Mar Biotechnol (2012) 14:155–166
DOI 10.1007/s10126-011-9398-z

ORIGINAL ARTICLE

Genomic Resources for Sea Lice: Analysis of ESTs and Mitochondrial Genomes

Motoshige Yasuike · Jong Leong · Stuart G. Jantzen · Kristian R. von Schalburg · Frank Nilsen · Simon R. M. Jones · Ben F. Koop

Received: 9 March 2011 / Accepted: 22 June 2011 / Published online: 12 July 2011

Abstract Sea lice are common parasites of both farmed and wild salmon. Salmon farming constitutes an important economic market in North America, South America, and Northern Europe. Infections with sea lice can result in significant production losses. A compilation of genomic information on different genera of sea lice is an important resource for understanding their biology as well as for the study of population genetics and control strategies. We report on over 150,000 expressed sequence tags (ESTs) from five different species (Pacific *Lepeophtheirus salmonis* (49,672 new ESTs in addition to 14,994 previously reported ESTs), Atlantic *L. salmonis* (57,349 ESTs), *Caligus clemensi* (14,821 ESTs), *Caligus rogercresseyi* (32,135 ESTs), and *Lernaeocera branchialis* (16,441 ESTs)). For each species, ESTs were assembled into complete or partial genes and annotated by comparisons to known proteins in public databases. In addition, whole mitochondrial (mt) genome sequences of *C. clemensi* (13,440 bp) and *C. rogercresseyi* (13,468 bp) were determined and compared to *L. salmonis*. Both nuclear and mtDNA genes show very high levels of sequence divergence between these ectoparasitic copepods suggesting that the different species of sea lice have been in existence for 37–113 million years and that parasitic association with salmonids is also quite ancient. Our ESTs and mtDNA data provide a novel resource for the study of sea louse biology, population genetics, and control strategies. This genomic information provides the material basis for the development of a 38K sea louse microarray that can be used in conjunction with our existing 44K salmon microarray to study host–parasite interactions at the molecular level. This report represents the largest genomic resource for any copepod species to date.

Keywords *Lepeophtheirus salmonis* · *Caligus clemensi* · *C. rogercresseyi* · *Lernaeocera branchialis* · Expressed sequence tags (ESTs) · Mitochondrial genome

Electronic supplementary material The online version of this article (doi:10.1007/s10126-011-9398-z) contains supplementary material, which is available to authorized users.

M. Yasuike · J. Leong · S. G. Jantzen · K. R. von Schalburg · B. F. Koop (✉)
Department of Biology, University of Victoria,
PO Box 3020 STN CSC, Victoria, BC V8W 3N5, Canada
e-mail: bkoop@uvic.ca

F. Nilsen
Department of Biology, University of Bergen,
5020 Bergen, Norway

S. R. M. Jones
Pacific Biological Station, Fisheries and Oceans Canada,
3190 Hammond Bay Road,
Nanaimo, BC V9T 6N7, Canada

Present Address:
M. Yasuike
Aquatic Genomics Research Center, National Research Institute of Fisheries Science, Fisheries Research Agency,
2-12-4 Fukuura, Kanazawa,
Yokohama, Kanagawa 236-8648, Japan

Introduction

Copepods (Copepoda) are a group of small crustaceans found in various aquatic environments and they are described as the most abundant metazoans on earth (Humes 1994). The subclass Copepoda consists of over 250 described families, 2,600 genera, and 21,000 described species classified into ten orders (Walter and Boxshall 2008). Their life histories are diverse; planktonic and benthic copepods are an important ecological link in the

aquatic food chain (Gee 1987; Ohman and Hirche 2001), but approximately one third of marine copepod species live as associates, commensals, or parasites on invertebrates and fishes (Humes 1994).

Parasitic copepods are commonly found both on farmed and wild marine finfish (Johnson and Fast 2004). They feed on host mucus, epidermal cells, tissues, and blood, the result of which causes physiological stress, immune dysfunction, impairment of swimming ability, and possibly death (Boxaspen 2006; Costello 2006; Johnson and Fast 2004; Tully and Nolan 2002). Members of the family Caligidae, especially the genera *Caligus* and *Lepeophtheirus*, are commonly referred to as sea lice (Costello 2006; Johnson et al. 2004; Pike and Wadsworth 1999). They are the most economically important parasites of the world salmon farming industry and may cause direct and indirect economic losses in the industry of €300 million (US $480 million) annually (Costello 2009). In addition, there is concern that salmon farms elevate the risk of sea lice infections on wild salmon beyond that which naturally occurs and lead to a depression in the abundance of wild salmon stocks (Costello 2006; Heuch et al. 2005; Krkošek et al. 2007a; Krkošek et al. 2007b; Todd et al. 2006).

In the North Atlantic Ocean, *Lepeophtheirus salmonis* and *Caligus elongatus* account for the most serious infestations of cultured and wild salmonids (Johnson et al. 2004; Pike and Wadsworth 1999). In the eastern north Pacific Ocean, *L. salmonis* and *Caligus clemensi* have been found on farmed Atlantic salmon (*Salmo salar*) and wild Pacific salmon (*Oncorhynchus* spp.; Beamish et al. 2009; Beamish et al. 2005; Saksida et al. 2007). While *L. salmonis* is prevalent in both Atlantic and Pacific coasts, earlier studies suggested that the Pacific and Atlantic populations of *L. salmonis* are genetically distinct (Tjensvoll et al. 2006; Todd et al. 2004). More recent genomic studies strongly suggest that distinct species of *L. salmonis* exist in the Pacific and Atlantic Oceans following a separation that occurred from 2.5 to 11 million years ago (Boulding et al. 2009; Yazawa et al. 2008). These parasites are referred to herein as the Pacific and Atlantic forms of *L. salmonis*, respectively. In the southern hemisphere, *Caligus rogercresseyi* is the dominant species affecting salmonid aquaculture in Chile where the parasites were found on farmed salmon in 99% of the established cultured cages (Boxshall and Bravo 2000; Carvajal et al. 1998).

Lepeophtheirus and *Caligus* species are distinguished from each other based on morphological characters (Kabata 1979). The life cycle in *L. salmonis* has a total of ten developmental stages, while *C. elongatus* and *C. rogercresseyi* are similar but appear to lack pre-adult stages (Piasecki and MacKinnon 1995; González and Carvajal 2003). The host range of *L. salmonis* mainly includes salmonids but the parasite has also been reported from non-salmonid hosts, including sticklebacks, that co-occur with salmon (Jones et al. 2006). In contrast, some *Caligus* species have a broad host range of salmonids and non-salmonids (Costello 2006; Johnson et al. 2004). Among its salmonid hosts, *L. salmonis* displays clear preferences, with heaviest infestations and greatest impacts occurring on Atlantic salmon (*S. salar*) and sea trout (*Onchorhynchus trutta*) followed by rainbow trout (*Onchorhynchus mykiss*), chinook (*Onchorhynchus tshawytscha*), and coho salmon (*Onchorhynchus kisutch*; Dawson et al. 1997; Fast et al. 2002; Johnson and Albright 1992). In contrast, *C. rogercresseyi* occurs in higher numbers on caged rainbow trout compared with Atlantic or coho salmon (González et al. 2000). Thus, while *L. salmonis* and *Caligus* species exhibit similar parasitic life history strategies, they display considerable differences in morphology, life cycle, and host range.

Another parasite, *Lernaeocera branchialis* belongs to the copepod family Pennellidae and is distantly related to the caligid copepods, and this species is commonly found on gadoids, particularly Atlantic cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*) in the North Atlantic Ocean and North Sea (Bricknell et al. 2006; Smith et al. 2007). This parasite has a negative impact on wild gadoids and is a potentially serious pathogen of farmed Atlantic cod (Smith et al. 2007). A compilation of genomic information on parasitic copepods is an important tool for understanding their biology as well as for the study of population genetics and control strategies.

