Genetic diversity of *Clonorchis sinensis* (Trematoda: Opisthoriidiae) in the Russian southern Far East based on mtDNA *cox1* sequence variation

Yulia V. Tatonova, Galina N. Chelomina and Vladimir V. Besprozvannykh

Institute of Biology and Soil Science, Far Eastern Branch, Russian Academy of Sciences, Vladivostok, Russia

**Abstract:** We examined the phylogeography and the variation of the mitochondrial DNA (mtDNA) cytochrome c oxidase subunit 1 gene (*cox1*) of the Chinese liver fluke *Clonorchis sinensis* (Cobbold, 1875) in two geographic localities in the Russian southern Far East and compared them with those from different geographical regions (China, Korea, Japan and Vietnam). The Russian samples differed from those of the other regions in haplotype frequencies, haplotype and nucleotide diversities, and AT/GC ratios. Only 4 of the 18 haplotypes were common to Russian and Chinese samples, and two haplotypes were common to Russia and other regions. The intraspecific genetic distances ranged from 0 to 1.58% for the entire dataset studied and from 0 to 1.25% among the samples from Russia. Phylogenetic trees revealed no significant genealogical clades of samples corresponding to sampling localities and no strong isolation by distance was estimated with Mantel test. Neutrality test analysis suggested a relatively recent population expansion for *C. sinensis*, whereas goodness-of-fit tests indicated deviation from the strict model of uniform expansion. Therefore, the sequences of the mtDNA *cox1* gene provide useful genetic markers for evaluating intraspecific diversity and generating phylogeographic reconstructions for this fish-borne trematode.

**Keywords:** liver fluke, *Clonorchis sinensis*, mtDNA, *cox1* gene, genetic diversity, phylogeography, Asia

The Chinese liver fluke *Clonorchis sinensis* (Cobbold, 1875) (Digenea: Opisthoriidiae) is endemic to Southeast Asia where it infects both humans and animals, and causes an important food-borne zoonosis, clonorchiasis (Lun et al. 2005). The infection is dangerous because of the negative impact of the parasite on the liver and bile ducts and the general intoxication of the host that results in a number of human diseases, including the life-threatening disease cholangiocarcinoma (Choi et al. 2004). Therefore, this species has recently been classified as a member of the group 1 biological agents that are carcinogenic to humans (Bouvard et al. 2009).

In the south of the Russian Far East, several reservoirs inhabited by the first intermediate hosts of the parasite (the freshwater snail Parafossarulus mantshouricus Bouriguignat) have been found. In some localities, the incidence of clonorchiasis has increased in recent years (own observations). Man-made introduction of intermediate hosts, the import of crude fish, tourism, immigration from Southeast Asia, and increasing interest in exotic eastern foods that may contain uncooked fish are factors that influence the epidemiological situation in Russia. Therefore, understanding the genetic structure and the geographical patterns of liver fluke populations in the Russian Far East is of a great scientific, medical and veterinary importance. Currently, there is little information on the levels of genetic variation within and among populations of *C. sinensis* throughout its geographical distribution based on mitochondrial DNA (mtDNA) genes (Park and Yong 2001, Lee and Huh 2004, Park 2007, Liu et al. 2012). mtDNA markers are typically used in population and ecological genetic studies on parasites as well as in phylogenetic studies (Saijuntha et al. 2008, Zhao et al. 2009, Ai et al. 2010, Martínez-Salazar and León-Régagnon 2010, Liu et al. 2012, Scholz et al. 2011).

Phylogeographic analyses based on genetic diversity data of contemporary populations elucidate the processes that determine the geographical distribution of genealogical lineages and allow the reconstruction of the history of a species (Avise 2000). Although a number of phylogeographic studies on animals have been reported, parasite species have been inadequately investigated, with limited information on helminths (Iwagami et al. 2000, Wickström et al. 2003, Nieberding et al. 2004, Attwood et al. 2008, Koehler et al. 2009, Ichikawa et al. 2011).

The aim of the present study is to provide results of the first population genetic analysis of the liver fluke *C. sinensis* in the Russian Far East and to present new data on the phylogeography of this fish-borne trematode in southeastern Asia.
Materials and Methods

Metacercariae of _Clonorchis sinensis_ were obtained by dissection of freshwater cyprinid fish, _Pseudorasbora parva_ (Temminck et Schlegel), collected from two localities in the southern region of the Russian Far East (Fig. 1). Metacercariae were fed to rats, which were examined approximately one month after infection, and the adult flukes were recovered.

