DOI: 10.1002/vms3.995

ORIGINAL ARTICLE

WILEY

# Assessment of genetic markers for multilocus sequence typing (MLST) of *Fasciola* isolates from Iran

Naser Nazari[1] | Mohamad Bagher Rokni[2] | Madoka Ichikawa-Seki[3] | Saber Raeghi[4] | Homa Hajjaran[2] | Shahab Falahi[5] | Yazdan Hamzavi[1] | Peyman Heydarian[6] | Afshin Davari[2] | Keyphobad Ghadiri[1] | Arezoo Bozorgomid[1]

[1]Infectious Diseases Research Center, Health Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran

[2]Department of Medical Parasitology and Mycology, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

[3]Laboratory of Veterinary Parasitology, Faculty of Agriculture, Iwate University, Morioka, Japan

[4]Department of Laboratory Sciences, Maragheh University of Medical Sciences, Maragheh, Iran

[5]Zoonotic Diseases Research Center, Ilam University of Medical Sciences, Ilam, Iran

[6]Department of Medical Parasitology and Mycology, School of Medicine, Qazvin University of Medical Sciences, Qazvin, Iran

**Correspondence**
Arezoo Bozorgomid and Keyphobad Ghadiri, Infectious Diseases Research Center, Health Institute, Kermanshah University of Medical Sciences, Kermanshah, Iran.
Email: arezoobozorgomid@yahoo.com; K_ghadiri@yahoo.com

**Funding information**
Kermanshah University of Medical Sciences, Kermanshah, Iran, Grant/Award Number: 990716

## Abstract

**Background:** Several markers have been described to characterise the population structure and genetic diversity of *Fasciola* species (*Fasciola hepatica (F. hepatica)* and *Fasciola gigantica* (*F. gigantica*). However, sequence analysis of a single genomic locus cannot provide sufficient resolution for the genetic diversity of the *Fasciola* parasite whose genomes are ~1.3 GB in size.

**Objectives:** To gain a better understanding of the gene diversity of *Fasciola* isolates from western Iran and to identify the most informative markers as candidates for epidemiological studies, five housekeeping genes were evaluated using a multilocus sequence typing (MLST) approach.

**Methods:** MLST analysis was developed based on five genes (*ND1*, *Pepck*, *Pold*, *Cyt b* and *HSP70*) after genomic DNA extraction, amplification and sequencing. Nucleotide diversity and phylogeny analysis were conducted on both concatenated MLST loci and each individual locus. A median joining haplotype network was created to examine the haplotypes relationship among *Fasciola* isolates.

**Results:** Thirty-three *Fasciola* isolates (19 *F. hepatica* and 14 *F. gigantica*) were included in the study. A total of 2971 bp was analysed for each isolate and 31 sequence types (STs) were identified among the 33 isolates (19 for *F. hepatica* and 14 for *F. gigantica* isolates). The STs produced 44 and 42 polymorphic sites and 17 and 14 haplotypes for *F. hepatica* and *F. gigantica*, respectively. Haplotype diversity was $0.982 \pm 0.026$ and $1.000 \pm 0.027$ and nucleotide diversity was 0.00200 and $0.00353 \pm 0.00088$ for *F. hepatica* and *F. gigantica*, respectively. There was a high degree of genetic diversity with a Simpson's index of diversity of 0.98 and 1 for *F. hepatica* and *F. gigantica*, respectively. While *HSP70* and *Pold* haplotypes from *Fasciola* species were separated by one to three mutational steps, the haplotype networks of *ND1* and *Cyt b* were more complex and numerous mutational steps were found, likely due to recombination.

**Conclusions:** Although *HSP70* and *Pold* genes from *F. gigantica* were invariant over the entire region of sequence coverage, MLST was useful for investigating the phylogenetic relationship of *Fasciola* species. The present study also provided insight into markers more suitable for phylogenetic studies and the genetic structure of *Fasciola* parasites.

**KEYWORDS**
*Fasciola hepatica*, *Fasciola gigantica*, genetic diversity, Iran, multilocus sequence typing

## 1 | BACKGROUND

Fasciolosis, a geographically widespread disease, is caused by the liver flukes *Fasciola hepatica* (*F. hepatica*) Linnaeus, 1758 and *Fasciola gigantica* (*F. gigantica*) Cobbold, 1856. The lifecycle of both species requires a mammal's definitive host harbouring adult worms in its bile duct and lymnaeid snails as intermediate hosts infected by the larval stage (Mas-Coma et al., 2019; Saadatnia et al., 2022).

Fasciolosis is considered as the most important helminth infection of the livestock in Iran with a prevalence ranging from 0.1% to 91.4% (Ashrafi, 2015). The endemic nature of fasciliasis leads to economic concerns not only because of decreased a growth, weight, fertility and condemnation of the infected liver but also for public health reasons as two large outbreaks of human fascioliasis occurred during the past two decades in Iran (Bashiri et al., 2021; Bozorgomid et al., 2018). A combination of slaughterhouse and molecular studies revealed the presence of *F. hepatica* and *F. gigantica* as well as hybrids between the two species in Iran; however, few studies have reported the genetic diversity of these parasites (Bozorgomid et al., 2018; Rokni et al., 2018).