In this study, we report on over 150,000 expressed sequence tags (ESTs) obtained from Pacific *L. salmonis* (49,672 new ESTs in addition to 14,994 previously reported ESTs), Atlantic *L. salmonis* (57,349 ESTs), *C. clemensi* (14,821 ESTs), *C. rogercresseyi* (32,135 ESTs), and *L. branchialis* (16,441 ESTs). These ESTs were assembled into complete or partial genes and annotated by comparisons to known proteins in public databases. In addition, whole mitochondrial (mt) genome sequences of two *Caligus* species, *C. clemensi* and *C. rogercresseyi*, were determined and compared to each other and to *L. salmonis*. These studies show high levels of sequence divergence in nuclear and mtDNA genes. This report describes the production and characteristics of the largest genomic resource for copepods.

Materials and Methods

EST Analysis

Specimens belonging to the Pacific (British Columbia, Canada (BC)) and Atlantic forms of *L. salmonis* (Norway and New Brunswick, Canada), *C. clemensi* (BC), *C. rogercresseyi* (Chile), and *L. branchialis* (Scotland, UK)

were collected and stored at −80°C or in RNAlater (Invitrogen) until RNA extraction. Total RNA was extracted from whole bodies (from various life stages and both sexes) using TRIzol reagent (Invitrogen) and spin-column purified using RNeasy Mini kits (Qiagen). The purified RNAs were then quantified and quality checked by spectrophotometer (NanoDrop Technologies) and agarose gel, respectively. Approximately 1.0–3.0 μg of total RNA was converted into cDNA and normalized and was directionally cloned into pAL 17.3 vector (Evrogen Co.).

Clones from each library were robotically arrayed in 384-well microtiter plates as detailed previously (Koop et al. 2008). Plasmid DNAs were extracted and sequenced on an ABI 3730 DNA analyzer (Applied Biosystems) with M13 forward and M13 reverse primers (*L. salmonis* and *C. rogercresseyi*) or with M13 forward and SP6 primers (*C. clemensi* and *L. branchialis*). These sequence primers are shown in supplemental Table 1. The resulting ESTs were assembled with CAP3 (Huang and Madan 1999) with default parameters. The assembled total contigs (clusters + singletons) were annotated using RPS-BLAST and BLASTX comparisons with the Conserved Domain Database (CDD) and SwissProt (Bairoch and Apweiler 1996), respectively. The best BLAST match (*E* value threshold of 1 E^{-10}) was used to identify contigs. Contigs that did not meet this threshold were annotated as "unknown."

Reference full-length cDNAs (FLcDNAs) were identified as detailed previously (Leong et al. 2010). A single clone containing an entire coding sequence (CDS) for a gene product is considered a reference FLcDNA.

Complete Mitochondrial Genome Sequences of *C. clemensi* and *C. rogercresseyi*

The total genomic DNAs were extracted from an adult male *C. clemensi* and *C. rogercresseyi* as previously described (Yazawa et al. 2008). A sample placed in 5% Chelex-100 resin (Sigma) solution (5% Chelex-100 resin, 0.2% SDS in TE, with proteinase K (100 μg/ml)) was incubated for 30 min at 55°C, and the proteinase K was then inactivated for 10 min at 90°C. The sequence determination of the complete *C. rogercresseyi* mt genome was carried out as previously described (Yazawa et al. 2008). The PCR primer sets that were used were designed for 15 fragments (Supplemental Table 1) based on the EST sequences encoding mtDNA. PCR amplification was performed using 1.0 μl of extracted total genomic DNA of *C. rogercresseyi* with an initial denaturation step of 2 min at 95°C and then 30 cycles as follows: 30 s of denaturation at 95°C, 30 s of annealing at 55°C, and 3 min of extension at 72°C. PCR products were cloned into pCR2.1 vector (TA Cloning Kit, Invitrogen) with the manufacturer's protocol, and each positive PCR product was sequenced as described above.

Table 1 Sea lice EST project summary

	L. salmonis (P)[a]	*L. salmonis* (A)[b]	*C. ciemensi*	*C. rogercresseyi*	*L. branchialis*
Number of clones[c]	38,880[e]	51,607	7,680	19,200	8,448
Number of sequences[d]	64,666[e]	5734.9[i]	14,821	32,135	16,441
Average trimmed EST length (bp)[f]	756	644	790	730	749
Number of contigs[g]	11,922	*9,113*	4,392	8,251	4,239
Number of singletons	4,186	5,145	1,662	3,106	2,199
Number of putative transcripts	16,108	14,466	6,054	11,357	6,438
Maximum contig size (no. of ESTs)	554	1482	15	34	21
Average contig size (no. of ESTs)	4.0	4.0	2.5	2.8	2.6
Number of transcripts with BLAST hits[h]	7,157	6,726	3,775	5,830	3,951
Percent with significant BLAST hits	44.4%	46.5%	62.4%	51.3%	61.4%

[a] *L. salmonis* Pacific form

[b] *L. salmonis* Atlantic (Canada, Norway) form

[c] Number of clones which from at least one sequence (5′ or 3′) was obtained

[d] Number of 5′ and 3′ EST sequences obtained

[e] Twenty-eight thousand thirty-two clones and 49,672 sequences were obtained from this study, while 5,760 clones and 14,994 sequences were previously reported (Yazawa et al. 2008)

[f] Vector, low quality, and contaminating bacterial sequences are trimmed

[g] A contig (contiguous sequence) contains two or more ESTs

[h] Number of transcripts that have a RPS-BLAST or BLASTX hit of less than 1 E^{-10} to the Conserved Domain Database (CDD) or SwissProt databases

[i] 28K sequences were obtained from F. Nilsen (University of Bergen, Norway)

The entire mt genome for *C. clemensi* was amplified by a long PCR method for three long fragments (5.4, 5.0, and 3.0 kb) and by PCR as described above for one short fragment (0.8 kb). The three PCR fragments were amplified using the PCR primer sets shown in Supplemental Table 1 and by using Long PCR Enzyme mix (Fermentas) following the manufacturer's protocol. The long PCR amplification was performed using 100 ng of extracted total genomic DNA of *C. clemensi* with an initial denaturation step of 2 min at 94°C and then a two-step PCR procedure (40 cycles of 95°C for 10 s and 68°C for 7 min), and 10 min of final extension. The three long PCR products were cloned into pCR-XL-TOPO vector (Invitrogen) with the manufacturer's protocol, and each positive PCR product was sequenced by primer walking (supplemental Table 1). The one short fragment was cloned into pCR2.1 vector and sequenced as described above.

Protein-coding and rRNA genes of *C. clemensi* and *C. rogercresseyi* were identified by alignment with the Pacific *L. salmonis* mt gene sequences (GenBank: EU288200). The majority of the tRNA genes was identified using tRNAscan-SE 1.21(Lowe and Eddy 1997), using the same parameters as described by Tjensvoll et al. (2005). The remaining tRNA genes were identified based on the sequence homology with *L. salmonis* tRNA sequences.

Pair-wise Kimura two-parameter (K2P) distances (Kimura 1980) of 16S rRNA and *cox1*genes for *C. clemensi*, *C. rogercresseyi*, and Pacific *L. salmonis* were calculated in MEGA5 (Tamura et al. 2007), with default settings.

