Molecular characteristics and genetic diversity of *Fasciola hepatica* from sheep in Xinjiang, China

Wang Xifeng¹, Zhang Kai¹, Zhang Guowu¹, Li Zhiyuan¹, Shang Yunxia¹, Ning Chengcheng¹, Ji Chunhui¹, Qiao Jun¹, Meng Qingling¹, Cai Xuepeng²

¹College of Animal Science & Technology, Shihezi University, Shihezi, Xinjiang, 832003, PR China
²State Key Laboratory of Veterinary Etiological Biology, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou, Gansu, 730046, PR China

xjmqlq@sina.com

Received: September 14, 2021  Accepted: March 18, 2022

Abstract

**Introduction:** *Fasciola hepatica* is a trematode infecting ruminants worldwide and occasionally affecting other animal species, including humans. It causes significant economic losses. Geographic distribution and patterns of infection must be considered before control and management measures are developed for this parasite. DNA molecular markers are useful for the identification of flukes and elucidation of their genetic evolution. Therefore, the population structure of *F. hepatica* was studied using this method in sheep in Xinjiang, China. **Material and Methods:** The molecular characteristics, genetic relationships within the population and dispersal patterns of *F. hepatica* isolates were analysed based on the *cox1* and *nad1* genes. The population structure of *F. hepatica* from three regions of Xinjiang was explored and a neutrality test was conducted. **Results:** The *cox1* and *nad1* genes have 21 and 42 variable sites, respectively, which can be classified into 34 and 33 haplotypes. Median-joining network and phylogenetic tree analyses showed that there was no significant variation in *F. hepatica* isolates between the three geographical regions. Analysis of variance revealed that the genetic variation of *F. hepatica* was mainly present within the populations. The neutrality test indicated that the populations were relatively stable but the Hami population may have undergone short-term expansion. **Conclusion:** This study revealed for the first time the molecular characteristics, genetic diversity and dispersal patterns of *F. hepatica* isolates from sheep in Xinjiang, thus providing new insights into the genetic variation and haplotype diversity of *F. hepatica* from indigenous sheep.

**Keywords:** *F. hepatica*, *cox1*, *nad1*, genetic diversity, population structure.

Introduction

Fascioliasis is a parasitic disease caused by the *Fasciola hepatica* and *Fasciola gigantica* trematodes and their hybrids. *Fasciola hepatica* is widespread in temperate regions, while *F. gigantica* is mainly distributed in tropical areas (20). So far, more than 70 countries have noted *Fasciola spp* infection. The intermediate host of the *Fasciola* flukes is the snail, and the main terminal hosts are sheep, cattle and goats; wildlife and humans may be involved in its transmission. Livestock are often infected via accidental ingestion of aquatic plants that carry metacercariae (30). The global prevalence of infection is estimated to be more than 250 million sheep and 350 million cattle, causing more than 3 billion dollars in annual productivity losses in the livestock industry (3). Infection with *F. hepatica* can cause liver damage, slow growth, and an apparent reduction in meat and milk production, which is seriously jeopardizing the healthy and stable development of the livestock industry. Although *F. hepatica* has been prevalent worldwide for many years, there is still insufficient detailed information on the genetic structure of populations in various regions. Therefore, understanding the molecular characteristics, population structure and diversity of genetic variation of *F. hepatica* is essential to control the spread of the disease.
Currently, molecular genetic techniques based on DNA analysis are used for the identification of lamellipod species and elucidation of the genetic structure of populations. Marker molecules, such as the internal ribosomal transcribed spacers (ITS1 and ITS2), the phosphoenolpyruvate carboxykinase (pepck) gene, the DNA polymerase δ (polδ) gene, and the 28S rRNA and 18S rRNA genes, have been used to identify *F. hepatica*, *F. gigantica* and their ‘hybridisation’ species (14, 15, 26). In addition, mitochondrial DNA with high mutation rate, simple molecular structure and rapid evolution (7, 20), such as the NADH dehydrogenase subunit 1 (nad1) and the cytochrome c oxidase 1 (cox1) genes, have been widely used to analyse the genetic structure and dispersal pattern of *F. hepatica* populations (24, 25, 27).

