Title: Species-specific PCR RFLP for identification of early life history stages of squid and other applications to fisheries research.

Author names and affiliations: Niall J. McKeown\textsuperscript{a}, Jean-Paul Robin\textsuperscript{b}, Paul W. Shaw\textsuperscript{a}

\textsuperscript{a}Institute of Biological, Environmental and Rural Sciences (IBERS), Aberystwyth University, Penglais, Aberystwyth, SY23 3FG, UK
\textsuperscript{b}UMR M100 IFREMER-UCBN, Physiologie et Ecophysiologie des Mollusques Marins, Universite de Caen Basse-Normande, 14032 Caen Cedex, France

Corresponding author: Niall J. McKeown (Email: njm2@aber.ac.uk)

Abstract

As cephalopods fulfil important roles in marine ecosystems and may be especially susceptible to overfishing the predicted expansion of cephalopod fisheries will require improved assessment and management of stocks to ensure ecosystem compatible exploitation. Genetic markers facilitating high throughput accurate species identification, particularly for early life history stages, would considerably benefit cephalopod fisheries research. Reported here is the development of a PCR-RFLP assay for identification of four squid species (\textit{Loligo vulgaris}, \textit{L. forbesi}, \textit{Alloteuthis media} and \textit{A. subulata}) of growing interest as fisheries resources. The assay was used to type morphologically indistinguishable paralarvae collected from Seine Bay and revealed 99\% to be \textit{A. media} despite \textit{a priori} expectations that \textit{L. vulgaris} (present at 1\%) would predominate. As the method can be applied to various life history stages and tissue types it offers considerable potential for use in studies of life history, stock structure, reproduction, recruitment and abundance that are necessary for sustainable management.

Keywords: Cephalopoda – \textit{Loligo} – \textit{Alloteuthis} – species identification – sustainable – fisheries management

Running title: Molecular species identification of squid
1. Introduction

As many traditionally exploited fin fish stocks continue to decline there is growing interest in the expansion of cephalopod fisheries (Boyle 1990; Young et al. 2006). For example, they are becoming an increasingly important fisheries resource in the North East Atlantic (Sacau et al. 2005), an area where their exploitation was previously described as relatively low (Caddy & Rodhouse 1998). The typical short life cycle of cephalopods renders them vulnerable to overfishing (Bravo de Laguna 1989) and as they fulfil important roles in marine ecosystems improved assessment and management of stocks will be vital to ensure ecosystem compatible exploitation (Pierce et al. 1998; Young et al. 2006). Accurate, high throughput, species identification of early life history stages would greatly benefit essential studies of life history, stock structure, reproduction, recruitment and abundance; however, for many cephalopods such stages are morphologically indistinguishable. Species specific genetic markers may be applied to many areas of fisheries science (Teletchea 2009) including the identification of early life history stages (Fox et al. 2005).

*Loligo forbesi* and *L. vulgaris* are, from a fisheries viewpoint, the most important squid species in the northeast Atlantic (Boyle & Pierce 1994). Within the region their ranges are largely overlapping (Guerra & Rocha 1994), however, *L. forbesi* is more abundant in northern waters while *L. vulgaris* dominates the southern part of its range and is regarded as largely absent from Scottish waters (Pierce et al. 1994a, b). In the English Channel fishery landings data suggest that the life cycles of the two species are out of phase (Robin & Boucard-Camou, 1995), proposed to be a mechanism of for reducing competition. *Alloteuthis media* and its congener *A. subulata* are also of growing commercial interest with both species reported in abundance in the English Channel and Irish Sea (Hastie et al. 2009).

The aim of this research was initially to develop a genetic assay permitting species identification of *L. vulgaris, L. forbesi* and *A. media*, and to validate it through analysis of
specimens of known type. The method is also predicted to distinguish *A. subulata* and
although *A. subulata* control types were not available the RFLP information is also reported.
The method was applied to the analysis of 96 wild caught paralarvae from the English
Channel with results validated by DNA sequencing.

2. Materials and methods

Cytochrome oxidase I (COI) sequences available for the four species on GenBank together
with additional sequences collected as part of ongoing squid population genetic research
within our group, were aligned using BioEdit (Hall 1999). The alignment was used to identify
suitable primer sites for cross species PCR amplification of a section of the COI region and
locate restriction enzyme cleavage sites within amplicons using NEBcutter (Vincze et al.
2003). Patterns of cleavage site presence were compared across sequences to identify
potential species-specific diagnostic enzyme combinations for RFLP genotyping. Restriction
digests were performed individually for enzymes following manufacturers (New England
Biolabs) recommendations with products separated on 2% agarose gels and visualised by
Gel-Red staining. To validate the PCR-RFLP assay it was performed on control specimens of
adult *L. vulgaris* (Bay of Biscay [n = 16]; English Channel [n = 24]; Irish Sea [n = 8]), *L.
forbesi* (Irish Sea [n = 6]; Moray Firth [n = 12]) and *A. media* (English Channel [n =19];
Moray Firth [n =7]) that had been morphologically classified and preserved in absolute
ethanol. DNA was extracted from control specimens following Winnepenninckx et al.,
(1993). The PCR-RFLP assay was then employed to genotype paralarvae (*n = 96*) sampled
from Seine Bay as part of the JUVECEPH 2011 survey. DNA was extracted from individual
paralarvae using the Chelex-Proteinase K method in McKeown & Shaw (2008).
3. Results and Discussion