Results and Discussion

EST Analysis and Comparison of the Nuclear Genes

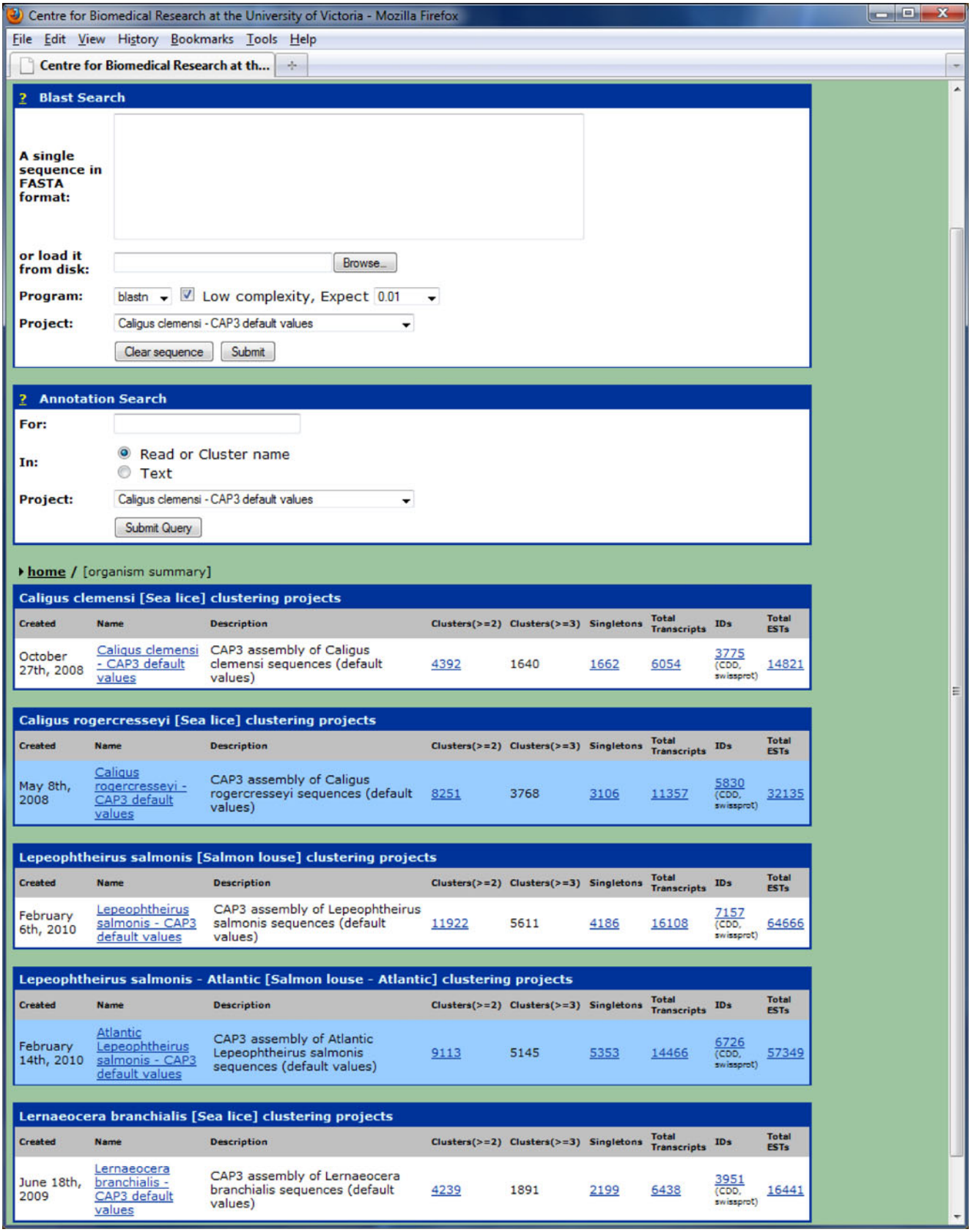
Normalized cDNA libraries were constructed for Pacific *L. salmonis*, Atlantic *L. salmonis*, *C. clemensi*, *C. rogercresseyi*, and *L. branchialis*. The 114,967 clones obtained from these cDNA libraries (28,032 Pacific *L. salmonis*, 51,607 Atlantic *L. salmonis*, 7,680 *C. clemensi*, 19,200 *C. rogercresseyi*, and 8,448 *L. branchialis*) were sequenced with M13 forward and M13 reverse (*L. salmonis* and *C. rogercresseyi*) or with M13 forward and SP6 primers (*C. clemensi* and *L. branchialis*). A summary of the EST project is shown in Table 1. From these clones, 153,977 high-quality ESTs were obtained from Pacific *L. salmonis* (49,672 ESTs), Atlantic *L. salmonis* (57,349 ESTs), *C. clemensi* (14,821 ESTs), *C. rogercresseyi* (32,135 ESTs), and *L. branchialis* (16,441 ESTs). The average trimmed length of these ESTs was 734 bp. These EST sequences are available in GenBank.

The 49,672 Pacific *L. salmonis* ESTs obtained in this study along with 14,994 Pacific *L. salmonis* ESTs from our previous study (Yazawa et al. 2008) were assembled into 11,922 contigs and 4,186 singletons (16,108 putative transcripts). There is a total of 14,466 putative transcripts for Atlantic *L. salmonis*, 6,054 for *C. clemensi*, 11,357 for *C. rogercresseyi*, and 6,438 for *L. branchialis*. These putative transcripts were annotated using RPS-BLAST and BLASTX comparisons with the CDD and SwissProt (Bairoch and Apweiler 1996), respectively. The best match (*E* value threshold of 1 E^{-10}) was used to identify putative transcripts. Of the 16,108 Pacific *L. salmonis* putative transcripts, 7,157 (44.4%) matched at least one entry in the databases while the others remain unidentified. Similarly, 6,726 (46.5%) Atlantic *L. salmonis*, 3,775 (62.4%) *C. clemensi*, 5,830 (51.3%) *C. rogercresseyi*, and 3,951 (61.4%) *L. branchialis* putative transcripts have significant BLAST hits (Table 1).

A collection of reference FLcDNA clones is an important resource for identifying genes, determining their structural features and for experimental analysis of gene functions. Possible reference FLcDNAs were defined as having an entire open reading frame (ORF) corresponding to a full-length protein and were identified as described previously (Leong et al. 2010). Using an *E* value filter of $E \leq 10^{-5}$, the top ten SwissProt high-scoring segment pairs (HSPs) from BLASTX for each putative transcript were analyzed in succession to identify the correct ORF. Of the 16,108 Pacific *L. salmonis* putative transcripts, 1,435 transcripts were identified as possible FLcDNAs. There are 1,086 Atlantic *L. salmonis* FLcDNAs, 1,223 *C. clemensi* FLcDNAs, and 1,574 *C. rogercresseyi* FLcDNAs. These reference FLcDNAs were submitted to NCBI's FLIC database.

A relational database with an intuitive web interface was developed to process and display the large quantities of EST data, their assemblies, and their associated annotation information (Fig. 1). This interface provides the ability to search using sequence data, identifiers, accession numbers, and descriptive keywords. The BLAST search allows users to perform homology searches with sequences of interest, identifying potential transcripts names, and then visualizing these sequences and EST alignments. These EST contigs have predicted ORFs and BLASTX HSPs displayed in a single view. This database contributes to the identification and analysis of proteins and to the development of microarrays for gene expression analyses.

Fig. 1 Screenshot of sea lice EST contig summary and search tools. The *top panel* allows users to perform homology searches for sequences of interest. The *second* provides the ability to search using sequence data, identifiers, accession numbers, and descriptive keywords. The *third* to *seventh panels* show a summary of the EST clustering results of *C. clemensi*, *C. rogercressyi*, Pacific *L. salmonis*, Atlantic *L. salmonis*, and *L. branchialis*, respectively