Liver flukes were washed several times in physiological solution immediately after dissection and stored in 96% ethanol. Genomic DNA was extracted from 26 individual worms using the HotSHOT technique (Truett et al. 2000). Partial fragments of the mitochondrial gene of the cytochrome c oxidase subunit 1 (cox1) were amplified by polymerase chain reaction (PCR) using the following universal primers: forward, 5\'-GGG-CAT-CCT-GAG-GTT-TAT-G-3\' and reverse, 5\'-AAC-AAA-TCA-TGA-TGC-AAA-AGG-3\' (Katokhin et al. 2008). PCR was performed in a total volume of 20 μl containing 0.25 mM of each primer pair, 1 μl of DNA in water, 1× Taq buffer, 1.25 mM dNTP, 1.5 mM magnesium and 1.5 units of Taq polymerase (Medigen, Novosibirsk, Russia). Amplification of the cox1 gene was performed on a GeneAmp 9700 (Applied Biosystems) using the following cycling conditions: a 1-min initial denaturation step at 95 °C; 35 cycles of 30 sec at 94 °C, 1 min at 55 °C, and 2 min at 72 °C; and a 7-min extension at 72 °C. Negative and positive controls with both primers were included.

The PCR products of the cox1 genes from 26 specimens of _C. sinensis_ were purified by ethanol precipitation and sequenced directly on an ABI 3130 Genetic Analyzer (at the Institute of Biology and Soil Sciences FEB RAS) using the ABI BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, USA).

The 26 mtDNA cox1 sequences that were obtained were deposited in the EMBL GenBank database under accession numbers JX040546–JX040571. The relevant sequences of this species from other regions (China, Korea, Vietnam and Japan) were downloaded from GenBank. Overall, we analysed 52 324-bp sequences of the mtDNA cox1 gene (Table 1, Fig. 1). The sequences were assembled manually and aligned using Clustal X (Thompson et al. 1997).

Molecular diversity was estimated using DnaSP version 5.10 (Librado and Rozas 2009) and ARLEQUIN version 3.11 (Excoffier et al. 2006). The number of haplotypes (H), polymorphic sites (S), both haplotype (k) and nucleotide (\(\pi\)) diversity (Nei 1987), and the average numbers of pairwise differences (\(\theta\)) (Tajima 1983) were assessed. Distances among haplotype sequences were calculated in MEGA version 5 (Tamura et al. 2011).

The substitution model and the gamma distribution shape parameter for the rate of heterogeneity among sites were determined with Modeltest 3.07 (Posada and Crandall 1998) based on the Hierarchical Likelihood Ratio Tests (hLRTs). The HKY model (Hasegawa et al. 1985) of evolution with the gamma shape parameter (HKY+G) was selected for phylogenetic analysis.

AMOVA (Analysis of Molecular Variance) statistics were calculated in ARLEQUIN (Excoffier et al. 2006) to determine the amount of genetic variability in geographical populations using \(\phi\) genetic distances between different haplotypes. The null hypothesis of population panmixia was tested using an exact test of haplotype differentiation among populations (Raymond and Rousset 1995). Probabilities were estimated by permutation analyses using 10,000 randomly permuted populations (\(\tau\)) with different haplotypes (\(\epsilon\)).

Genetic relationships among haplotypes were reconstructed using Bayesian inference analysis (BI) using MrBayes 3.1.1. (Ronquist and Huelsenbeck 2003), maximum likelihood (ML), and neighbour-joining (NJ) (Saitou and Nei 1987) methods in PAUP 4.0 Beta 10 program (Swofford 2002) with _Opisthorchis_
Table 2. Summary of genetic diversity descriptive statistics for cox1 sequences for Clonorchis sinensis studied.