Fascioliasis is widespread throughout China and the prevalence in animals is quite high (32). Approximately 28.5% of sheep were found to be infected with *F. hepatica* between 1999 and 2003 in Heilongjiang Province (29). The prevalence of this parasite in yak was 21.8–39.1% in Gansu province (34). In Guangxi province in southern China, the infection rate of *F. gigantica* in buffaloes is 87.35% (33). However, there are few reports of the Fasciola fluke infection in animals in Xinjiang. Xinjiang province is one of the important sheep breeding bases in northwest China and features rich natural pasture resources. At present, there are 42 million sheep in Xinjiang, most of which have been grazed mainly on pasture. In recent years, the high prevalence of fascioliasis has caused huge economic loss to the livestock industry and posed a great threat to the development of the sector of that industry which rears sheep. However, little is known about the molecular characteristics, population structure or genetic variation of *F. hepatica* isolates in Xinjiang, and their dispersal patterns are not yet clear. Therefore, the main purpose of this study was to investigate the level of genetic diversity and population structure of *F. hepatica* endemic strains in different geographical regions of Xinjiang to reveal these strains’ genetic dynamics and dispersal patterns in Xinjiang sheep. It is hoped that the knowledge obtained will provide a scientific basis for the development of reasonable management measures and control strategies to prevent the transmission of fascioliasis.

Material and Methods

**Natural overview of sampling sites.** Urumqi is located at the northern foot of the middle part of the Tianshan Mountains, surrounded by mountains on three sides, with extensive water resources such as the Urumqi River and over 600,000 hectares of natural pastures. Kuqa is located in the middle of the southern foot of the Tianshan Mountains and north of the Tarim Basin, an area abundant in waters such as the Tarim River, with vast forests and summer alpine pastures. Hami is located in the eastern part of Xinjiang, with a typical temperate continental climate and more virgin forests and natural pastures in the mountainous areas (Supplementary Fig. S1).

**Collection of trematodes.** A total of 144 worms were collected from the livers of sheep in Urumqi (Fh1–45), Kuqa (Fh46–54) and Hami (Fh55–144) regions of Xinjiang from 2016 to 2020. These flukes were washed in phosphate buffered saline (0.01M, pH 7.2–7.4) and stored in 70% alcohol at –80°C in a refrigerator for subsequent use.

**Primer design and synthesis.** The ITS1 gene forward F1: 5′-TTGCGCTGATTACGTCCTCTG-3′ and reverse R1: 5′-TGGCTGCCTTTCTTCGAC-3′ primers were used to identify the species of the collected worms, and the cox1 gene forward F2: 5′-TTGTGTTTGTGSGCATCT-3′ and reverse R2: 5′-AGGCCACCAAAAAGA-3′ and nad1 gene forward F3: 5′-AAGGATGTCGTITTTGCGG-3′ and reverse R3: 5′-GGATACGGTTACATCC-3′ primers were used to identify the species of the collected worms. The cox1 and nad1 genes were used for genetic diversity analysis. All primers were designed by Primer Premier 5 software (Premier Biosoft, San Francisco, CA, USA) and synthesised by Shanghai Sangon Biotech (Shanghai, China).

**Identification of trematodes.** Following Periago’s method (21), the worms were morphologically identified by microscopy according to size, shape, whether the ‘shoulders’ were obvious, the size and position of the oral and ventral suckers, the position of the genital pore, the yolk glands and lateral branches of the intestinal canal, and the position of the excretory canal. After morphological identification, each worm tissue sample was cropped to 20 mg and ground into a pulp, and the genomic DNA of Fasciola fluke was extracted using a TaKaRa miniBEST Universal Genomic DNA Extraction Kit (TaKaRa Bio, Dalian, China). The ITS1 sequence was amplified by PCR using the F1 and R1 primers. The amplification procedure was as follows: pre-denaturation at 94°C for 4 min, denaturation at 94°C for 40 s, annealing at 50°C for 40 s and extension at 72°C for 50 s through 35 cycles, and extension at 72°C for 10 min. The amplified product was identified by 1.5% agarose gel electrophoresis and purified using TIANgel Purification Kit (TIANGEN, China) and then sent to Shanghai Sangon Biotech for sequencing.

