Utility of 18S rDNA and ITS sequences as population markers for *Lepeophtheirus salmonis* (Copepoda: Caligidae) parasitising Atlantic salmon (*Salmo salar*) in Scotland

1A.P. Shinn, 1B.A. Banks, 2N. Tange, 1J.E. Bron, 1C. Sommerville, 2T. Aoki & 1R. Wootten
1Institute of Aquaculture, University of Stirling, Stirling FK9 4LA, Scotland; 2Laboratory of Genetics and Biochemistry, Department of Aquatic Bioscience, Tokyo University of Fisheries, Konan 4-5-7, Minato-ku, Tokyo 108, Japan

Keywords: Sea louse, *Lepeophtheirus salmonis*, copepod, Crustacea, Atlantic salmon, sequencing, 18S ribosomal DNA, internal transcribed spacer, population genetics

Abstract

Genetic differentiation within the salmon louse *Lepeophtheirus salmonis* (Krøyer, 1837), was investigated by the sequencing of specific nucleotide regions. Partial sequences of the 18S ribosomal RNA gene and the ribosomal internal transcribed spacer (ITS-1) region from single sea lice were amplified by the polymerase chain reaction (PCR). Lice were collected from wild and farmed Atlantic salmon (*Salmo salar* L., 1758) from nine selected localities around the Scottish coastline. A 0.9kb fragment of the 18S ribosomal RNA gene was amplified and compared for several samples of lice which showed no observable differences between the lice from different collection sites confirming the absence of cryptic species. The 454 nucleotide ITS-1 sequence showed differences between derived sequences from 13 sea lice samples from 4 collection sites which included 2 farm sites and 2 sites where lice were taken from wild fish. Across all samples, there was a 92.14% similarity in the ITS-1 sequence. The percentage similarity in the ITS-1 sequence in samples of lice from two fish farms were 99.71% (site A) and 95.72% (site D) but only 86.90% (site B) and 86.03% (site C) similarity was shown in lice samples taken from sites where wild salmonids were caught. The greater similarity between the ITS-1 sequence within farm sites may be attributed to a restricted gene flow within lice populations in Atlantic salmon cage sites.

Contents

| Material and methods | 89 |
| Sample collection and preparation | 90 |
| Genomic DNA extraction | 90 |
| PCR amplification | 91 |
| 18S ribosomal DNA | 91 |
| Internal Transcribed Spacer (ITS) | 91 |
| Cloning into pMOS Blue T-vector | 91 |
| Dideoxynucleotide sequencing and analysis | 92 |

Introduction

The characterization of salmon louse (*Lepeophtheirus salmonis* (Krøyer, 1837)) populations has recently become the subject of considerable interest as a tool to investigate louse population origins and interactions. Nordhagen, Heuch & Schram (1998) investigated experimentally the usefulness of morphometric parameters to identify the origins of the salmon lice populations in Norway. They concluded that it was not possible to discriminate populations of lice on the basis of size, as lice of different origins raised at the same temperature on a common host showed no significant differences in dimensions and similar life-cycle features such as egg production and growth rate. Similarly, in an unpublished experiment, the present authors applying multivariate analyses to morphometric data of several populations of lice taken from naturally infected wild and farmed salmonids sampled throughout a year, concluded that the discrimination of lice on the basis of size was not sufficiently clear to be useful.

Other studies have investigated the allelic frequencies of polymorphic enzymes (allozymes) of populations of the genus *Lepeophtheirus*. For example, Shinn et al. (1998) found that, using nine
polymorphic enzymes, it was possible to obtain some discrimination of lice from wild fish on the east coast of Scotland from both wild and farmed fish sampled on the west coast of Scotland. Isdal, Nylund & Naadvall’s (1997) allozyme study of Norwegian populations of lice similarly found differences between northern and southern populations of farmed lice. De Meeus et al. (1992) used allozymes to demonstrate an adaptive polymorphism to salinity in L. europaensis Zeddam, Berrebi, Renaud, Raibaut & Gabrion, 1988 parasitic on Scophthalmus rhombus L., 1758 and Platichthys flesus L., 1758 which they suggested, when coupled with a restricted movement of hosts between environments of differing salinity, could in time give rise to genetically isolated sub-populations.