<table>
<thead>
<tr>
<th>Region/Parameter</th>
<th>N</th>
<th>H</th>
<th>S</th>
<th>h</th>
<th>π</th>
<th>k</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Russia</td>
<td>28</td>
<td>10</td>
<td>9</td>
<td>0.854±0.040</td>
<td>0.004679±0.003227</td>
<td>1.515873±0.939111</td>
<td>0−0.0125 (0.0047)</td>
</tr>
<tr>
<td>China</td>
<td>19</td>
<td>11</td>
<td>13</td>
<td>0.830±0.085</td>
<td>0.005162±0.003540</td>
<td>1.672515±1.026370</td>
<td>0−0.0158 (0.0052)</td>
</tr>
<tr>
<td>Japan, Korea, Vietnam</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>0.900±0.161</td>
<td>0.00673±0.004847</td>
<td>2.00006±1.343457</td>
<td>0−0.0125 (0.0063)</td>
</tr>
<tr>
<td>Total samples</td>
<td>52</td>
<td>18</td>
<td>20</td>
<td>0.868±0.034</td>
<td>0.005398±0.003536</td>
<td>1.748689±1.032657</td>
<td>0−0.0158 (0.0054)</td>
</tr>
</tbody>
</table>

N – sample size; H – number of haplotypes; S – number of polymorphic (segregating) sites; h – haplotype diversity (±SD); π – nucleotide diversity (±SD); k – mean pair-wise difference (±SD); D – genetic distances (±SD).

Table 3. Parameters of genetic differentiation between samples of Clonorchis sinensis from different regions.

<table>
<thead>
<tr>
<th>Pairwise comparison/ Parameter</th>
<th>Fst (P)</th>
<th>D (± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Russia vs China</td>
<td>0.13726 (P &lt; 0.0001)</td>
<td>0.00578 (0.00196)</td>
</tr>
<tr>
<td>Russia vs Japan, Korea and Vietnam</td>
<td>0.16174 (P &lt; 0.05)</td>
<td>0.00634 (0.00222)</td>
</tr>
<tr>
<td>China vs Japan, Korea and Vietnam</td>
<td>0.02346 (P &lt; 0.5)</td>
<td>0.00581 (0.00169)</td>
</tr>
</tbody>
</table>

Fst – coefficient of gene fixations; P – probability; D – genetic distances; SD – standard deviation.

RESULTS

Genetic diversity

The 324-bp sequences of the mtDNA cox1 gene (positions 739-1062 bp of the complete gene sequence) were determined for 26 specimens of Clonorchis sinensis from two isolated localities of the Russian Far East. In combination with available GenBank data, the total dataset included 52 nucleotide sequences from 11 geographical localities in Russia, China, Korea, Vietnam and Japan.

Table 4. Results of neutrality tests for different samples of Clonorchis sinensis.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fs</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td></td>
</tr>
</tbody>
</table>

The sequence divergence among haplotypes ranged from 0% to 1.58%, with an average of 0.54% for the species (Table 2). The number of polymorphic sites was higher in the Chinese samples compared with those in the Russian samples (13 vs 9, respectively). The genetic distances between the most divergent haplotypes and the average genetic distance among Russian samples were lower than those of Chinese samples. The highest genetic distances were estimated among haplotypes from Korea, Vietnam and Japan (Table 2).

Genetic diversity indices for samples from Russia and China and for the total dataset are summarised in Table 3. The overall haplotype diversity (h) of 0.868 ± 0.034 and a nucleotide diversity (π) of 0.005398 ± 0.003536 indicate high levels of haplotype diversity and low levels of nu-
cleotide diversity. Additionally, all parameters of genetic diversity for the Russian samples were somewhat lower than those for the Chinese samples despite the greater sample size of the former set. Both the genetic distances and AMOVA analysis, including the haplotype frequencies and nucleotide divergence between the haplotypes of the groups under comparison, did not reveal strong differences between regional groups (Table 2). The exact test of population differentiation also showed no significant differences among the populations (P = 1.00). For this reason, the null hypothesis (i.e. that the C. sinensis population is panmictic) could not be rejected. Additionally, genetic isolation by distance was not found among populations (Mantel test P-value > 0.05 for all models of IBD analysis), i.e. geographic distances do not serve as a barrier to gene flow.