Investigating the genetic diversity of *Fasciola* isolates in the livestock is of interest because they play a key role in spreading infection. *F. hepatica* and *F. gigantica* can generally be distinguished on the basis of morphological criteria but the existence of parthenogenic (aspermic) *Fasciola* flukes with intermediate morphological characteristics between the two species can cause confusion (Aryaeipour et al., 2017). Several markers have been developed and used to assess the population structure of *Fasciola* species such as the ribosomal internal transcribed spacer (ITS), phosphoenolpyruvate carboxykinase (*Pepck*), DNA polymerase delta (*Pold*), cytochrome c oxidase I (*COI*) and the NADH dehydrogenase (*ND*) genes (Bozorgomid et al., 2020; Heydarian et al., 2022; Kasahara et al., 2021; Shafiei et al., 2014). It has been reported that the genetic diversity of this parasite may have a role in drug resistance, virulence, pathogenicity, clinical characteristics of the disease and certain epidemiological features (Lalor et al., 2021). However, sequence analysis of a single genomic locus cannot provide sufficient resolution for the genetic diversity of the *Fasciola* parasite whose genomes are ~1.3 GB in size (Cwiklinski et al., 2015).

Multilocus sequence typing (MLST) was originally recommended as a power discriminatory method and reproducibility based on single nucleotide polymorphism (SNP) in 5–7 housekeeping genes for phylogenetic study of bacterial pathogens (Maiden et al., 2013). It was later used to characterise population genetic structure of different parasite species such as *Enterocytozoon bieneusi*, *Giardia duodenalis*, *Cryptosporidium* sp., *Leishmania* spp., *Trichomonas vaginalis* and *Trypanosoma cruzi* (Banu et al., 2019; Costache et al., 2020; Herrera et al., 2017; Li et al., 2021; Malatyali et al., 2020). To gain a better understanding of the phylogenetic and population structure of *Fasciola* species, the MLST scheme was developed based on five genes of *Fasciola* isolates in livestock from western Iran.

## 2 | METHODS

### 2.1 | Samples collection and genomic DNA extraction

A total of 68 flukes were obtained from the infected livers (sheep $n = 16$, cattle $n = 48$, and goat $n = 4$) during randomly repeated visits of slaughterhouses in Kermanshah Province, Iran from April 2015 to July 2017. The specimens were preserved in ethanol (70%) and transferred to the Helminthology Laboratory, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran. Thirty-five adult fluke were randomly selected for molecular studies. Genomic DNA was extracted from the posterior part of each fluke using the DNGTM-plus Kit (CinnaGen, Tehran, Iran) according to the manufacturer's instructions and stored at – 20°C until use. Samples were previously identified as *F. hepatica* or *F. gigantica* based on PCR-RFLP of the ITS-1 region with RsaI enzyme and the samples of *F. gigantica* were sequenced based on the *ND1* gene (Rokni et al., 2020).

### 2.2 | Selection of loci, amplification and sequencing

Amplification was performed based on five genes including heat shock protein 70 (*HSP70*), DNA polymerase delta (*Pold*), phosphoenolpyruvate carboxykinase (*Pepck*), cytochrome b (*Cyt b*) and NADH dehydrogenase subunit 1 (*ND1*). The primers for the target genes were designed using the publicly available Primer3plus program and checked for specificity using the NCBI Primer-BLAST tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). The primers were synthesised by Macrogen Inc., Korea. All primers used in this study are listed in Table 1.

PCR amplification reactions were performed in a final volume of 30 μl containing 0.5 μM of each primer, 15 μl Red Master Mix

**TABLE 1** Characterisation of gene markers and primers used in this study

| Target gene | Primer sequence | Amplicon size (bp) | Size of aligned fragment analysed (bp) | Annealing temp, (°C) | Reference orsource |
|---|---|---|---|---|---|
| **Heat shock protein 70 (*HSP70*)** | F: GATTGTTCTCGTTGGTGGTTCG<br>R: GCATTTGCTTGATATGGTGCTG | 724 | 668 | 62 | This study |
| **DNA polymerase delta (*Pold*)** | F: ATCCTATCCCCGTCGTATGCTG<br>R: GCTTATGGTGAGAGGCGAGG | 630 | 398 | 62 | This study |
| **Phosphoenolpyruvate carboxykinase (*Pepck*)** | F: TCGCCTGCGTGCTATCAATC<br>R: GCATCACCAACGGCACAC | 822–876 | 810–811 | 62 | This study |
| **Cytochrome b (Cyt b)** | F: CAGGTGTTGTCTGGGGTTG<br>R: CCAACCACAATCCCTTAAACAT | 970 | 523 | 64 | This study |
| **NADH dehydrogenase subunit 1 (*ND1*)** | F: AAGGATGTTGCTTTGTCGTGG<br>R: GGAGTACGGTTACATTCACA | 660 | 570 | 55 | This study |

(Ampliquon, Denmark), 0.5 μM of each primer and 1.25 μl of DNA (<250 ng). Amplification was conducted in a thermal cycler (Eppendorf, Germany) with an 5-min initial denaturation at 95°C, followed by 30 amplification cycles [denaturation at 94°C for 45 seconds, annealing at the appropriate temperature for each gene for 45 s (Table 1) and elongation at 72°C for 1 min] followed by a final extension step at 72°C for 5 min. The amplification products (3 μl) were electrophoresed on a 1.5% agarose gel containing the DNA stain Simply Safe (Eurx, Cat. No. E4600-01). The PCR products were purified using the QIAquick PCR Purification Kit (Qiagen). DNA sequencing of purified PCR products was carried out at Macrogen Inc., Korea using similar primers used for the PCR.

## 2.3 | Genetic differentiation and haplotype network analysis

The sequences for each gene were visually checked for the detection of any missing or ambiguous (heterozygous) sites when two peaks overlapped in a chromatogram. The sequences were edited and aligned using the Bioedite software v. 1.9 (Hall, 1999). To determine the genetic diversity, the following indices were calculated for each gene and each species using the DnaSP v5.1 software (Rozas et al., 2017): the number of segregating sites (polymorphic) (S), number of haplotypes (H), haplotype diversity, nucleotide diversity ($\pi$) and their corresponding standard deviations (SD). Tajima's *D* values were calculated to test the neutral theory of evolution. Positive values correspond to positive or balancing selection, whereas negative values correspond to negative or purifying selection. A median joining haplotype network was created to examine the haplotypes relationship among *Fasciola* isolates using the PopART software (Leigh & Bryant, 2015).