Other approaches in the investigation of the population characteristics of L. salmonis include the use of inductively coupled plasma mass spectrometry (ICPMS) to ascertain the elemental signature of adult female lice (Shinn et al., this volume). In this study it was demonstrated that adult female lice from seven localities could be discriminated from each other on the basis of 16 elements and lice from farmed fish could be discriminated from lice on wild salmonids within the same locality.

Molecular techniques permit finer resolution of genetic differences between populations and individuals that cannot be detected using techniques such as morphometrics and allozyme analysis. Todd et al. (1997) using RAPDs, stated that genetic differentiation between populations of lice on farmed salmon and wild salmonids and between individual farm sites was evident to the extent that the presence of putative “farm markers” within certain individual wild lice suggested they originated from farm sites. While RAPD allow for the comparison of amplified DNA bands of similar molecular weight, without direct sampling it is unknown whether these similar sized bands correspond to the same portion of the genome. This study describes an investigation which characterizes populations through determination of the nucleotide sequence of selected areas of the L. salmonis genome, namely the 18S ribosomal RNA gene and the internal transcribed spacer (ITS).

Information resulting from analysis of the 18S rRNA gene sequence has previously been used to determine whether or not the species under investigation contains cryptic species or genetically isolated populations. Similarly, determination of the nucleotide sequence of internal transcribed spacers (ITS) have been successfully implemented as a taxonomic tool to discriminate closely related or morphologically indistinguishable species of parasite e.g apicomplexans (Homan et al., 1997); gyrodactylids (Monogenea) (Cunningham, McGillivray & MacKenzie, 1995; Cunningham, 1997); schistosomes (Digenea) (Littlewood & Johnston, 1995; Van Herwerden, Blair & Agatsuma, 1998); and nematodes (Chilton, Gasser & Beveridge, 1995). Nucleotide variations in ITS regions have also been used to characterise strains or isolates of Cryptocaryon irritans (Diggles & Adlard, 1997); Enterocytozoon bieneusi (Rinder, Katzwinkel-Wladarsch & Loscher, 1997a); Eimeria maxima (Barta et al., 1998); Echinostoma caproni (Morgan & Blair, 1998) and Hymenolepis diminuta (Okamoto et al., 1997). Sequence similarity in the ITS region has also provided evidence for synonymy e.g. Teladorsagia spp. (Trichostrongylidae) (Steven- son, Gasser & Chilton, 1996) and Fasciola spp. (Hashimoto et al., 1997); sub-speciation e.g. Echino- coccus multilocularis (Rinder et al., 1997b) and Paragonimus westermani (Blair et al., 1997) or the identification of cryptic species as has been suggested for geographically distanced isolates of Necator americanus (Romstad et al., 1998). Overall, these studies suggest that the 18S and ITS regions may be used to provide a means of population characterization and it is with this in mind that this study has been undertaken.

Materials and methods

Sample collection and preparation

Sea lice were collected from eight sites in Scotland. Samples were collected from wild Atlantic salmon (Salmo salar L., 1758) at three sites and from farmed Atlantic salmon at five sites (Fig. 1). After collection sea lice were maintained in sea water (10°C) for 48 hrs to remove host products from the gut before use and then homogenised in
buffer (1mM Tris-HCl pH7.0, EDTA 1mM, 2-mercaptoethanol 0.1mM). Samples were then frozen in liquid nitrogen and stored at –70°C.

Genomic DNA extraction

Frozen lice samples were kept on ice, homogenised in a cell lysis buffer (10mM Tris-HCl pH 8.0; 150mM NaCl; 10mM EDTA-NaOH pH 8.0; and 1% SDS) and then treated with protease K (1mg/ml) overnight at 55°C and RNaseA (0.5mg/ml) for one hour at 37°C to digest cellular proteins and RNA. Genomic DNA was then precipitated using 0.3M NaOH-Ac and isolated using a standard phenolic-chloroform extraction (Sambrook et al., 1989) and precipitated in ice-cold 75% ethanol. The purified DNA pellets were dried at room temperature overnight and resuspended in TE buffer (10mM Tris-HCl pH 8.0, 1mM EDTA).