Centre for Biomedical Research at the University of Victoria - Mozilla Firefox
File Edit View History Bookmarks Tools Help
Centre for Biomedical Research at th...
? Blast Search
A single sequence in FASTA format:
or load it from disk:
Browse...
Program:
blastn
Low complexity, Expect
0.01
Project:
Caligus clemensi - CAP3 default values
Clear sequence
Submit
? Annotation Search
For:
In:
Read or Cluster name
Text
Project:
Caligus clemensi - CAP3 default values
Submit Query
home / [organism summary]
Caligus clemensi [Sea lice] clustering projects
Created Name Description Clusters(>=2) Clusters(>=3) Singletons Total Transcripts IDs Total ESTs
October 27th, 2008 Caligus clemensi - CAP3 default values CAP3 assembly of Caligus clemensi sequences (default values) 4392 1640 1662 6054 3775 (CDD, swissprot) 14821
Caligus rogercresseyi [Sea lice] clustering projects
Created Name Description Clusters(>=2) Clusters(>=3) Singletons Total Transcripts IDs Total ESTs
May 8th, 2008 Caligus rogercresseyi - CAP3 default values CAP3 assembly of Caligus rogercresseyi sequences (default values) 8251 3768 3106 11357 5830 (CDD, swissprot) 32135
Lepeophtheirus salmonis [Salmon louse] clustering projects
Created Name Description Clusters(>=2) Clusters(>=3) Singletons Total Transcripts IDs Total ESTs
February 6th, 2010 Lepeophtheirus salmonis - CAP3 default values CAP3 assembly of Lepeophtheirus salmonis sequences (default values) 11922 5611 4186 16108 7157 (CDD, swissprot) 64666
Lepeophtheirus salmonis - Atlantic [Salmon louse - Atlantic] clustering projects
Created Name Description Clusters(>=2) Clusters(>=3) Singletons Total Transcripts IDs Total ESTs
February 14th, 2010 Atlantic Lepeophtheirus salmonis - CAP3 default values CAP3 assembly of Atlantic Lepeophtheirus salmonis sequences (default values) 9113 5145 5353 14466 6726 (CDD, swissprot) 57349
Lernaeocera branchialis [Sea lice] clustering projects
Created Name Description Clusters(>=2) Clusters(>=3) Singletons Total Transcripts IDs Total ESTs
June 18th, 2009 Lernaeocera branchialis - CAP3 default values CAP3 assembly of Lernaeocera branchialis sequences (default values) 4239 1891 2199 6438 3951 (CDD, swissprot) 16441

Sequence similarities and putative transcripts were compared among the nuclear genes of the five copepods (Pacific *L. salmonis*, Atlantic *L. salmonis*, *C. clemensi*, *C. rogercresseyi*, and *L. branchialis*) by BLASTN for nucleotide (nt) sequences and tBLASTX for amino acid (aa) sequences (Table 2). We previously reported that a total of 155 nuclear genes from Pacific and Atlantic *L. salmonis* showed an average of 96.8% nt identity over an average of 756 bp (Yazawa et al. 2008). In this study, a total of 8,121 nucleotide and 8,827 translated aa sequences matched between the Pacific and Atlantic *L. salmonis* putative transcripts. These sequences showed an average of 96% identity at the nt level over an average of 626 bp and 88% at the aa level over an average of 187 aa (Table 2). Nuclear gene sequences were quite different not only between the genera *Caligus* and *Lepeophtheirus* (81–82% nt, 70–72% aa identities), but also between the two *Caligus* species (83% nt, 71% aa identities; Table 2). The range of nuclear gene sequence divergence was quite similar among these species (17–19% nt and 28–30% aa sequence divergences). As expected, nucleotide sequences of *L. branchialis*, the only species examined from the family Pennellidae, were very different from the caligid sequences: only 4–6% of the total queries (254–405 sequences) matched the nuclear genes of the four other copepods. We speculate that the matched genes are conserved among copepods and therefore we could not determine the divergence between nt sequences of *L. branchialis* and the four caligid copepods. However, the 2,634–3,375 translated aa sequences of *L. branchialis* (44–52% of query sequences) did show significant matches with sequences of the four other copepods. These translated aa sequences showed 59–62% identities over averages of 121–132 aa (Table 2). Although these comparisons provide only a very rough estimate of overall sequence similarity, they clearly indicate a high level of sequence divergence among these copepods nuclear genes.

Mitochondrial Genome Sequences of *L. salmonis*, *C. clemensi*, and *C. rogercresseyi*

Metazoan mt genomes typically range between 15 and 20 kb in size, containing 37 genes: 13 protein-encoding genes (PCGs), 22 transfer RNA (tRNA) genes, two ribosomal RNA (rRNA) genes and a major non-coding region (NCR; Boore 1999). In this study, whole mt genome sequences of two *Caligus* species, *C. clemensi* and *C. rogercresseyi*, were determined. The sizes of the entire mt genomes were 13,440 bp for *C. clemensi* [Genbank: HQ157566] and 13,468 bp for *C. rogercresseyi* [Genbank: HQ157565], and thus, these mt genomes are the shortest among 57 crustacean mt genomes (average length: 15,785 bp) reported so far (Genbank: November 2010). There are two reasons for the small size of these mt genomes. First, the major NCRs of the *C. clemensi* (104 bp) and *C. rogercresseyi* (129 bp) mt genomes were much shorter than that of *L. salmonis* (Pacific form, 1,441 bp; Atlantic form, 2,146 bp) and that of other crustaceans (average length, 875 bp), except for that of the amphipod *Metacrangonyx longipes* (76 bp; Bauzà-Ribot et al. 2009). Second, while both *Caligus* mt genomes contained the typical set of 12 protein-encoding, 21 tRNA and two rRNA genes found in other animal mt genomes, both mt genomes lacked the PCG, *nad4L*, and a tRNA gene, $trnL_2$ (CUN).

Interestingly, the *C. clemensi* mt genome is adenine and thymine (A + T)-rich (PCG, 74.5%; whole genome, 75.6%) compared to *C. rogercresseyi* and *L. salmonis* (PCG, 63.6–64.9%; whole genome, 65.2–66.5%; Supplemental Table 2). In crustaceans, the mt genomic A–T content values range from 60.9% for *Ligia oceanica* (Isopoda; Kilpert and Podsiadlowski 2006) to 77.8% for *Argulus americanus* (Branchiura; Lavrov et al. 2004). The reason for the variability in A–T richness within the mitochondrial genome among taxa is not clear.

Like the nuclear genes, the mtDNA gene sequences also exhibited large divergence, not only between *L. salmonis* and the two *Caligus* species (66.7–68.8% nt and 64.2–65.4% aa identities), but also between the two *Caligus* species (68.8% nt and 63.6% aa identities). The range of mtDNA sequence divergence was quite similar among the three caligid copepods. The percent nt and aa identities among the *L. salmonis*, *C. clemensi*, and *C. rogercresseyi* sequences are 63.6–68.8% (Table 3). The *cox1* gene is the most conserved PCG among the three mt genomes (79.1–82.6% nt and 91.2–94.1% aa identities), while *nad2*, *nad4*, *nad5*, and *nad6* exhibit a large sequence divergence (56.1–62.2% nt and 40.0–51.9% aa identities; Table 3).

Hebert et al. (2003) reported that *cox1* divergences among the 13,320 species in the animal kingdom ranged from a low of 0.0% to a high of 53.7% and the mean divergence value of 11.3%. The *cox1* divergences in the Crustacea showed the mean species divergence value of 15.4% (Hebert et al. 2003). Interestingly, our present study showed that the *cox1* divergences among the three caligid copepods were higher than the mean divergence value of Crustacea. The *cox1* interspecific divergence between *C. clemensi* and *C. rogercresseyi* is 20.2% and between the genera *Caligus* and *Lepeoptheirus* 26.0%. Øines and Schram (2008) compared among the *cox* fragment (a total 504 aligned base pairs) of 18 caligid copepods and the 16S rRNA fragment (a total of 438 aligned base pairs) of 11 caligid copepods. They found that an average K2P distance of *cox1* were 0.218 and those of 16S rRNA were 0.221 (Øines and Schram 2008). In the present study, the K2P distance of *cox1* (a total of 1,539 aligned base pairs) among the *L. salmonis*, *C. clemensi*, and *C. rogercresseyi* is 0.202–

Table 2 Comparison of the Pacific and Atlantic *L. salmonis*, *C. clemensi*, *C. rogercressyi*, and *L. branchialis* nuclear genes