## 2.4 | Phylogenetic analysis

The MEGA 6.0 software was used to choose the best-fitting models of molecular evolution (Kumar et al., 2018). Phylogenetic relationships among individual gene sequences and concatenated sequences were constructed using the Maximum Likelihood (ML) method. Furthermore, for the concatenated sequences, Neighbour-Net analysis was performed with SplitsTree version 4.11.3 using uncorrected p-distances (Huson & Bryant, 2006). Both phylogenetic analyses were done with 1000 bootstrap replicates.

## 2.5 | Discriminatory power

To assess discriminatory power of different genes, Simpson's index of diversity was calculated based on the probability of two unrelated strains being characterised as the same type (Hunter & Gaston, 1988). The Simpson's index of diversity values are between 0.0 and 1.0. A value of 0.0 indicates that all strains in a population are of the same type and a value of 1.0 indicates high species diversity and complex community.

# 3 | RESULTS

## 3.1 | PCR amplification and DNA sequencing

Of 35 DNA samples extracted, 33 were amplified by PCR and sequenced successfully for the five selected markers (*F. hepatica*: $n = 19$, *F. gigantica*: $n = 14$). Two samples were excluded from further analysis because of the poor quality of the sequences of one or more markers.

## 3.2 | SNP analysis and nucleotide diversity

Polymorphic sites ranged between 4 (*Cyt b* and *HSP70*) and 19 (*Pepck*) for *F. hepatica* and between 0 (*Pold* and *HSP70*) and 21 (*ND1*) for *F. gigantica*. The nucleotide diversity varied between 0.00103 ± 0.00041 (*HSP70*) and 0.00630 ± 0.00216 (*Pepck*) for *F. hepatica* and between 0 (*Pold* and *HSP70*) and 0.00899 ± 0.00182 (*Cyt b*) for *F. gigantica*. For *F. hepatica*, the highest discriminatory power and haplotype diversity were observed in the *ND1* gene (Table 2). For *F. gigantica*, the highest

**TABLE 2** Assessment of DNA polymorphism and neutrality test for the *F. hepatica* and *F. gigantica* isolates

| | Gene | Length (bp) | S | DP | H | $\pi \pm SD$ | *k* | Hd ± SD | Tajima's D |
|---|---|---|---|---|---|---|---|---|---|
| *F. hepatica* $n = 19$ | *Cyt b* | 523 | 4 | 0.505 | 5 | 0.00161 ± 0.00054 | 0.84211 | 0.4620 ± 0.136 | −0.77791 |
| | *ND1* | 570 | 11 | 0.8129 | 9 | 0.00369 ± 0.00089 | 2.10526 | 0.813 ± 0.081 | −1.19054 |
| | *HSP70* | 668 | 4 | 0.4561 | 4 | 0.00103 ± 0.00041 | 0.69006 | 0.4561 ± 0.132 | −1.16911 |
| | *Pold* | 398 | 6 | 0.5556 | 3 | 0.00262 ± 0.00119 | 1.04094 | 0.556 ± 0.073 | −1.69834 |
| | *Pepck* | 810 | 19 | 0.7368 | 5 | 0.00630 ± 0.00216 | 4.69006 | 0.737 ± 0.082 | −0.52559 |
| | Concatenated | 2971 | 44 | 0.9825 | 17 | 0.00323 ± 0.00063 | 9.36842 | 0.982 ± 0.026 | −1.03539 |
| *F. gigantica* $n = 14$ | *Cyt b* | 523 | 14 | 0.9011 | 8 | 0.00899 ± 0.00182 | 4.70330 | 0.9011 ± 0.058 | 0.27687 |
| | *ND1* | 570 | 21 | 0.6484 | 4 | 0.00626 ± 0.00370 | 3.56044 | 0.648 ± 0.116 | −1.93337* |
| | *HSP70* | 668 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| | *Pold* | 398 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| | *Pepck* | 811 | 7 | 0.6593 | 4 | 0.00272 ± 0.00048 | 2.20879 | 0.736 ± 0.109 | 0.01278 |
| | Concatenated | 2971 | 42 | 1 | 14 | 0.00353 ± 0.00088 | 10.47253 | 1.000 ± 0.027 | −0.90230 |

**Abbreviations:** S, number of polymorphic sites; DP, discriminatory power; H, number of haplotype; $\pi$, nucleotide diversity; *k*, average number of nucleotide differences; Hd, haplotype diversity.

discriminatory power and haplotype diversity are observed in the *Cyt b* gene.

The Tajima's *D* neutrality test was also conducted on the loci sequences. Tajima's *D* tests showed no statistical significance for all genes in *F. hepatica* and for four out of five genes in *F. gigantica*, indicating that a neutral model of sequence evolution could not be rejected. The details of DNA polymorphisms found in the five MLST loci obtained from all 33 isolates are shown in Table 2. The sequences of the *F. hepatica* ($n = 19$) and *F. gigantica* ($n = 14$) strains were deposited in the GenBank as described in Table 3

## 3.3 | Haplotype network analysis

Each gene fragment produced a haplotype network with a different pattern (Figure 1). The *HSP70* and *Pold* network were the simplest and a straight-line pattern was formed rather than a network. Both *HSP70* and *Pold* loci had only one haplotype for *F. gigantica*. For *F. hepatica*, the haplotypes were separated by one to three mutational steps for the *HSP70* locus and one to five mutational steps for the *Pold* locus. As for the *Pepck* network, one or two branches emerged from a highly frequent haplotype (Hap 1 and Hap 6). The most distant haplotypes (Hap 5 and Hap7) were separated by 19 and 5 mutational steps, respectively. The *ND1* and *Cyt b* networks were the most variable loci.