The patterns of haplotype distribution are presented in Fig. 2. Among 18 haplotypes identified, more than half were unique and represented by a single specimen. The most common haplotype (accounting for 30% of all samples) was found in all regions except Korea. The patterns of haplotype frequencies in China and Russia were estimated to be significantly different, with only four common haplotypes. A total 10 of 18 haplotypes were found among the Russian samples and only three of these were singletons for the total dataset. Most of the sequences (18 of 26, i.e. 67%) represented three haplotypes, one specific for Russia, the second in common with the northern China sample and the third in common with the samples from southern China, Vietnam and Japan. Twelve of 18 haplotypes were revealed among Chinese samples; most of these (8 of 22, i.e. 36%) represented one widespread haplotype and six haplotypes were singletons for the total dataset. In the remaining samples comprising Korea, Vietnam and Japan, a single haplotype (from Vietnam) was unique for the species.

**Intraspecific phylogeny and phylogeographic structure of mtDNA haplotypes**

Modeltest revealed that the Hasegawa-Kishino-Yano (Hasegawa et al. 1985) substitution model fits with invariable sites and the gamma shape parameter (HKY+G, G = 0.279) provides the best fit for the cox1 data. From these results, the Bayesian, ML and NJ trees were reconstructed using HKY genetic distances to estimate the phylogenetic relationships among individuals (Fig. 3). The topologies of all trees were very similar, with small differences in bootstrap values. The species within the outgroup were well differentiated, and all sequences of C. sinensis formed a monophyletic group, which was sister to Opisthorchis felineus. The intraspecific topology was unresolved, with no significant genealogical branches or clusters of haplotypes corresponding to sampling localities. Only a few small groups of individuals from different geographical localities were statistically supported.

To further investigate the shallow phylogeny of the species and to identify ancestral haplotypes, we used the minimum spanning tree (MST) reconstruction, an approach explicitly developed for intraspecific data. The MST (Fig. 4) showed a large single star-like structure with two small substructures, one of which comprised the Russian and northern China haplotypes, and another included Russian and southern China haplotypes. These results suggest a relatively recent population expansion, a single ancestral haplotype, and no phylogeographic structuring, which implies a high level of gene flow among local populations and geographical groups.
Fig. 3. Phylogenetic relationships between isolates of *Clonorchis sinensis* inferred from Bayesian inference (Bi), maximum likelihood (ML) and neighbour-joining (NJ) analyses for cytochrome c oxidase subunit 1 with other trematodes as outgroup. Bootstrap support of > 50% is shown at nodes. *Abbreviations:* 1.1–1.17 – isolates sampled at Russia, Kronshtadtka; 2.1–2.17 – isolates sampled at Russia, Kondratenovka; JapOk – Japan, Okayama; ChGng – China, Guangdong; ChHun – China, Hunan; ChHei – China, Heilongjiang; ChGxi – China, Guangxi; VietNA – Vietnam, Nghe An; VietTH – Vietnam, Thanh Hoa; RusKha – Russia, Khabarovsk; Kor – Korea; KorKim – Korea, Kimhae.

Fig. 4. Minimum spanning tree for cytochrome c oxidase subunit 1 haplotypes of *Clonorchis sinensis*. The length of the branches corresponds to genetic distances; n – number of haplotypes.
Mismatch distributions and demographic histories

The demographic history of *C. sinensis* was investigated using mismatch distributions. The unimodal mismatch distribution (Fig. 5) matched the expected distribution under the sudden expansion model, suggesting population expansion in the past. Both Tajima’s *D* (-1.917, *p* = 0.01) and Fu’s *F* (*s* -27.477, *p* = 0.00) indices were negative and significant, which is typical for populations that experienced a recent demographic expansion. However, the goodness-of-fit tests (Rogers and Harpending 1992) with sum-of-squared deviation (0.003, *P* = 0.31) and Raggedness index (0.0663, *P* = 0.0160) indicate a deviation from the strict model of uniform expansion.

DISCUSSION

The mitochondrial genomes of trematodes, with the exception of *Paragonimus westermani*, are AT-rich and the coding chain is particularly rich in thymine, but deficient in cytosine (Shekhovtsov et al. 2010), which agrees well with our data. The variation in the A+T content of the *cox1* gene among the geographical groups revealed in this study is also in agreement with the data on complete mitochondrial genomes of *C. sinensis* from China and Korea (Cai et al. 2012).