The primers CRESH-F (5’-GAGCAGGCTTAGTTGGTACT-3’) and CRESH-R (5’-ATGGCTCCAGCTAACACAGG-3’) permitted PCR amplification of a 544bp fragment of the COI gene in all individuals tested. PCR’s were performed in 25µl total reaction volumes consisting of 1X PCR Buffer, 2mM MgCl2, 200 µM of each dNTP, 0.2 units of Taq (BioLine), 0.5 µM of each primer, and 2 µl of DNA template. Amplifications involved an initial denaturation step (95 °C for 3 min) followed by 35 cycles of 30 s at 95 °C, 30 s at 52 °C and 30 s at 72 °C. Based on sequence comparisons the combination of SfcI and BccI restriction enzymes was predicted to reciprocally differentiate L. vulgaris, L. forbesi and Alloteuthis spp. by a minimum of two site differences (Table 1) with BccI predicted to distinguish A. media and A. subulata by a single site difference (Table 1). Analysis of control specimens for the three species produced clear restriction patterns with fragment sizes corresponding to expectations from sequences (Fig. 1). The method could not be empirically tested for A. subulata due to the unavailability of control samples, however, the absence of any nucleotide differences within the primer sites, and conservation of restriction enzyme cleavage sites among GenBank sequences support the likelihood of the predicted species specific RFLP. Furthermore, the PCR-RFLP assay successfully excluded A. subulata from the paralarvae samples (described below). Although dependent on a small number of site differences, analysis of reference individuals and GenBank sequences spanning a wide geographical range did not indicate any ambiguity due to intraspecific polymorphisms. However, as the main focus of this research was the English Channel further analysis of individuals from throughout the respective ranges of each species is recommended, particularly for A. media and A. subulata as some localised interspecific haplotype sharing has been reported (Lefkaditou et al. 2011).
Unambiguous genotypes were obtained for all paralarvae (i.e. there were no reaction failures or unrecognised genotypes). Of the 96 paralarvae, 95 were classified as *A. media* and one individual was identified as *L. vulgaris*. The validity of species assignments was confirmed by DNA sequencing of the COI amplicon (using the PCR primers) and use of BLASTn for a subset of individuals. Inaccurate species identification of early life history stages has been reported to severely compromise estimates of stock biomass in other commercially important taxa (e.g. Cod, Fox et al. 2005) and the species ratios reported here were surprising as *L. vulgaris* was expected to predominate. However, when early development stages are correctly identified they can be used to estimate stock sizes and accurate estimates of sustainable catches. Egg identification in plankton surveys is carried out using molecular markers for stock assessment in cod, whiting and haddock in the southern North Sea (Taylor et al. 2002). Species identification of early life history stages can also be used to construct spatial/temporal spawning maps (Fox et al. 2008; Munk et al. 2009). This is highly relevant to cephalopods as widespread human activities, particularly bottom fishing operations, but also shipping, oil exploration and production have been highlighted as potentially damaging to cephalopod spawning areas (Hastie et al. 2009).

At present, for *L. vulgaris* and *L. forbesi*, fishery statistics do not distinguish between the two species since they are of similar appearance and equal commercial value (Chen et al. 2006). In the case of *A. media* and *A. subulata* commonly used morphometric characters (e.g. relative fin length) have been shown to provide inaccurate species identification of adults (Anderson et al. 2008). Species misidentification within fishery landings or stock assessments can severely compromise stock sustainability (Garcia-Vazquez et al. 2012). Application of species specific genetic markers in catch estimates and population surveys should therefore be considered for more accurately estimating stock sizes of the two congener studied here,
particularly in light of recent range shifts reported for *Loligo* (Chen et al. 2006) and suggested for *Alloteuthis* (Anderson et al. 2008) that may rapidly alter mixed fisheries.

DNA based species identification methodologies can be applied to all the different life stages of a marine species and a range of sample types (e.g. whole individuals, biopsies, processed and dried tissue) and thus offer considerable potential to fisheries science. The integration of such methodologies has often been hampered by relatively high per sample costs and by the time and specialist skills needed (Lindeque et al. 2006). By omitting the need for sequencing the PCR-RFLP method described here represents a simple, cost effective means for high throughput screening. Furthermore, although details for two enzymes are reported, digestion with BccI alone is sufficient to discriminate among the four species. The PCR-RFLP method may also be extended to other cephalopod species through judicious selection of cutting enzymes that may be identified from analysis of sequences on GenBank and/or by exploratory sequencing.

**Acknowledgements**

This work was funded as part of the EU CRESH (*Cephalopod Recruitment from English Channel Spawning Habitats*) project. We are indebted to all members of the group for their assistance in obtaining samples and expertise.

**References**


Table 1. Location of restriction enzyme cleavage sites and expected fragment sizes for each enzyme/species combination.

<table>
<thead>
<tr>
<th>Species</th>
<th>Sfc I Cleavage sites – Fragment sizes</th>
<th>Bcc I Cleavage sites – Fragment sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loligo vulgaris</td>
<td>277, 352 – 277,192,75</td>
<td>None</td>
</tr>
<tr>
<td>Loligo forbesi</td>
<td>277 – 277,267</td>
<td>385 – 385,159</td>
</tr>
<tr>
<td>Alloteuthis media</td>
<td>None</td>
<td>254 – 290,254</td>
</tr>
<tr>
<td>Alloteuthis subulata</td>
<td>None</td>
<td>254, 436 – 254, 182, 108</td>
</tr>
</tbody>
</table>
Figure 1. PCR-RFLP patterns for digestion with Sfc I \((3 = L. vulgaris, 4 = L. forbesi, 5 = A. media)\) and Bcc I \((8 = L. vulgaris, 9 = L. forbesi, 10 = A. media)\) and unrestricted PCR products (2 & 7) and molecular weight marker Bioline Hyperladder II (1,6,11).