	Type	No. of queries	No. of matches	Percentage with match[b]	Average length	Maximum length	Standard deviation for length	Average identities	Maximum identities	Minimum identities	Standard deviation for% identities	Average positive AAs
Atlantic form *L. salmonis* vs Pacific form *L. salmonis*	blastn	14,466	8,121	56%	626 bp	2,891 bp	365.25	96%	100%	78%	4.41	–
	tblastx	14,466	8,827	61%	187 aa	820 aa	110.46	88%	100%	22%	18.15	91%
C. clemensi vs Pacific form *L. salmonis*	blastn	6,054	1,598	26%	327 bp	1,316 bp	208.89	81%	98%	76%	2.83	–
	tblastx	6,054	3,852	64%	151 aa	569 aa	79.38	72%	100%	12%	17.09	83%
C. clemensi vs Atlantic form *L. salmonis*	blastn	6,054	1,595	26%	318 bp	1,316 bp	197.67	81%	98%	76%	2.98	–
	tblastx	6,054	4,096	68%	145 aa	539 aa	75.68	72%	100%	16%	16.65	83%
C. clemensi vs *C. rogercressyi*	blastn	6,054	1,893	31%	338 bp	1,305 bp	202.55	83%	100%	77%	3.11	–
	tblastx	6,054	3,715	61%	142 aa	456 aa	71.72	71%	100%	21%	17.09	82%
C. clemensi vs *L. branchialis*	blastn	6,054	257	4%	278 bp	1,226 bp	186.83	81%	97%	77%	3.01	–
	tblastx	6,054	2,634	44%	125 aa	436 aa	63.27	59%	100%	21%	16.47	74%
C. rogercressyi vs Pacific form *L. salmonis*	blastn	11,357	1,931	17%	301 bp	1,309 bp	178.46	82%	99%	76%	2.86	–
	tblastx	11,357	5,937	52%	139 aa	557 aa	72.40	70%	100%	18%	17.45	81%
C. rogercressyi vs Atlantic form *L. salmonis*	blastn	11,357	1,973	17%	292 bp	1,498 bp	175.48	82%	100%	76%	3.40	–
	tblastx	11,357	6,383	56%	133 aa	521 aa	68.41	70%	100%	21%	16.81	82%
L. branchialis vs Pacific form *L. salmonis*	blastn	6,438	417	6%	264 bp	1,079 bp	168.51	81%	99%	76%	2.99	–
	tblastx	6,438	3,284	51%	132 aa	466 aa	69.24	62%	100%	22%	16.22	77%
L. branchialis vs Atlantic form *L. salmonis*	blastn	6,438	405	6%	260 bp	1,079 bp	163.04	81%	98%	77%	2.75	–
	tblastx	6,438	3,375	52%	126 aa	466 aa	64.90	62%	100%	22%	16.31	76%
L. branchialis vs *C. rogercressyi*	blastn	6,438	254	4%	250 bp	1,211 bp	156.59	81%	99%	76%	2.73	–
	tblastx	6,438	3,021	47%	121 aa	465 aa	59.81	59%	100%	17%	16.38	74%

[a] The number of queries that had BLASTN hit with an *E* value <1 E^{-10} and 100 bp minimum of alignment length or that had tBLASTX with an *E* value <1 E^{-10} and 50 aa minimum of alignment length

[b] First match that conformed to parameters was taken from the top five hits of blast output. If no suitable match was found in the top five hits, it was not included in the results

Table 3 Comparison of the *L. salmonis*, *C. clemensi*, and *C. rogercressyi* mtDNA genes

Genes	In nucleic sequence (%)				In deduced amino acid sequence (%)			
	Pacific form *L. salmonis* vs *C. clemensi*	Pacific form *L. salmonis* vs *C. rogercressyi*	*C. clemensi* vs *C. rogercressyi*	Atlantic form *L. salmonis* vs Pacific form *L. salmonis*	Pacific form *L. salmonis* vs *C. clemensi*	Pacific form *L. salmonis* vs *C. rogercressyi*	*C. clemensi* vs *C. rogercressyi*	Atlantic form *L. salmonis* vs Pacific form *L. salmonis*
rrnS	77.2	76.4	74.9	98.8	–	–	–	–
rrnL	68.3	67.2	71.2	96.9	–	–	–	–
atp8[a]	72.0	67.7	72.0	96.8	–	–	–	–
atp6	63.9	65.4	65.3	91.9	61.5	66.7	60.3	95.9
cob	71.0	70.8	71.3	93.7	79.7	80.3	77.6	98.5
coxi	79.1	77.9	82.6	92.9	90.8	91.2	94.1	99.2
cox2	76.6	75.7	78.5	93.5	75.4	81.2	85.5	100.0
cox3	73.2	71.7	72.4	92.0	75.6	82.5	79.2	98.2
nadi	72.2	71.6	70.5	92.8	69.2	75.3	66.8	96.9
nad2	57.9	57.8	59.3	90.9	40.5	45.8	43.2	94.8
nad3	68.5	57.8	65.0	91.6	64.1	54.7	59.3	97.5
nad4	61.9	58.8	58.7	91.2	49.1	49.3	44.0	92.6
nad4L[b]	N.A.	N.A	N.A.	94.3	–	–	–	97.3
nad5	62.2	58.8	61.8	90.7	51.9	50.4	49.4	95.5
nad6	59.6	56.1	59.1	93.8	48.3	42.0	40.0	97.3
Average	68.8	66.7	68.8	93.5	64.2	65.4	63.6	97.0

[a] Comparisons of amino acid sequences of *atp8* genes were not conducted because these sequences are very short in size (31 aa)

[b] *nad4L* genes are absent in the two *Caligus* species

0.270 (Supplemental Table 3), which is similar to an average K2P distance found by Øines and Schram (2008). However, the 16S rRNA among the three copepods showed a very high genetic distance. The K2P distance of the 16S rRNA (a total of 1,085 aligned base pairs) were 0.333 between *C. clemensi* and *C. rogercresseyi* and 0.422 (Supplemental Table 3). These molecular distance values support an ancient separation between *C. clemensi* and *C. rogercresseyi* as well as between *Lepeoptheirus* and *Caligus*.

In our previous study, a molecular clock based on 16S rRNA and calibrated by copepod data suggested that the forms of *L. salmonis* existing in the Pacific and Atlantic Oceans evolved from a common ancestor following a separation that occurred from 4.6–11 million years ago (Yazawa et al. 2008). In this study, the molecular estimates of the age of divergence between the *L. salmonis* (Pacific) and the two *Caligus* species were calculated based on the 16S rRNA gene using the same method as previously reported (Yazawa et al. 2008). The results suggest that the separation between the *L. salmonis* (Pacific) and the two *Caligus* species occurred approximately 45–113 million years ago (Table 4). In addition, the separation between the two *Caligus* species was estimated to have occurred 37–87 million years ago (Table 4). Salmonids are believed to have evolved from an ancestor in which a whole genome duplication event occurred 25–100 million years ago (Ohno 1970). Thus, our present results suggest that the *L. salmonis* and *C. clemensi* have been in existence for 45–106 million years and that parasitic association with salmonids is likely also quite ancient (Table 4).