**PCR amplification of cox1 and nad1 genes.** The cox1 and nad1 genes were amplified using the F2 and R2 and F3 and R3 primers, respectively. Twenty microlitres of PCR reaction mixture were prepared, including 10 μL of double-distilled H2O, 7 μL of PCR Mixture (Thermo Fisher Scientific, Waltham, MA, USA), 2 μL of DNA, and 0.5 μL of each forward and reverse primer. The PCR amplification procedure for the cox1 and nad1 genes was as follows: pre-denaturation at 94°C for 4 min, denaturation at 94°C for 40 s, annealing at 50°C (or 55°C) for 40 s and
extension at 72°C for 50 s through 35 cycles, followed by extension at 72°C for 10 min. The PCR amplification products were analysed by 1.5% agarose gel electrophoresis and purified using TIANgel Purification Kit for sequencing. Each PCR product was sequenced three times. Only the sequence from identical sequencing results was used for alignment analysis.

**Genetic analysis of different geographical populations of F. hepatica isolates.** DNAMAN software (Lynnon Biosoft, San Ramon, CA, USA) was used to align the sequencing results, and the sequences were analysed for homology using BLAST in the NCBI database (https://www.ncbi.nlm.nih.gov). The sequences were analysed for base composition and variable sites using MEGA 7.0 software (17). The number of haplotypes, haplotype diversity (Hd), nucleotide diversity (π), average number of nucleotide differences (K) were calculated and mismatch distribution curves were generated using DnaSP software (23). Haplotype network diagrams and a neighbour-joining phylogenetic tree were constructed using PopART (19) and MEGA 7.0 software, respectively, and the genetic variation of different geographical populations of F. hepatica isolates was analysed using the analysis of variance in Arlequin version 3.5.2 software (12).

**Results**

**Morphological identification of parasites.** These worms are flattened dorsoventrally and have a leaf-like appearance; the surface of the body is covered with dermal spines, and there is a triangular conical protrusion at the anterior end of the body with a distinct ‘shoulder’, which gradually narrows posteriorly. The oral sucker is rounded and located at the anterior end of the ventral sucker, the ventral sucker is larger than the oral sucker, and there is a genital pore between the oral and ventral suckers. There are many lateral branches in the intestines on both sides of the worm body, and the yolk glands composed of brown particles overlapping the lateral branches are distributed with excretory tubes at the rear end of the body. The parasites collected in Xinjiang, preliminarily identified as F. hepatica, were found to have an average length of 2.3 cm and width of 0.98 cm (Fig. 1).

**Molecular identification of parasites.** The ITS1 gene was amplified by PCR and sequenced (Supplementary Fig. S2), and all of the collected worms were identified as F. hepatica. The PCR amplification results showed that the cox1 and nad1 genes were 496 bp and 660 bp in size, respectively (Supplementary Fig. S3), and were consistent with the expected amplification sizes. The sequence analyses based on 144 sequences of the cox1 and nad1 genes revealed that the base composition of the sequences contained significantly more A+T bases (63.8% and 62.3%, respectively) than C+G (36.2% and 37.7%, respectively), with a clear A/T preference in the sequence base composition. The cox1 gene sequences contained 21 variable sites (comprising 10 parsimony informative sites and 11 singleton sites) (Supplementary Table S1), while 42 variable sites (comprising 13 parsimony informative sites and 29 singleton sites) were harbored in the nad1 genes (Supplementary Table S2). However, there were no base insertions or deletions in any of the sequences, and only base conversions and reversals occurred.