## 3.4 | Analysis of MLST data

For both *Fasciola* species, the sequence data of all five genes were concatenated making a multilocus gene of 2971 bp length. Thirty-one sequence types (STs) were identified from 33 isolates (17 for *F. hepatica* and 14 for *F. gigantica* isolates), 30 (90%) of which were present only once, with the most common ST (ST-1) occurring 3 times in the dataset (Table 3). The 31 STs produced 44 and 42 polymorphic sites and 17 and 14 haplotypes with a haplotype diversity of 0.982 ± 0.026 and 1.000 ± 0.027 and nucleotide diversity of 0.00200 and 0.00353 ± 0.00088 for *F. hepatica* and *F. gigantica*, respectively. *F. hepatica* and *F. gigantica* had a high degree of genetic diversity with a Simpson's index of diversity of 0.98 and 1, respectively.

## 3.5 | Phylogenetic analysis

The constructed phylogenetic tree inferred by each of the individual gene sequences of *Fasciola* isolates is shown in Supplementary Figure S1a–e. Phylogenetic analysis based on the STs of *F. gigantica* revealed two major clusters, cluster A and cluster B. A clade was divided into two sub-clusters with a high statistical support. Although the bootstrap values of phylogenetic tree for STs of *F. hepatica* were generally very poor, two separate clusters were identifiable (Figure 2).

The results from maximum-likelihood phylogram were broadly consistent with the clusters delineated by SplitsTree. SplitsTree decomposition analysis revealed a network-like tree for both *Fasciola* species

**TABLE 3** Details of *Fasciola* samples used in the study, concatenated haplotypes (*ND1*, *Pepck*, *Pold*, *Cyt b* and *HSP70*) and GenBank accession numbers of the corresponding newly generated sequences

| No. | Specimen | Species | Sequence types (ST) | Host | Marker | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | *Cyt b* | | *ND1* | | *HSP70* | | *Pepck* | | *Pold* | |
| | | | | | Accession no. | H | Accession no. | H | Accession no. | H | Accession no. | H | Accession no. | H |
| **1** | 2 | *F. hepatica* | ST1 | Cattle | KY246447 | 1 | MF428477 | 1 | MN370893 | 1 | LC516432 | 1 | MN340359 | 1 |
| **2** | 5 | *F. hepatica* | ST2 | Cattle | KY246448 | 2 | MF428469 | 2 | MN370860 | 2 | LC516429 | 2 | MN386687 | 2 |
| **3** | 6 | *F. hepatica* | ST3 | Sheep | KY246446 | 1 | MF428476 | 3 | MN370861 | 1 | LC516430 | 1 | MN386688 | 1 |
| **4** | 7 | *F. hepatica* | ST4 | Goat | KY246452 | 1 | MF428470 | 4 | MN370862 | 3 | LC516433 | 3 | MN386689 | 1 |
| **5** | 9 | *F. hepatica* | ST5 | Goat | KY246462 | 1 | MF428477 | 1 | MN370863 | 2 | LC516428 | 1 | MN386690 | 1 |
| **6** | 11 | *F. hepatica* | ST6 | Goat | KY246453 | 1 | MF428471 | 5 | MN370864 | 1 | LC516428 | 1 | MN386691 | 2 |
| **7** | 12 | *F. hepatica* | ST7 | Sheep | KY246454 | 1 | MF428476 | 3 | MN370865 | 1 | LC516431 | 3 | MN386692 | 2 |
| **8** | 13 | *F. hepatica* | ST8 | Sheep | KY246454 | 1 | MF428472 | 6 | MN370866 | 1 | LC516434 | 2 | MN386693 | 2 |
| **9** | 17 | *F. hepatica* | ST1 | Sheep | KY246461 | 1 | MF428477 | 1 | MN370867 | 1 | LC516428 | 1 | MN386694 | 1 |
| **10** | 19 | *F. hepatica* | ST9 | Cattle | KY246450 | 3 | MF428473 | 7 | MN370868 | 1 | LC516435 | 4 | MN386695 | 2 |
| **11** | 22 | *F. hepatica* | ST10 | Cattle | KY246449 | 3 | MF428474 | 8 | MN370869 | 1 | LC516435 | 4 | MN386696 | 1 |
| **12** | 23 | *F. hepatica* | ST11 | Sheep | KY246460 | 1 | MF428477 | 1 | MN370870 | 4 | LC516429 | 2 | MN386697 | 3 |
| **13** | 28 | *F. hepatica* | ST1 | Sheep | KY246455 | 1 | MF428477 | 1 | MN370871 | 1 | LC516428 | 1 | MN386698 | 1 |
| **14** | 30 | *F. hepatica* | ST12 | Cattle | KY246451 | 4 | MF428475 | 9 | MN370872 | 1 | LC516428 | 1 | MN386699 | 1 |
| **15** | 33 | *F. hepatica* | ST13 | Cattle | KY246456 | 1 | MF428477 | 1 | MN370873 | 3 | LC516436 | 5 | MN386700 | 2 |
| **16** | 42 | *F .hepatica* | ST14 | Cattle | KY246445 | 5 | MF428477 | 1 | MN370874 | 1 | LC516428 | 1 | MN386701 | 1 |
| **17** | 43 | *F .hepatica* | ST15 | Sheep | KY246457 | 1 | MF428476 | 3 | MN370875 | 1 | LC516428 | 1 | MN386702 | 2 |
| **18** | 48 | *F. hepatica* | ST16 | Cattle | KY246458 | 1 | MF428477 | 1 | MN370876 | 1 | LC516436 | 5 | MN386703 | 1 |
| **19** | 49 | *F. hepatica* | ST17 | Sheep | KY246459 | 1 | MF428472 | 6 | MN370877 | 1 | LC516429 | 2 | MN386704 | 1 |

