

Query 421  AATTCTTGGGGCCATTAATTTTATTAACACAATATTAATAGAAACCCCTGCAAATCTC  480
Sbjct  421  AATTCTTGGGGCCATTAATTTTATTAACACAATATTAATAGAAACCCCTGCAAATCTC  480

Query 481  ACAAATACCAACCCGGGGCTATTGTGATCCGGAGCTCTTTGAACCGCCGTTCTCCTCTTCT  540
Sbjct  481  ACAAATACCAACCCGGGGCTATTGTGATCCGGAGCTCTTTGAACCGCCGTTCTCCTCTTCT  540

Query 541  CTCACTCTCTCTCTCGCTCGCTGCCTGGGATTACAAAATGCTCTCTCAACAGACGGAAATCTAATAC  600
Sbjct  541  CTCACTCTCTCTCTCGCTCGCTGCCTGGGATTACAAAATGCTCTCTCAACAGACGGAAATCTAATAC  600

Query 601  AACCTTCTTTGACCCGCGAGGGGGAGGGGACCAAATTATAACCAACACCTA  652
Sbjct  601  AACCTTCTTTGACCCGCGAGGGGGAGGGGACCAAATTATAACCAACACCTA  652

>gb|KJ768202.1| Alosa fallax voucher MLFP1252 cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
Length=652

Score = 1155 bits (625),  Expect = 0.0
Query 601  AACCTTCTTTGACCCGGCAGGGGGAGGGGACCCAATTTTATACCAACACCTA  652

Sbjct 6130  AACCTTCTTTGACCCGGCAGGGGGAGGGGACCCAATTTTATACCAACACCTA  6181

>gb|KC500181.1| Alosa alosa voucher TR212EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500182.1| Alosa alosa voucher TR213EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500183.1| Alosa alosa voucher TR214EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500184.1| Alosa alosa voucher TR215EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500185.1| Alosa alosa voucher TR216EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500186.1| Alosa alosa voucher TR219EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500188.1| Alosa alosa voucher TR217EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial
gb|KC500189.1| Alosa alosa voucher TR218EK cytochrome oxidase subunit 1 (COI) gene, partial cds; mitochondrial

Length=654
Score =  1147 bits (621),  Expect = 0.0
Identities = 641/651 (98%),  Gaps = 0/651 (0%)
Strand=Plus/Plus
Lambda      K      H
1.33    0.621    1.12

Gapped
Lambda      K      H
1.28    0.460    0.850

Matrix: blastn matrix:i -2
Gap Penalties: Existence: 0, Extension: 0
Number of Sequences: 29086194
Number of Hits to DB: 0
Number of extensions: 0
Number of successful extensions: 0
Number of sequences better than 10: 40
Number of HSP's better than 10 without gapping: 0
Number of HSP's gapped: 40
Number of HSP's successfully gapped: 40
Length of query: 652
Length of database: 82816317834
Length adjustment: 34
Effective length of query: 618
Effective length of database: 81827387238
Effective search space: 50569325313084
Effective search space used: 50569325313084
A: 0
X1: 13 (25.0 bits)
X2: 32 (59.1 bits)
X3: 54 (99.7 bits)
S1: 13 (25.1 bits)
S2: 23 (43.6 bits)
Data Archive Appendix
Data outputs associated with this project are archived as project number 461 media number 1552 on server–based storage at Natural Resources Wales.

The data archive contains:

[A]  The final report in Microsoft Word and Adobe PDF formats.

[B]  An Excel spreadsheet, River Tywi shad egg sampling locations.xlsx detailing the mtDNA Cyt b PCR results for each egg.

[C]  A single FASTA format file, R Tywi COI gene.fas, for the partial COI gene sequences

Metadata for this project is publicly accessible through Natural Resources Wales’ Library Catalogue http://194.83.155.90/olibcgi by searching ‘Dataset Titles’. The metadata is held as record no 115891