The order of the genes in the two *Caligus* mt genomes is identical despite extensive sequence divergence. In contrast, the order of genes in the two *Caligus* mt genomes is quite different from that in the *L. salmonis* mt genome. The gene arrangement in the region between *nad4* and $trnL_1$ (UUR; approximately 10 kb) is well conserved between *L. salmonis* and the *Caligus* species. However, the gene arrangements adjacent to their control regions (CRs) are very distinct, and the *Caligus* mt genomes show a novel gene arrangement (Fig. 2). The region around the CR is more prone to gene rearrangement in both vertebrate (Macey et al. 1997) and invertebrate (Dowton and Austin 1999) mt genomes. In the *L. salmonis* mt genomes, the region between $trnK_2$ and *trnR* (six tRNA and *atp6* genes) is in a row (Tjensvoll et al. 2005; Yazawa et al. 2008). However, in the *Caligus* mt genomes, this region is separated by *rrnS-nad6-trnA-*$trnK_1$*-trnQ-trnT-cytb-CR*, and divided into $trnK_2$*-trnN-trnG-trnV* and *atp6-trnY-trnR* (*trnY* also had a position change; Fig. 2). As mentioned above, the *nad4L* and $trnL_2$ (CUN) genes are absent in the *Caligus* mt genomes. These two genes normally reside in this region

Table 4 Ranges of 16S rRNA gene divergence based on Kimura two-parameter distance and crustacean molecular clock calibrations

	Distance (K2P)	Divergence Range (Myr)			
		Ano	Fid	Gra (low)	Gra (high)
Pacific form *L. salmonis* vs. *C. clemensi*	0.405	106.2	45.0	62.3	46.0
Pacific form *L. salmonis* vs. *C. rogercressyi*	0.431	113.0	47.8	66.2	48.9
C. clemensi vs. *C. rogercressyi*	0.333	87.4	37.0	51.2	37.8

The values for "Distance" are the Kimura two-parameter (K2P) distance between the species. Rates of molecular evolution used for the 16S rRNA gene include 0.38% K2P/million year (Myr) for anomurans (Ano; Cunningham et al. 1992), 0.90% K2P/Myr for fiddler crabs (Fid; Sturmbauer et al. 1996), and 0.65 (low)–0.88% (high) K2P/Myr obtained from grapsid crabs (Gra; Schubart et al. 1998)

and have probably been lost due to rearrangement. It is likely that this rearrangement event also has led to the trimming of their CRs in the two *Caligus* mt genomes.

In the mt genomes of most animals, *nad4L* and *atp8* are located together with *nad4* and *atp6*, respectively (*nad4L-nad4* and *atp8-atp6*), and *nad4L- nad4* and *atp8-atp6* are translated from a single mRNA (Amalric et al. 1978; Berthier et al. 1986). In contrast, several genes separate *nad4* and *nad4L* in the mt genomes of *L. salmonis* and in the mt genomes of all copepods characterized so far: *Tigriopus japonicas* (Machida et al. 2002), *Tigriopus californicus* (Burton et al. 2007), *Paracyclopina nana* (Ki et al. 2009), and the partially sequenced mt genomes of *Eucalanus bungii* and *Neocalanus cristatus* (Machida et al. 2004). The *atp6* and *atp8* are also separated in the two *Caligus* species and in *L. salmonis* (Fig. 2). In addition, it has been reported that *atp8* is absent in the mt genome of *P. nana* (Ki et al. 2009). Thus, it is most likely that these separations of *nad4-nad4L* and *atp6-atp8* occurred during copepod evolution and led to the loss of *nad4L* in the two *Caligus* species and to the loss of *atp8* in the *P. nana*.

In summary, the mtDNA genes of the two *Caligus* species showed high levels of sequence divergence (Table 3). The A+T content is also quite different between

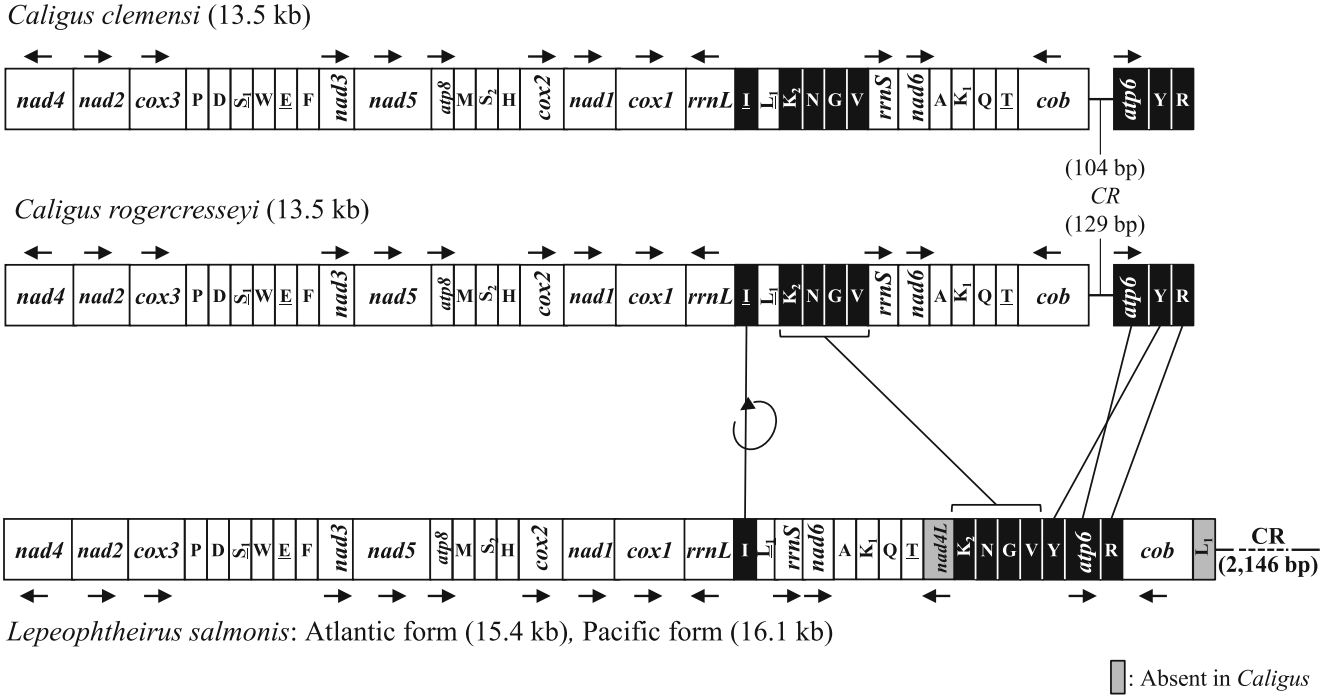


Fig. 2 Genomic organization of the *C. clemensi* (13,440 bp) and the *C. rogercressyi* (13,468 bp) mt genomes. The complete mt genomes of the Atlantic (15,445 bp) and Pacific (16,148 bp) *L. salmonis* were previously reported, and these mt genomes are identical in gene organization (Tjensvoll et al. 2005; Yazawa et al. 2008). *Boxes* represent mtDNA genes. tRNA genes are denoted by the single letter amino acid code, and an *underline* indicates tRNA genes located on negative strand. *rrnL* and *rrnS* refer to 16S and 12S rRNA; *cox1*, *cox2*, and *cox3* refer to cytochrome oxidase subunit I, II, and III; *cob* refers to cytochrome b; *nad1–6* and *nad4L* refer to NADH dehydrogenase subunits 1–6 and 4 L, *atp6* and *atp8* refer to ATP synthase subunits 6 and 8, respectively, and *CR* refers to control region. Transcription directions for the protein-coding and rRNA genes are shown by *arrowheads*

the two *Caligus* mt genomes (Supplemental Table 2). In addition, the orders of the genes in the two *Caligus* mt genomes are identical to each other, but different from the order in the *L. salmonis* mt genome (Fig. 2).

Sea Lice as Ectoparasite Model System

Since parasites by definition depend on a live host for growth and survival, in vitro culture system is typically very difficult to establish. Although procedures for experimental infections are established for some parasitic species, manipulation of the parasites may still be very difficult since removing them from the host is lethal for the parasite in general. Sea lice have life cycle features that make them promising as a model system. The life cycle features, consisting of both free-living larval developmental stages and pre-adults and adult stages that can move unrestricted on host surface, enable manipulation of these parasites. For *L. salmonis*, recent advances in larval production systems and infection procedures (see Hamre et al. 2009) have been crucial for the establishment of defined laboratory strains of the salmon louse with different properties (e.g., drug-resistant strains, inbred strains). Stable and predictable production conditions further enables specific breeding to create various types of hybrids (e.g., susceptible and drug-resistant family groups). The improvement of rearing facilities has been a crucial facilitator for establishment of RNAi in *L. salmonis* (Dalvin et al. 2009). Systemic RNAi is easily achieved in pre-adult or adult lice by injection of dsRNA in the animal. In addition, soaking free-living larval stages (e.g., copepodids) in dsRNA enables RNAi in copepodids (Campell et al. 2009). In addition, the genomes of both the Pacific and Atlantic variants of *L. salmonis* are currently being sequenced and together with the present cDNA resources this will open up for a new avenue in sea lice research. There is a wide diversity of arthropod parasites and good experimental parasite model systems are scarce, and we anticipate that experimental studies on salmon louse and other sea lice species will contribute to increase our knowledge about ectoparasites in general, particularly when more parasite genomes become available.