(Continues)

**TABLE 3** (Continued)

| No. | Specimen | Species | Sequence types (ST) | Host | Marker | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | *Cyt b* | | *ND1* | | *HSP70* | | *Pepck* | | *Pold* | |
| | | | | | Accession no. | H | Accession no. | H | Accession no. | H | Accession no. | H | Accession no. | H |
| **20** | 1 | *F. gigantica* | ST18 | Sheep | KY246463 | 6 | MF428464 | 10 | MN370878 | 5 | LC516441 | 6 | MN386705 | 4 |
| **21** | 2T | *F. gigantica* | ST19 | Sheep | KY246466 | 7 | MF428465 | 11 | MN370879 | 5 | LC516441 | 6 | MN386706 | 4 |
| **22** | 6T | *F. gigantica* | ST20 | Cattle | KY246467 | 8 | MF428468 | 12 | MN370880 | 5 | LC516438 | 7 | MN386707 | 4 |
| **23** | 9T | *F. gigantica* | ST21 | Cattle | KY246475 | 9 | MF428468 | 12 | MN370881 | 5 | LC516439 | 8 | MN386708 | 4 |
| **24** | 14 | *F. gigantica* | ST22 | Cattle | KY246472 | 10 | MF428467 | 13 | MN370883 | 5 | LC516437 | 6 | MN386710 | 4 |
| **25** | 22A | *F. gigantica* | ST23 | Cattle | KY246469 | 11 | MF428468 | 12 | MN370884 | 5 | LC516441 | 6 | MN386711 | 4 |
| **26** | 24 | *F. gigantica* | ST24 | Cattle | KY246470 | 9 | MF428467 | 13 | MN370885 | 5 | LC516442 | 7 | MN386712 | 4 |
| **27** | 25 | *F. gigantica* | ST25 | Cattle | KY246471 | 10 | MF428468 | 12 | MN370886 | 5 | LC516441 | 6 | MN386713 | 4 |
| **28** | 27 | *F. gigantica* | ST26 | Cattle | KY246473 | 12 | MF428468 | 12 | MN370887 | 5 | LC516441 | 6 | MN386714 | 4 |
| **29** | 30T | *F. gigantica* | ST27 | Cattle | KY246474 | 9 | MF428468 | 9 | MN370888 | 5 | LC516442 | 7 | MN386715 | 4 |
| **30** | 39A | *F. gigantica* | ST28 | Cattle | KY246475 | 9 | MF428468 | 9 | MN370889 | 5 | LC516443 | 7 | MN386716 | 4 |
| **31** | 48A | *F. gigantica* | ST29 | Cattle | KY246465 | 13 | MF428465 | 13 | MN370890 | 5 | LC516444 | 9 | MN386717 | 4 |
| **32** | 54 | *F. gigantica* | ST30 | Cattle | KY246468 | 8 | MF428468 | 8 | MN370891 | 5 | LC516441 | 6 | MN386718 | 4 |
| **33** | 58A | *F. gigantica* | ST31 | Cattle | KY246464 | 13 | MF428467 | 13 | MN370892 | 5 | LC516442 | 7 | MN386719 | 4 |

Abbreviation: H, haplotype.

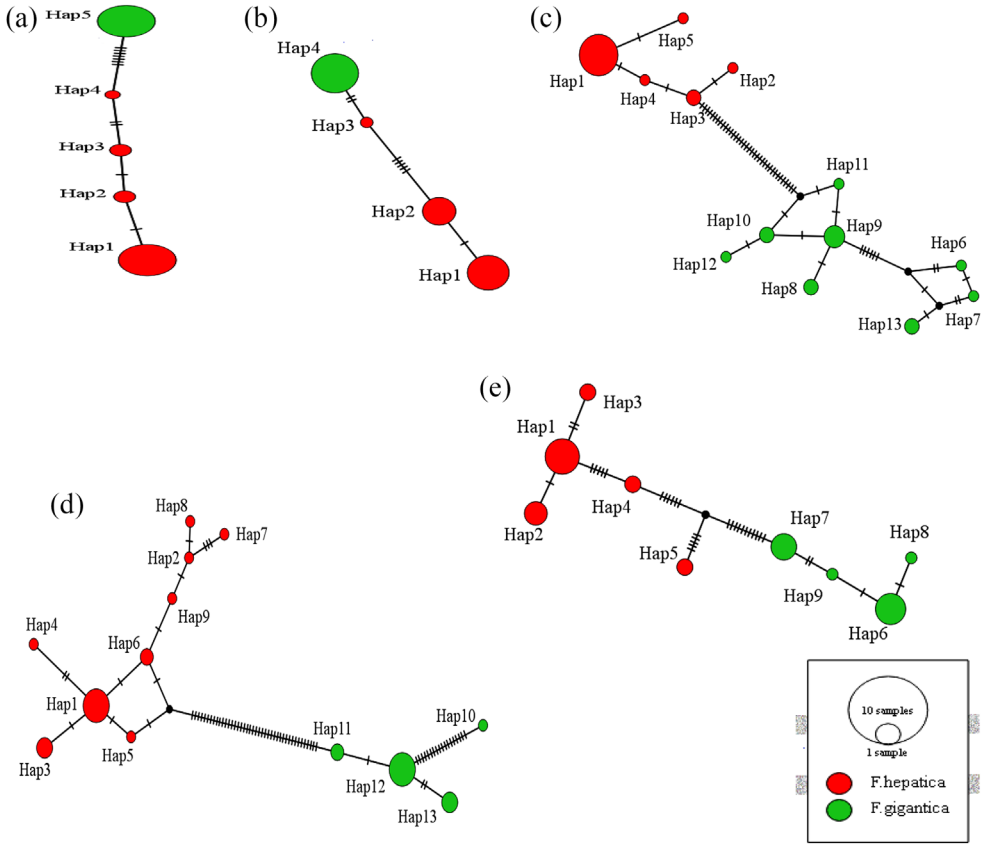


**FIGURE 1** TCS haplotype network obtained from *HSP70* (a), *Pold* (b), *Cyt b* (c), *ND1* (d) and *Pepck* (e) sequences. The area of each circle is proportional to the haplotype frequency, and each branch represents one mutation