PCR amplification

Universal and species specific primers (given below) were selected for each gene and amplified using PCR (Saiki et al., 1988).

18S ribosomal DNA

The nematode Caenorhabditis elegans rRNA primers complementary to the annealing positions 2074-2098 5’-GGGCAAGTCTGGTGCC-3’ (18S-1) and 2931-2946 5’-GGTCTGTGATGCCCTT-3’ (18S-2) were used (Ellis, Sulston & Coulson, 1986). A 25µl reaction mix was used, containing 25pM of each primer, 50-100ng DNA and a PCR bead (Amersham Pharmacia-Biotech)(1.5 units of Taq, 10mM Tris-HCl, 50mM KCl, 1.5mM MgCl₂, 200µM of each dNTP). The thermocycle programme was 5 min at 95°C, 30 cycles of 30s at 95°C, 30s at 55°C, 90s at 72°C, then 5min at 72°C.

Internal Transcribed Spacer (ITS)

The Gyrodactylus salmonis 18S rDNA forward primer ITS-1 5’-TTTCGGTAGGTGAACCT-3’ corresponding to the annealing position 1,938-1955 (Cunningham, McGillivray & MacKenzie, 1995) and the Schistosoma 28S rDNA reverse primer ITS-2 5’-TCCTCCGCTTAGTATA-3’ (Littlewood & Johnston, 1995) were used as primers in the PCR. The PCR reaction mix was as for the 18S rDNA amplification. The thermocycling programme was 5min at 95°C, 30 cycles of 60s at 95°C, 90s at 52°C, 120s at 72°C, then 5min at 72°C.

Cloning into pMOS Blue T-vector

PCR products were purified and ligated into pMOS Blue T-vector (pT7Blue) (Amersham, UK) using
Novagen blunt cloning kits. Transformations were carried out using NovaBlue competent cells and screened on LB agar plates containing ampicillin (100 µg/ml final conc.), tetracycline (25 µg/ml final conc.), X-gal (0.5 µl/ml final conc.) and IPTG (0.1 µl/ml final conc.).

Dideoxynucleotide sequencing and analysis

Direct colony PCR was used as a rapid screening method to ensure correct insertion of the selected product. Plasmid DNA samples were sequenced using the dideoxy method (Sanger, Nicklen & Coulson, 1977) using fluorescently labelled nucleotides. Either a LI-COR Model 4000L sequencer or a Perkin Elmer ABI PRISM 377 was used following amplification with M13 forward and reverse primers. Sequence data was analysed using MT Navigator ver. 1.02b3 (1994) (Applied Biosystems Inc., Perkin Elmer) and ClustalX (1.5b) for the multiple alignment of genetic sequences. A phylogenetic tree was constructed using Saitou & Nei’s Neighbour Joining Method (NJ) (Saitou & Nei, 1987) which calculates distances (% divergence) between all pairs of sequence from a multiple alignment. The NJ is then applied to the distance matrix and the bootstrap tree which provides confidence estimates for the groupings was drawn using Treeview 1.5 (Page, 1996).

Results

a) 18S ribosomal DNA

The precise sequence for a 0.9kb fragment of the 18S ribosomal RNA gene was determined from eight sea lice (2 from each site – 2 wild and 2 farmed sites), using two clones per specimen (i.e. a total of 16 clones) and from both forward and reverse sequences for each clone. Further, the DNA for this region was extracted, purified, cloned and sequenced for lice taken from the same samples on two separate occasions placing increased confidence in the sequence identity. A 904 nucleotide (nt) sequence was determined (Fig. 2) for each clone with no differences between clones, between individual lice or between sites. The most similar crustacean 18S previously recorded is that of the branchiuran parasite Argulus nobilis with a similarity of 83.92% (914 nt) (Abele, Kim & Felgenhauer, 1989). The nucleotide sequence for this partial sequence of the 18S ribosomal RNA gene has been submitted to the GenBank Data Library under the acquisition number AF043979.