Conclusions

We sequenced over 150,000 ESTs from Pacific *L. salmonis* (49,672 new ESTs in addition to 14,994 previously reported ESTs), Atlantic *L. salmonis* (57,349 ESTs), *C. clemensi* (14,821 ESTs), *C. rogercresseyi* (32,135 ESTs), and *L. branchialis* (16,441 ESTs; Table 1). A relational database with an intuitive web interface was developed to process and display the large quantities of EST data, their assemblies and associated annotation information, as well as possible full-length gene information (Fig. 1). This database provides a novel resource for the study of sea louse biology, population genetics, and control strategies. This genomic resource represents the largest compilation of any copepod species and provides the material basis for the development of a 38K microarray that can be used in conjunction with our existing salmon 44K microarray to study host–parasite interactions at the molecular level.

The nuclear genes showed a high level of sequence divergence among the caligid copepods examined: *L. salmonis*, *C. clemensi*, *C. rogercresseyi*, and *L. branchialis* (Table 2). In addition, whole mt genome sequences of two *Caligus* species, *C. clemensi* (13,440 bp) and *C. rogercresseyi* (13,468 bp), were determined and compared. The *L. salmonis*, *C. clemensi*, and *C. rogercresseyi* mtDNA genes also exhibited extensive sequence divergence, ranging among these species from 66.7 to 68.8% nt and from 63.6% to 65.4% aa identities (Table 3). Both nuclear and mtDNA genes showed very high levels of sequence divergence between these ectoparasitic copepods which suggested that they have been in existence for 37–113 million years and that parasitic association with marine organisms is likely also quite ancient. However, while the order of the genes in the two *Caligus* mt genomes is the same, they are different from *L. salmonis* (Fig. 2). The large sequence divergence observed among these copepods may help to explain an extensive variety of morphology, life history, and host association in copepods.

Acknowledgments This project (GiLS—Genomics in Lice and Salmon) was supported by Genome BC, Microtek Intl., Marine Harvest, Mainstream Canada, Greig Seafoods, and the University of Victoria. We would like to thank Rob Holt (Head of Sequencing, Genome Sciences Centre, Vancouver, BC, Canada), Richard Moore (Sequencing Group Leader, Genome Sciences Centre), Sarah Munro, Mike Mayo, and Susan Wagner (Genome Sciences Centre) for plating and sequencing. We also would like to thank John Burka (University of P.E.I., Canada), Frank Nilsen, and Heidi Kongshaug (University of Bergen, Norway) for Atlantic forms of *L. salmonis*; the Salmones Maullin Company (Chile) for *C. rogercresseyi*; Brendan Conners (Salmon Coast Field Station, Simoom Sound, BC, Canada) for *C. clemensi*; and James Bron and Sarah Barker (University of Stirling, Scotland, UK) for *L. branchialis*.

References

Amalric F, Merkel C, Gelfand R, Attardi G (1978) Fractionation of mitochondrial RNA from HeLa cells by high-resolution electrophoresis under strongly denaturing conditions. J Mol Biol 118:1–25

Bairoch A, Apweiler R (1996) The SWISS-PROT protein sequence data bank and its new supplement TREMBL. Nucl Acids Res 24:21–25

Bauzà-Ribot MM, Jaume D, Juan C, Pons J (2009) The complete mitochondrial genome of the subterranean crustacean *Metacrangonyx longipes* (Amphipoda): A unique gene order and extremely short control region. Mitochondrial DNA 20:88–99

Beamish R, Wade J, Pennell W, Gordon E, Jones S, Neville C, Lange K, Sweeting R (2009) A large, natural infection of sea lice on juvenile Pacific salmon in the Gulf Islands area of British Columbia, Canada. Aquaculture 297:31–37

Beamish RJ, Neville CM, Sweeting RM, Ambers N (2005) Sea lice on adult Pacific salmon in the coastal waters of Central British Columbia, Canada. Fish Res 76:198–208

Berthier F, Renaud M, Alziari S, Durand R (1986) RNA mapping on *Drosophila* mitochondrial DNA: precursors and template strands. Nucl Acids Res 14:4519–4533

Boore JL (1999) Animal mitochondrial genomes. Nucl Acids Res 27:1767–1780

Boulding GE, deWaard RJ, Ang PK, Hebert DNP (2009) Population genetic structure of the salmon louse, *Lepeophtheirus salmonis* (Krøyer) on wild and farmed salmonids around the Pacific coast of Canada. Aquac Res 40:973–979

Boxaspen K (2006) A review of the biology and genetics of sea lice. ICES J Mar Sci 63:1304–1316

Boxshall AG, Bravo S (2000) On the identity of the common *Caligus* (Copepoda: Siphonostomatoida: Caligidae) from salmonid netpen systems in southern Chile. Contrib Zool 69:137–146

Bricknell IR, Bron JE, Bowden TJ (2006) Diseases of gadoid fish in cultivation: a review. ICES J Mar Sci 63:253–266

Burton RS, Byrne RJ, Rawson PD (2007) Three divergent mitochondrial genomes from California populations of the copepod *Tigriopus californicus*. Gene 403:53–59

Campell EM, Pert CC, Bowman AS (2009) RNA-interference methods for gene-knockdown in the sea louse, *Lepeophtheirus salmonis*: studies on a putative prostaglandin E synthase. Parasitology 136:867–874

Carvajal J, González L, George-Nascimento M (1998) Native sea lice (Copepoda: Caligidae) infestation of salmonids reared in netpen systems in southern Chile. Aquaculture 166:241–246

Costello JM (2009) The global economic cost of sea lice to the salmonid farming industry. J Fish Dis 32:115–118

Costello MJ (2006) Ecology of sea lice parasitic on farmed and wild fish. Trends Parasitol 22:475–483

Cunningham CW, Blackstone NW, Buss LW (1992) Evolution of king crabs from hermit crab ancestors. Nature 355:539–542

Dalvin S, Frost P, Biering E, Hamre L, Eichner C, Krossøy B, Nilsen F (2009) Functional characterisation of the maternal yolk-associated protein (LsYAP) utilising systemic RNA interference in the salmon louse (*Lepeophtheirus salmonis*) (Crustacea: Copepoda). Int J Parasitol l39:1407–1413

Dawson LHJ, Pike AW, Houlihan DF, McVicar AH (1997) Comparison of the susceptibility of sea trout (*Salmo trutta* L.) and Atlantic salmon (*Salmo salar* L.) to sea lice (*Lepeophtheirus salmonis* (Kroyer, 1837)) infections. ICES J Mar Sci 54:1129–1139

Dowton M, Austin AD (1999) Evolutionary dynamics of a mitochondrial rearrangement "hot spot" in the Hymenoptera. Mol Biol Evol 16:298–309

Fast MD, Ross NW, Mustafa A, Sims DE, Johnson SC, Conboy GA, Speare DJ, Johnson G, Burka JF (2002) Susceptibility of rainbow trout *Oncorhynchus mykiss*, Atlantic salmon *Salmo salar* and coho salmon *Oncorhynchus kisutch* to experimental infection with sea lice *Lepeophtheirus salmonis*. Dis Aquat Organ 521:57–68

Gee JM (1987) Impact of epibenthic predation on estuarine intertidal harpacticoid copepod populations. Mar Biol 96:497–510