(Figure 3). The presence of parallelograms in SplitsTrees was evidence for recombination, which was statistically confirmed by the phi test ($p < 0.05$) for both *Fasciola* species ($p < 0.05$). Thus, recombination is believed to be a relatively common evolutionary process in *Fasciola* flukes. Due to recombination events, *Fasciola* species may be close to one another hampering accurate identification, especially in areas with an overlapping distribution of both species.

## 4 | DISCUSSION

Phylogenetic studies based on both mitochondrial and nuclear DNA genes such as ITS, *Pepck*, *Pold*, *COI* and *ND* are widely used for *Fasciola* species. Although the heat shock protein gene has been used for specific identification of a number of protozoan and helminth species and their differentiation, based on our current knowledge, heat-shock protein 70 (*HSP70*) gene has not been used to evaluate the genetic variation and diversity of *Fasciola* species. Therefore, in the present study, an MLST tool targeting *HSP70* + *ND1* + *Pold* + *Pepck* + *Cytb* genes was developed for *F. hepatica* and *F. gigantica*. The analysis revealed that less 1.5% of the 2971 nucleotides at the five sequenced gene fragments were polymorphic. This tool allowed the identification of at least 17 MLST subtypes of *F. hepatica* and 14 MLST subtypes of *F. gigantica*.

Only three of the five loci examined in this study were polymorphic in *F. gigantica*. Unlike what is already reported for *F. hepatica* and *F. gigantica*, both species were almost identical in nucleotide polymorphism. However, evaluation of more genetic markers and a large number of samples from different geographical areas and more diverse hosts are needed to have a better understanding of the factors affecting the genetic diversity of the *Fasciola* parasite

*HSP70*, a protein with a molecular weight of 70 kDa, is a suitable biomarker for environmental contamination monitoring, and drug resistance (Fuqua et al., 1994; Moreira-de-Sousa et al., 2018). It has also been successfully amplified and employed as a DNA marker for molecular investigations of *Onchocerca fasciata* and some protozoa such as *Giardia*, *Cryptosporidium* and *Leishmania* (Mirzaei et al., 2018; Nemati et al., 2017; Squire & Ryan, 2017). Nevertheless, HSP70 was approximately invariant over the entire region of sequence coverage; hence, it is not recommended for genetic characterisation and phylogenetic analysis of *Fasciola* species.

In previous studies, *Pepck* and *Pold* genes were used to distinguish the aspermic *Fasciola* flukes from those of *F. hepatica* and *F. gigantica* (Itagaki et al., 2022; Thang et al., 2020). The results showed that these genes successfully discriminated *Fasciola* species using the PCR–RFLP and multiplex PCR techniques. By contrast, Kasahara et al. (2021) recently reported that multiplex PCR based on *Pepck* gene was unable to discriminate between *F. hepatica* flukes accurately. They also

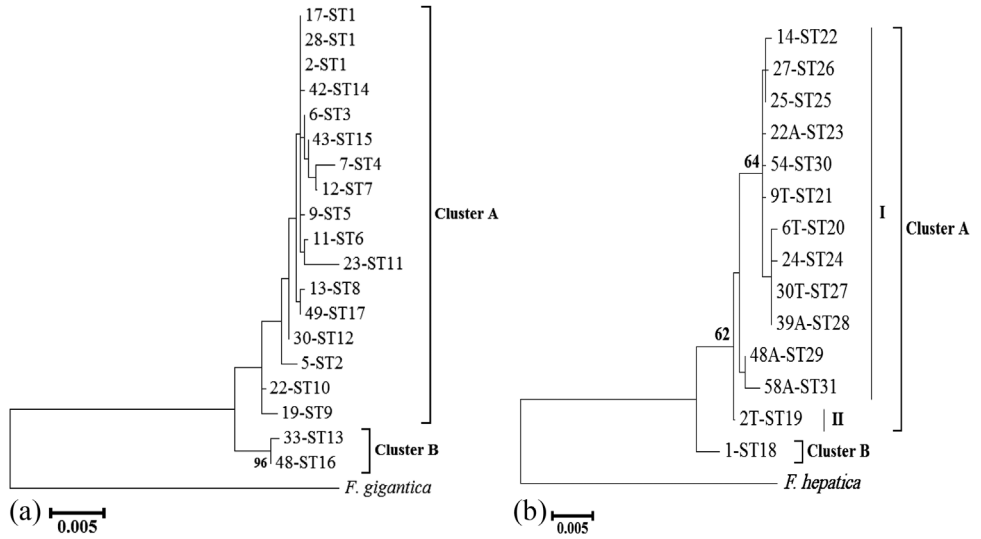


**FIGURE 2** Maximum likelihood phylogeny of *F. hepatica* (a) and *F. gigantica* (b) concatenated data calculated using the Hasegawa-Kishino-Yano (HKY + G) model. The numbers on the branches are per cent bootstrap values from 1000 replicates