An average intraspecific divergence of 0.03–1.3% at the mtDNA level was observed in eight trematode species (Vilas et al. 2005). The intraspecific sequence variations within *C. sinensis* samples from China have previously been estimated as 0–1.6% for *cox1* (Li et al. 2011). The levels of genetic differentiation observed in the present study confirmed these data; we also have shown that the levels of *cox1* differentiation among the Russian samples is lower than among the Chinese samples.

Paleontological data (Starobogatov 1970) indicate that the intermediate hosts of *C. sinensis*, species of the genus *Parafossarulus* Annandale inhabited China and the Russian Far East during the Pliocene. The climate in the south of the Russian Far East during the Pliocene was warm temperate with the prolonged dry seasons but cold at the end of this period (Korotky et al. 1996). In the middle Pleistocene, there were significant changes in river systems, possibly due in part to volcanic activity (Korotky et al. 2011). These factors could have formed an ecological barrier for freshwater snails, the intermediate hosts of *C. sinensis*, which may have resulted in haplotype differentiation of the liver fluke.

Comparisons of nucleotide and gene diversity parameters can provide insights into the demographic history of populations and permit speculation concerning past demographic events (Avise 2000). High genetic and low nucleotide diversities, which were found in our study, can result from the long-term isolation of a relatively small population (Avise 2000, Painter et al. 2007). The lower genetic diversity in northern (Russian) samples might be the consequence of a founder effect, which would imply an ancestral population for the species in the more southern territories. The phylogenetic investigation of postglacial expansion of free-living species has shown that populations usually have less genetic variation in previously glaciated areas, whereas ancestral populations in regions south of the species’ territory maintain higher genetic diversity than those to the north (Hewitt 1999). The trend of northward decline of genetic diversity has also been found in parasite species during their dispersal or expansion (Li et al. 2011).

Based on the mtDNA *cox1* sequence data, we found a single haplotype lineage and detected population expansion of the Chinese liver fluke inhabiting a vast territory of South East Asia. The phylogeographic continuity that was revealed correlated with reciprocal monophyly, and the presence of a widespread haplotype can be evidence of an absence of refugial divisions in species during the glacial ages of the Pleistocene (Hewitt 1999).

In theory, populations that have undergone a sudden expansion or exponential growth are characterised by a smooth, unimodal distribution (Rogers and Harpending 1992), and a peak in the range of 0–1 (the L-shaped curve) is an indicator of a recent bottleneck event (Excoffier and Schneider 1999). The star-like phylogeny is more clearly explained by a sudden population expansion after a bottleneck (due to the small number of founders or a reduction in the historically effective population size) when molecular markers can be homologous over a vast territory, and both nucleotide and haplotype diversities can be low even in structured populations (Slatkin and Hudson 1991, Walker et al. 1998).

Another explanation is that a single haplotype can be preferred by natural selection replacing other haplotypes (Rich et al. 1998). Populations are usually subject to spatial expansion if their area was initially strongly limited and then increased. The extensive spatial expansion often gives the same signal in the distribution of genetic differ-

---

![Fig. 5. Mismatch distribution of cytochrome c oxidase subunit 1 haplotypes for the total samples of *Clonorchis sinensis*. The frequency of pairwise differences is on the axis y. Abbreviations: Exp – expected frequency; Obs – observed frequency.](image-url)
ences as a pure demographic expansion in the panmicotic population, but only when the adjacent populations have a large (≥ 50) number of migrants (Calvo et al. 2009).

Therefore, the sequences of the mtDNA cox1 gene provide useful genetic markers for evaluating intraspecific diversity and generating phylogeographic reconstructions for the zoonotic trematode *C. sinensis*. However, to a certain extent, the phylogeographic data can be regarded as preliminary. Extending the geographic range and increasing the DNA regions analysed will allow a more accurate reconstruction of the history of this parasite in South East Asia.

**Acknowledgements.** This research was partially supported by grant of Far Eastern Branch of RAS (project No. 11-III-D-06-014).

**REFERENCES**


Ichikawa M., Bawn S., Maw N.N., Htun L.L., Thein M., Gyi A., Sunn K., Takakura K., Itagaki T. 2011: Characterization of *Fasciola* spp. in Myanmar on the basis of spermatogenesis sta-


Received 20 June 2012

Accepted 19 November 2012