**Genetic diversity and haplotype network analysis.** Genetic diversity analysis of the three populations taking 144 sequences into consideration identified 34 and 33 haplotypes based on the cox1 and nad1 genes, respectively (Table 1). The haplotype diversity (Hd) of the Urumqi (0.909 ± 0.00089) and Kuqa (0.889 ± 0.00828) populations was higher, and the nucleotide diversity (π) of the Hami population was the highest.

A haplotype network map based on the cox1 gene sequences (Fig. 2A) shows that there are 13 exclusive and 21 shared haplotypes. The dominant ones are Hap_4 and Hap_12, having and being distributed in each geographical population, and the remaining haplotypes are connected to these dominant ones by mutation. The phylogenetic tree (Fig. 3A) presents the haplotypes clustered mainly into two branches, which is consistent with the results of the haplotype network map. There are 20 shared haplotypes based on the nad1 gene in the haplotype network map (Fig. 2B), and the rest are exclusive haplotypes. The frequency of the Hap_1 haplotype is the highest, and this one dominates the network linking outwards to other haplotypes in a star shape. The phylogenetic tree (Fig. 3B) presents two branches of haplotype clustering, with individuals of different geographical origin mixed together and not grouped by geographical region, which is generally consistent with the results of the haplotype network map. The study showed that F. hepatica isolated from the three geographic populations in Xinjiang did not display significant geographic variability.

**Fig. 1.** Morphological identification of parasites from sheep. A – parasites found in bile ducts of sheep; B – parasites fixed on slides with 75% alcohol after washing with phosphate buffered saline (0.01M, pH 7.2–7.4)
Table 1. Genetic diversity parameters of the three geographic populations of *F. hepatica*

<table>
<thead>
<tr>
<th>Genes</th>
<th>Populations</th>
<th>N</th>
<th>h</th>
<th>Hd</th>
<th>K</th>
<th>π</th>
</tr>
</thead>
<tbody>
<tr>
<td>coxl</td>
<td>Urumqi</td>
<td>45</td>
<td>18</td>
<td>0.909 ± 0.00089</td>
<td>3.160</td>
<td>0.00637</td>
</tr>
<tr>
<td></td>
<td>Hami</td>
<td>90</td>
<td>23</td>
<td>0.890 ± 0.00047</td>
<td>3.280</td>
<td>0.00661</td>
</tr>
<tr>
<td></td>
<td>Kuqa</td>
<td>9</td>
<td>6</td>
<td>0.833 ± 0.01600</td>
<td>2.722</td>
<td>0.00549</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>144</td>
<td>34</td>
<td>0.894 ± 0.00034</td>
<td>3.287</td>
<td>0.00663</td>
</tr>
<tr>
<td></td>
<td>Urumqi</td>
<td>45</td>
<td>12</td>
<td>0.835 ± 0.00118</td>
<td>2.638</td>
<td>0.00400</td>
</tr>
<tr>
<td></td>
<td>Hami</td>
<td>90</td>
<td>25</td>
<td>0.7590 ± 0.00218</td>
<td>2.778</td>
<td>0.00421</td>
</tr>
<tr>
<td></td>
<td>Kuqa</td>
<td>9</td>
<td>6</td>
<td>0.889 ± 0.00828</td>
<td>2.583</td>
<td>0.00391</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>144</td>
<td>33</td>
<td>0.805 ± 0.00086</td>
<td>2.969</td>
<td>0.00450</td>
</tr>
</tbody>
</table>

N – number of samples; h – number of haplotypes; Hd – haplotype (gene) diversity; K – average number of nucleotide differences; π – nucleotide diversity

Fig. 2. Median-joining network of *F. hepatica* isolates based on *cox1* (A) and *nad1* (B) genes’ haplotypes. The size of the circle indicates the frequency of the haplotype. The colour of the circle indicates the origin of the population. The green dot represents the default vertex.
Fig. 3. Phylogenetic tree of haplotypes of different geographic populations of *F. hepatica* in Xinjiang. The tree is based on *cox1* (A) and *nad1* (B) and constructed with the neighbour-joining method.