b) Internal Transcribed Spacer (ITS-1)

A total of twenty sea lice from six sites (3 wild and 3 farm) were prepared for cloning. Two or three clones from each louse were prepared for sequencing. Forward and reverse sequences for each clone were obtained, aligned and the sequences compared. Where differences occurred in the forward and reverse sequence and / or between clones from each louse, the sequences were rejected. A total of thirteen lice from four sites (2 wild and 2 farm) were selected and compared and a 454 nucleotide sequence was derived (Fig. 3). Between all samples (n = 13) there was an overall 92.14% similarity in nucleotide sequences with 28 transitions, 28 transversions, 17 indels and 3 inversions (between 6 nucleotide positions). Within the ITS-1 sequence, there are three highly conserved regions between nucleotides 1–32, 91/94–183/189 (x/y denotes the range across all specimens), and 410/411–446/456 and one hypervariable region between base pairs 289/299–401/411. The similarity between all sequences is given in Table 1 and a phylogenetic tree produced from determined bootstrap values (1000 iterations) is given in Fig. 4. To summarize, the percentage similarity between lice within farm sites was 99.71% and 95.72% for sites A and D respectively and 86.90% and 86.03% between lice from within wild sites B and C. This is clearly shown in Fig. 4 where lice from salmon held in farm cages cluster together in the dendrogram, whilst sea lice from wild salmonids are distributed throughout the dendrogram.

A sequence has been submitted to GenBank (acquisition no. AF043980 for a louse collected from a farmed salmonid (site D)). The most similar crustacean sequence in GenBank was the ITS sequence for the ostracod Darwinula brasiliensis.
Pinto & Kotzian (Schon, Butlin, Griffiths & Martins, 1998) with a similarity of 47.94% (315 nt).

Discussion

The ability to identify different sea lice populations will make a major contribution to understanding characteristics including the dispersion and epidemiology of sea lice infections and contribute information to the ongoing dispute between salmon farming and wild salmonid fishery interests. To begin investigating the population characteristics of *L. salmonis*, two regions of the sea louse genome, namely 18S rDNA and ITS-1 were targeted. No variation in a partial fragment (904 base pairs)
A.P. Shinn et al. – 18S rDNA and ITS sequences as population markers for Lepeophtheirus salmonis

of the 18S rRNA gene was found between samples of lice collected from four different sites. The 18S rRNA gene was found to be highly conserved between samples demonstrating that the “L. salmonis” population from farmed and wild salmonids collected in Scottish waters represents a single species i.e. without cryptic species present.

Sequences of the intergenic spacer ITS-1 however, were variable between sites with greater similarity demonstrated between individual lice taken from farmed salmonids than those taken from wild salmonids. Variation within the ITS-1 region is not uncommon and has been reported from Schistosoma where both size and sequence variation among copies of the ITS-1 repeat within individuals occurs (Van Herwerden et al., 1998). Similarly,
Fig. 3. Continued.
Diggles & Adlard (1997) found 11 variable nucleotides in the 169-170 sequence representing ITS-1 between 16 isolates of Cryptocaryon. Further, they reported genetic drift amounting to 2.9-3.5% divergence when the original samples of 3 isolates were compared to those maintained in a laboratory over a 36 month period. The 0.29-4.28% divergence between samples of lice collected from farmed salmonids when compared to background levels (i.e. wild lice) of 13.10-13.97% divergence between samples, may serve to illustrate that cage sites represent a bottleneck to genetic flow. A reduc-
Table 1. Percentage similarity between nucleotide sequences for the internal transcribed spacer (ITS-1) obtained from four populations of *Lepeophtheirus salmonis*. Lice collected at sites A and D were from farmed salmonids whilst those lice collected from sites B and C were from wild salmonids.