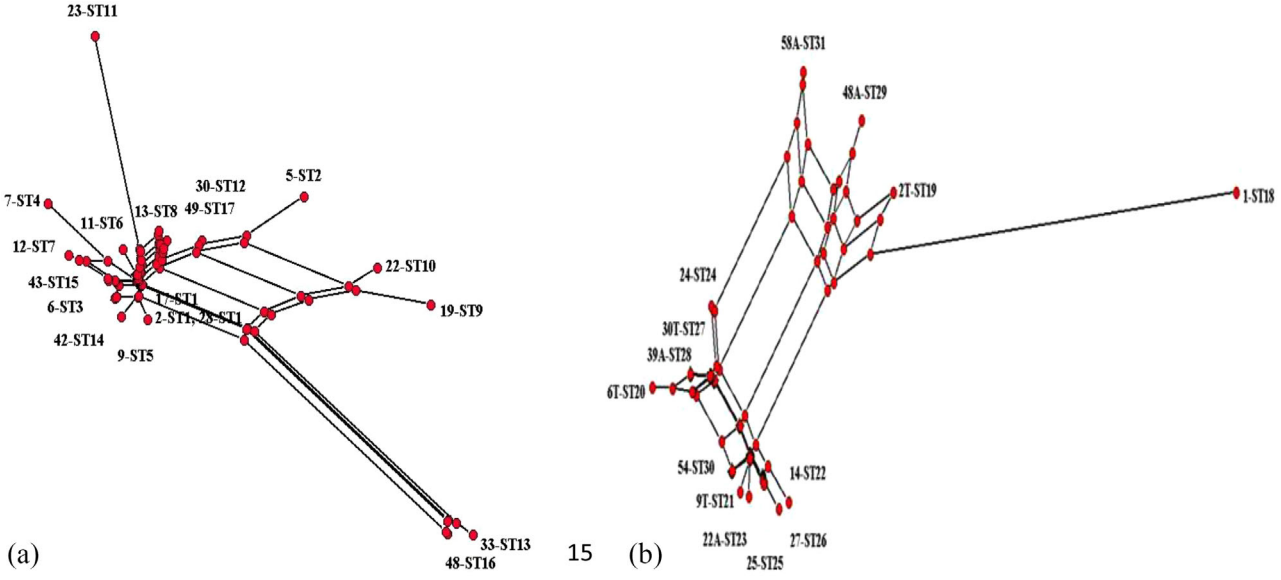


**FIGURE 3** SplitsTree decomposition network of concatenated DNA sequences of five genetic loci for 19 *F. hepatica* (a) and 14 *F. gigantica* (b) isolates obtained from livestock in western Iran. The network was built using the parameters Neighbour-Net and uncorrected p-distance. Bar = 0.001 and represents nucleotide substitutions per site

reported that unlike a number of eukaryotes in which the *Pepck* gene has a single copy, *F. hepatica* has multiple loci. However, the results of the present showed that the use of mitochondrial genes was more suitable for epidemiological studies of *Fasciola* species compared to *Pold* and *Pepck* genes due to the higher diversity of haplotypes and more complex haplotype networks.

The high genetic diversity of the *ND1* gene was evident in most branches of the ML tree and lower genetic diversity of *HSP70* and *Pold* genes were evident in a small number of phylogenetic tree branches. The results of the trees from the sequences of the *ND1*, *Cyt b* and *Pepck* genes showed that even the populations within each group were placed in separate subclades indicating the high accuracy of these three genes in showing the genetic relationships of the populations within each group thus suggesting their appropriateness for population genetic studies. On the other hand, ML trees resulting from the sequences of *HSP70* and *Pold* genes were unable to differentiate between populations in each clade, indicating the lack of necessary efficiency of these two genes in population genetic studies.

Simpson's index, which is used to assess the discriminative ability of typing methods, was higher for MLST than for the single locus typing

method. The small sample sizes and the widespread geographical origin of the samples studied here could lead to the overestimation of Simpson's index. However, the number of isolates used in this assessment was small ($n = 33$), and further investigation is needed to accurately assess the relative abilities of the two methods.

In conclusion, an MLST tool was developed to genotype *Fasciola* species. *HSP70* was approximately invariant over the entire region of the sequence coverage; hence, it is not recommended for genetic characterisation and phylogenetic analysis of *Fasciola* species. Thus, genetic variants should be further studied by searching new candidate genes. These studies should analyse a larger number of specimens from more diverse regions and assess the relationship among MLST subtypes, host specificity, morphological characteristics and geographical distribution.

## ACKNOWLEDGEMENTS

We extend our thanks to the Clinical Research Development Center of Imam Reza Hospital affiliated to Kermanshah University of Medical Sciences for their kind help. This study received financial support from Kermanshah University of Medical Sciences, Kermanshah, Iran (Grant No. 990716).

## ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

The protocol was approved by the Kermanshah University of Medical Sciences Ethics Committee [IR.KUMS.REC.1399.513].

## AUTHOR CONTRIBUTIONS

Conceptualisation, supervision, writing-original draft, data collection, formal analysis, methodology and software: Arezoo Bozorgomid. Conceptualisation, methodology and editing: Mohamad Bagher Rokni. Formal analysis and software: Madoka Ichikawa-Seki. Formal analysis, writing-original draft and software: Saber Raeghi. Conceptualisation, review and editing: Naser Nazari. Conceptualisation, methodology, resources: Homa Hajjaran. Data collection and methodology: Yazdan Hamzavi. Data collection and methodology: Afshin Davari. Data collection, review and editing: Keyphobad Ghadiri. Data collection, formal analysis, methodology and software: Peyman Heydarian. Data collection, review and editing: Shahab Fahahi.

## CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

## ORCID

*Keyphobad Ghadiri* https://orcid.org/0000-0003-1678-6610
*Arezoo Bozorgomid* https://orcid.org/0000-0003-2093-9317

## PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/vms3.995

## REFERENCES

Aryaeipour, M., Bozorgomid, A., Kazemi, B., Behnia, M., Azizi, H., & Rokni, M. B. (2017). Molecular and morphometrical characterization of *Fasciola* species isolated from domestic ruminants in Ardabil Province, Northwestern Iran. *Iranian Journal of Public Health*, *46*(3), 318.