### Analysis of the population’s genetic structure.

The results of the analysis of variance are shown in Table 2. The pairwise fixation indices (F–statistic, $-$ Fst) are 0.04451 and 0.03520 when targeting the *cox1* and *nad1* genes, respectively, and the P values are all greater than 0.05. The corresponding gene flows are 5.37 and 6.85 (31), respectively, indicating that the gene communication among the populations is sufficient and genetic differentiation is low. The respective percentages of intrapopulation genetic variation are 99.7% and 96.48%, while those of interpopulation variation are 0.3% and 3.52%, indicating that the genetic variation is mainly present within the populations rather than between them.

The results of the neutrality test (Table 3) showed that Tajima’s D ($-$1.01276, P > 0.10) and Fu’s Fs values ($-$19.204, P > 0.10) were not significant for the total population based on the *cox1* gene, the mismatch distribution curve (Fig. 4A) was bimodal, and the test of each geographical population did not yield any indicators of significance, inferring that the populations were relatively stable and did not expand. However, the Tajima’s D ($-$1.92702, P < 0.05) test based on the total population of the *nad1* gene was
significant and the mismatch distribution curve (Fig. 4B) was single-peaked. The Tajima’s D for this gene of the Hami population reached a significant value, while this indicator and Fu’s Fs tests did not for the other two populations, suggesting that the Hami population may have deviated from the neutral theory of evolution and undergone short-term expansion.

**Fig. 4.** Mismatch distribution of haplotypes of the *cox1* (A) and *nad1* (B) genes. The number of nucleotide differences between pairs of sequences is indicated along the x-axis, and their frequency along the y-axis. Obs – observed value; Exp – expected value

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Gene</th>
<th>Degrees of freedom</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage of variation (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within populations</td>
<td><em>cox1</em></td>
<td>141</td>
<td>62.933</td>
<td>0.44634 Vb</td>
<td>99.70</td>
</tr>
<tr>
<td></td>
<td><em>nad1</em></td>
<td>141</td>
<td>55.711</td>
<td>0.39511 Vb</td>
<td>96.48</td>
</tr>
<tr>
<td>Among populations</td>
<td><em>cox1</em></td>
<td>2</td>
<td>0.990</td>
<td>0.00133 Va</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td><em>nad1</em></td>
<td>2</td>
<td>1.844</td>
<td>0.01442 Va</td>
<td>3.52</td>
</tr>
<tr>
<td>Total</td>
<td><em>cox1</em></td>
<td>143</td>
<td>63.924</td>
<td>0.44767</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>nad1</em></td>
<td>143</td>
<td>57.556</td>
<td>0.40953</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.04451, P &gt; 0.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.03520, P &gt; 0.05</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.** Molecular variation analysis based on *cox1* and *nad1* gene sequences
Table 3. Neutrality indices based on cox1 and nad1 genes of F. hepatica in different geographical regions of Xinjiang

<table>
<thead>
<tr>
<th>Population</th>
<th>cox1</th>
<th></th>
<th>nad1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tajima’s D</td>
<td>Fu’s Fs</td>
<td>Tajima’s D</td>
<td>Fu’s Fs</td>
</tr>
<tr>
<td>Ürumqi</td>
<td>-0.74410</td>
<td>-7.514</td>
<td>-0.72079</td>
<td>-2.547</td>
</tr>
<tr>
<td>Hami</td>
<td>-0.36276</td>
<td>-8.659</td>
<td>-1.94511*</td>
<td>-13.448</td>
</tr>
<tr>
<td>Kuqa</td>
<td>-0.34120</td>
<td>-1.395</td>
<td>-0.55536</td>
<td>-1.522</td>
</tr>
<tr>
<td>Total</td>
<td>-1.01276</td>
<td>-19.204</td>
<td>-1.92702*</td>
<td>-19.881</td>
</tr>
</tbody>
</table>

* Represents significant deviation from neutrality (P < 0.05)