<table>
<thead>
<tr>
<th>Sample</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>B1</th>
<th>B2</th>
<th>C1</th>
<th>C2</th>
<th>D1</th>
<th>D2</th>
<th>D3</th>
<th>D4</th>
<th>D5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td></td>
<td>99.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>99.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>99.78</td>
<td>99.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>99.56</td>
<td>99.78</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>99.56</td>
<td>99.56</td>
<td>99.78</td>
<td>99.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>86.49</td>
<td>86.68</td>
<td>86.90</td>
<td>86.90</td>
<td>86.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>97.58</td>
<td>97.80</td>
<td>98.02</td>
<td>98.02</td>
<td>98.24</td>
<td>85.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>86.49</td>
<td>86.68</td>
<td>86.90</td>
<td>86.90</td>
<td>86.68</td>
<td>99.12</td>
<td>86.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>92.75</td>
<td>92.95</td>
<td>93.17</td>
<td>93.17</td>
<td>93.17</td>
<td>91.39</td>
<td>93.30</td>
<td>91.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D2</td>
<td>86.74</td>
<td>86.49</td>
<td>86.71</td>
<td>86.71</td>
<td>86.71</td>
<td>99.78</td>
<td>85.19</td>
<td>93.90</td>
<td>91.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D3</td>
<td>85.90</td>
<td>86.07</td>
<td>86.30</td>
<td>86.52</td>
<td>86.30</td>
<td>99.12</td>
<td>84.78</td>
<td>98.02</td>
<td>90.77</td>
<td>99.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D4</td>
<td>86.49</td>
<td>86.68</td>
<td>86.90</td>
<td>86.90</td>
<td>86.90</td>
<td>99.56</td>
<td>85.81</td>
<td>99.56</td>
<td>91.61</td>
<td>99.34</td>
<td>98.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D5</td>
<td>86.06</td>
<td>86.24</td>
<td>86.46</td>
<td>86.46</td>
<td>86.24</td>
<td>98.45</td>
<td>94.59</td>
<td>99.56</td>
<td>90.73</td>
<td>98.53</td>
<td>97.58</td>
<td>98.90</td>
<td></td>
</tr>
</tbody>
</table>

Summary statistics for lice within sites: within A (n = 6) 99.71 ± 0.23% similarity; within B (n = 1) 86.90% similarity; within C (n = 1) 86.03% similarity; and within D (n = 10) 95.72 ± 4.04% similarity. Across all lice within farm sites (n = 36) there was 91.97 ± 5.70% similarity; within wild sites (n = 6) there was 90.39 ± 6.45% similarity: and across all lice and all sites (n = 78) there was 92.14 ± 5.88% similarity.

in the genetic variability or population isolation within farm sites, may stem from a combination of environmental and anthropogenic induced factors including farming intensity, loch flush rates, fallowing history, the retention of lice larval stages and the frequency of chemotherapeutant application. While fallowing is anticipated to eliminate the resident lice population, it is not known to what extent the gene pool is propagated across the following period to the next stock of fish by either the interaction of the planktonic larvae with neighbouring farm sites or the activity of local salmonids, both of which may serve as a reservoir of genetic identity. Using the features of the hypervariable region identified within the ITS-1 region in this study, it may be possible to design population specific tags or primers that could subsequently be used to ascertain the origin of nauplii and copepodes on farmed and wild salmonids. The acquisition of such population markers could make significant inroads into our ability to monitor the dispersal, fate and impact of planktonic larvae not only on wild salmonids, but between farm sites and across fallow periods.

Acknowledgements

This work was supported by a BBSRC grant to Dr. C. Sommerville and Munbusho funding awarded to Professor T. Aoki.

References


Stevenson LA, Gasser RB, Chilton NB. 1996. The ITS-2 rDNA of Teladorsagia circumcincta, T. trifurcata and T. daviani (Nematoda: Trichostrongylidae) indicates that these taxa are one species. Int. J. Parasitol. 26: 1123-1126.


Received: 4 January 1999