González L, Carvajal J, George-Nascimento M (2000) Differential infectivity of *Caligus flexispina* (Copepoda, Caligidae) in three farmed salmonids in Chile. Aquaculture 183:13–23

González L, Carvajal J (2003) Life cycle of *Caligus rogercresseyi* (Copepoda: Caligidae) parasite of Chilean reared salmonids. Aquaculture 220:101–117

Hamre L, Glover K, Nilsen F (2009) Establishment and characterisation of salmon louse (*Lepeophtheirus salmonis*) laboratory strains. Parasitol Int 58:451–460

Hebert PD, Ratnasingham S, deWaard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc Biol Sci 270:S96–S99

Heuch PA, Bjørn PA, Finstad B, Holst JC, Asplin L, Nilsen F (2005) A review of the 'Norwegian National Action Plan Against Salmon Lice on Salmonids': the effect on wild salmonids. Aquaculture 246:79–92

Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. Genome Res 9:868–877

Humes AG (1994) How many copepods? Hydrobiologia 292–293:1–7

Johnson SC, Albright LJ (1992) Comparative susceptibility and histopathology of the response of naive Atlantic. Chinook and coho salmon to experimental infection with Lepeophtheirus salmonis Copepoda:Caligidae. Dis Aquat Organ 14:179–193

Johnson SC, Fast MD (2004) Interactions between sea lice and their hosts. Symp Soc Exp Biol 55:131–159

Johnson SC, Treasurer JW, Bravo S, Nagasawa K, Kabata Z (2004) A review of the impact of parasitic copepods on marine aquaculture. Zool Stud 432:229–243

Jones SR, Prosperi-Porta G, Kim E, Callow P, Hargreaves NB (2006) The occurrence of *Lepeophtheirus salmonis* and *Caligus clemensi* (Copepoda: Caligidae) on three-spine stickleback *Gasterosteus aculeatus* in coastal British Columbia. J Parasitol 92:473–480

Kabata Z (1979) Parasitic Copepoda of British fishes. The Ray Society, London

Ki JS, Park HG, Lee JS (2009) The complete mitochondrial genome of the cyclopoid copepod Paracyclopina nana: a highly divergent genome with novel gene order and atypical gene numbers. Gene 435:13–22

Kilpert F, Podsiadlowski L (2006) The complete mitochondrial genome of the common sea slater, *Ligia oceanica* (Crustacea, Isopoda) bears a novel gene order and unusual control region features. BMC Genomics 7:241

Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Koop BF, von Schalburg KR, Leong J, Walker N, Lieph R, Cooper GA, Robb A, Beetz-Sargent M, Holt RA, Moore R, Brahmbhatt S, Rosner J, Rexroad CIE, McGowan CR, Davidson WS (2008) A salmonid EST genomic study: genes, duplications, phylogeny and microarrays. BMC Genomics 9:545

Krkošek M, Ford JS, Morton A, Lele S, Myers RA, Lewis MA (2007a) Declining wild salmon populations in relation to parasites from farm salmon. Science 318:1772–1775

Krkošek M, Gottesfeld A, Proctor B, Rolston D, Carr-Harris C, Lewis MA (2007b) Effects of host migration, diversity and aquaculture on sea lice threats to Pacific salmon populations. Proc R Soc Lond B Biol Sci 274:3141–3149

Lavrov DV, Brown WM, Boore JL (2004) Phylogenetic position of the Pentastomida and (pan)crustacean relationships. Proc R Soc Lond B Biol Sci 271:537–544

Leong JS, Jantzen SG, von Schalburg KR, Cooper GA, Messmer AM, Liao NY, Munro S, Moore R, Holt RA, Jones SJM, Davidson WS, Koop BF (2010) *Salmo salar* and *Esox lucius* full-length cDNA clones reveal changes in evolutionary pressures on a posttetraploidization genome. BMC Genomics 11:279

Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucl Acids Res 25:955–964

Macey JR, Larson A, Ananjeva NB, Fang Z, Papenfuss TJ (1997) Two novel gene orders and the role of light-strand replication in

rearrangement of the vertebrate mitochondrial genome. Mol Biol Evol 14:91–104

Machida RJ, Miya MU, Nishida M, Nishida S (2002) Complete mitochondrial DNA sequence of *Tigriopus japonicus* (Crustacea: Copepoda). Mar Biotechnol 4:406–417

Machida RJ, Miya MU, Nishida M, Nishida S (2004) Large-scale gene rearrangements in the mitochondrial genomes of two calanoid copepods *Eucalanus bungii* and *Neocalanus cristatus* (Crustacea), with notes on new versatile primers for the srRNA and COI genes. Gene 332:71–78

Ohman MD, Hirche HJ (2001) Density-dependent mortality in an oceanic copepod population. Nature 412:638–641

Ohno S (1970) Evolution by gene duplication. Springer, New York

Øines Ø, Schram T (2008) Intra- or inter-specific difference in genotypes of *Caligus elongatus* Nordmann 1832? Acta Parasitol 53:93–105

Pike AW, Wadsworth SL (1999) Sealice on salmonids: their biology and control. Adv Parasitol 44:233–337

Piasecki W, MacKinnon BW (1995) Life cycle of a sea louse, *Caligus elongatus* von Nordmann, 1832 (Copepoda, Siphonostomatoida, Caligidae). Can J Zool 73:74–82

Schubart CD, Diesel R, Hedges SB (1998) Rapid evolution to terrestrial life in Jamaican crabs. Nature 393:363–365

Smith JL, Wootten R, Sommerville C (2007) The pathology of the early stages of the crustacean parasite, *Lernaeocera branchialis* (L.), on Atlantic cod, *Gadus morhua* L. J Fish Dis 30:1–11

Saksida S, Constantine J, Karreman GA, Donald A (2007) Evaluation of sea lice abundance levels on farmed Atlantic salmon (*Salmo salar* L.) located in the Broughton Archipelago of British Columbia from 2003 to 2005. Aquac Res 38:219–231

Sturmbauer C, Levinton JS, Christy J (1996) Molecular phylogeny analysis of fiddler crabs: test of the hypothesis of increasing behavioral complexity in evolution. Proc Natl Acad Sci USA 93:10855–10857

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Tjensvoll K, Glover KA, Nylund A (2006) Sequence variation in four mitochondrial genes of the salmon louse *Lepeophtheirus salmonis*. Dis Aquat Organ 68:251–259

Tjensvoll K, Hodneland K, Nilsen F, Nylund A (2005) Genetic characterization of the mitochondrial DNA from *Lepeophtheirus salmonis* (Crustacea; Copepoda). A new gene organization revealed. Gene 353:218–230

Todd CD, Walker AM, Ritchie MG, Graves JA, Walker AF (2004) Population genetic differentiation of sea lice (*Lepeophtheirus salmonis*) parasitic on Atlantic and Pacific salmonids: analyses of microsatellite DNA variation among wild and farmed hosts. Can J Fish Aquat Sci 61:1176–1190

Todd CD, Whyte BDM, MacLean JC, Walker AM (2006) Ectoparasitic sea lice (*Lepeophtheirus salmonis* and *Caligus elongatus*) infestations of wild, adult, one sea-winter Atlantic salmon *Salmo salar* returning to Scotland. Mar Ecol Prog Ser 328:183–193

Tully O, Nolan DT (2002) A review of the population biology and host–parasite interactions of the sea louse *Lepeophtheirus salmonis* (Copepoda: Caligidae). Parasitology 124:S165–S182

Walter TC, Boxshall GA (2008) World of Copepods database. Available online at http://www.marinespecies.org/copepoda: Consulted on 2010-01-06

Yazawa R, Yasuike M, Leong J, von Schalburg KR, Cooper GA, Beetz-Sargent M, Robb A, Davidson WS, Jones SRM, Koop BF (2008) EST and mitochondrial DNA sequences support a distinct Pacific form of salmon louse, *Lepeophtheirus salmonis*. Mar Biotechnol 10:741–749