Discussion

As a re-emerging neglected tropical disease, fascioliasis has become a major food safety and public health concern because of the increasing infection risks. Human fascioliasis especially has come to pose a great threat to public health. The unique biological properties of F. hepatica may be some of the factors driving its genetic diversity and shaping its population structure. Firstly, F. hepatica has not only normal spermatogenetic capability but also those of autofertilisation and cross-fertilisation. In some conditions, it can reproduce by parthenogenesis, and this particular mode of reproduction may cause changes in the allele frequencies in defined populations (10, 26, 27). Furthermore, the requirement of F. hepatica for intermediate hosts (mainly freshwater snails) and definitive hosts to complete its life cycle may also foster greater heterogeneity. Snails are a highly invasive and adaptive group of organisms with a wide range of distribution, and the potential for clonal reproduction during the asexual developmental stages of miracidia in snails, coupled with the extensive movement of definitive hosts, could facilitate the spread of F. hepatica populations and promote higher levels of gene exchange, leading to genetic diversity (4, 22, 28). In addition, adult worms of F. hepatica can survive in hosts for long periods of time, and large numbers of eggs can be excreted in feaces from untreated livestock, potentially resulting in genotypic diversity of the worms. Environmental changes such as temperature, humidity, altitude and farm management may also mould the population structure of F. hepatica (1, 13).

Genetic diversity can reflect the evolutionary potential of a species. It is generally accepted that the higher the level of genetic diversity, the greater the adaptability to the environment. Currently, it has been shown that the level of diversity and population structure varies considerably between geographical regions. Bozorgomid et al. (8) analysed the genetic diversity of the nad1 gene from F. hepatica isolates from cattle, sheep and goats in Iran, and found a total of 37 haplotypes, with a high level of population genetic diversity. Beesley et al. (2) found a high level of population genetic diversity in Argentinean F. hepatica isolates and identified a total of 263 unique genotypes. Ichikawa-Seki et al. (16) revealed that 11 haplotypes belonged to a F. hepatica population and 18 haplotypes belonged to a F. gigantica population in 211 lamellipodia from partial geographical regions of China, and Beja-Pereira et al. (5) observed that the ‘hybrid Fasciola fluke’ type harboured the highest number of haplotypes and presented high haplotype diversity and nucleotide diversity values. In the present study, 34 and 33 haplotypes were detected from 144 sequences of cox1 and nad1 genes, respectively, and the Hami population had the highest number of haplotypes and the highest π and K values. The occurrence of this phenomenon in the Hami population may be related to the mild and humid climate and the presence of a large number of water-filled depressions, which provide the best environment for intermediate hosts to survive (6, 9, 11).

The median-joining network and phylogenetic trees showed that 144 F. hepatica isolates from three different geographical sources in Xinjiang were mixed and distributed without significant geographical differentiation. The analysis of variance showed that the Fst ranged from 0 to 0.05, indicating a small genetic differentiation (18). The respective proportions of intrapopulation genetic variation for cox1 and nad1 were 99.7% and 96.48%, indicating that the three geographic populations in Xinjiang had a low degree of genetic differentiation but adequate gene exchange mainly deriving from intrapopulation variation. The results of the neutrality test indicated that the populations were relatively stable but the Hami population may have undergone short-term expansion. The present study has for the first time revealed the molecular characteristics, population genetic relationships and dispersal patterns of F. hepatica in sheep in Xinjiang, China, which is of great importance for the better understanding of its genetic diversity and the developing of reasonable management measures to control sheep fascioliasis.

Conflict of Interests Statement: The authors declare that there is no conflict of interests regarding the publication of this article.
Financial Disclosure Statement: This work was supported by a grant from the National Key Research and Development Program (No. 2017YFD0501200), the International Science & Technology Cooperation Program of XPC (No. 2016AH006) and the Xinjiang Autonomous Region graduate innovation project (No. XJ2020G085).

Animal Rights Statement: The experiments on animals were approved by the Research and Ethical Committee of Sihize University (RECSHZ2020166).

Acknowledgements: We thank the field staff for providing technical assistance for this study.

References


30. Webb C.M., Cabada M.M.: Recent developments in the epidemiology, diagnosis, and treatment of Fasciola infection.