Ashrafi, K. (2015). The status of human and animal fascioliasis in Iran: A narrative review article. *Iranian Journal of Parasitology*, *10*(3), 306.

Banu, S. S., Meyer, W., Ferreira-Paim, K., Wang, Q., Kuhls, K., Cupolillo, E., Schönian, G., & Lee, R. (2019). A novel multilocus sequence typing scheme identifying genetic diversity amongst *Leishmania donovani* isolates from a genetically homogeneous population in the Indian subcontinent. *International Journal for Parasitology*, *49*(7), 555–567.

Bashiri, H., Bozorgomid, A., & Hamzavi, Y. (2021). Biliary of fascioliasis in Kermanshah Province, *Western Iran. Iranian Journal of Public Health*, *50*(1), 215–216.

Bozorgomid, A., Nazari, N., Eshrat Beigom, K., Mohebali, M., Hajaran, H., Hydarian, P., Hamzavi, Y., Nemati, S., Aryaeipour, M., & Rokni, M. B. (2018). Epidemiology of fascioliasis in Kermanshah Province, western Iran. *Iranian Journal of Public Health*, *47*(7), 967.

Bozorgomid, A., Rouhani, S., Harandi, M. F., Ichikawa-Seki, M., & Raeghi, S. (2020). Genetic diversity and distribution of *Fasciola hepatica* haplotypes in Iran: Molecular and phylogenetic studies. *Veterinary Parasitology: Regional Studies and Reports*, *19*, 100359.

Costache, C., Kalmár, Z., Colosi, H. A., Baciu, A. M., Opriş, R. V., Györke, A., & Colosi, I. A. (2020). First multilocus sequence typing (MLST) of *Giardia duodenalis* isolates from humans in Romania. *Parasites & Vectors*, *13*(1), 1–11.

Cwiklinski, K., Dalton, J. P., Dufresne, P. J., La Course, J., Williams, D. J., Hodgkinson, J., & Paterson, S. (2015). The *Fasciola hepatica* genome: Gene duplication and polymorphism reveals adaptation to the host environment and the capacity for rapid evolution. *Genome Biology*, *16*(1), 1–13.

Fuqua, S. A., Oesterreich, S., Hilsenbeck, S. G., Von Hoff, D. D., Eckardt, J., & Osborne, C. K. (1994). Heat shock proteins and drug resistance. *Breast Cancer Research and Treatment*, *32*(1), 67–71.

Hall, T. (Ed.) (1999). BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series.

Herrera, G., Hernández, C., Ayala, M. S., Flórez, C., Teherán, A. A., & Ramírez, J. D. (2017). Evaluation of a multilocus sequence typing (MLST) scheme for *Leishmania* (Viannia) *braziliensis* and *Leishmania* (Viannia) *panamensis* in Colombia. *Parasites & Vectors*, *10*(1), 1–10.

Heydarian, P., Jajarmi, V., Spotin, A., Ashrafi, K., Mohebali, M., Aryaeipour, M., Bozorgomid, A., Hajialilo, E., Javad Abbaszadeh Afshar, M., Fadaei Tehrani, M., & Bagher Rokni, M. (2022). Molecular characterization of animal *Fasciola* spp. isolates from Lorestan Province, Western Iran. *Iranian Journal of Public Health*, *51*(8), 1847–1856.

Hunter, P. R., & Gaston, M. A. (1988). Numerical index of the discriminatory ability of typing systems: An application of Simpson's index of diversity. *Journal of Clinical Microbiology*, *26*(11), 2465–2466.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*(2), 254–267.

Itagaki, T., Jin, S., Takashima, S., Ohari, Y., & Angeles, J. M. M. (2022). Molecular analyses confirm the coexistence of *Fasciola gigantica* and parthenogenetic *Fasciola* in the Philippines. *Parasitology International*, *88*, 102562.

Kasahara, S., Ohari, Y., Jin, S., Calvopina, M., Takagi, H., Sugiyama, H., & Itagaki, T. (2021). Molecular characterization revealed *Fasciola* specimens in Ecuador are all *Fasciola hepatica*, none at all of *Fasciola gigantica* or parthenogenic *Fasciola* species. *Parasitology International*, *80*, 102215.

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547.

Lalor, R., Cwiklinski, K., Calvani, N. E. D., Dorey, A., Hamon, S., Corrales, J. L., Dalton, J. P., & De Marco Verissimo, C. (2021). Pathogenicity and

virulence of the liver flukes *Fasciola hepatica* and *Fasciola gigantica* that cause the zoonosis Fasciolosis. *Virulence*, *12*(1), 2839–2867.

Leigh, J. W., & Bryant, D. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6*(9), 1110–1116.

Li, S., Zou, Y., Wang, P., Qu, M. -R., Zheng, W. -B., Wang, P., Chen, X. Q., & Zhu, X. Q. (2021). Prevalence and multilocus genotyping of *Cryptosporidium* spp. in cattle in Jiangxi Province, southeastern China. *Parasitology Research*, *120*(4), 1281–1289.

Maiden, M. C., Van Rensburg, M. J. J., Bray, J. E., Earle, S. G., Ford, S. A., Jolley, K. A., & McCarthy, N. D. (2013). MLST revisited: The gene-by-gene approach to bacterial genomics. *Nature Reviews Microbiology*, *11*(10), 728–736.

Malatyali, E., Guclu, O., Yildiz, I., Bozdogan, B., Ertug, S., & Ertabaklar, H. (2020). Molecular characterisation of *Trichomonas vagi*nalis isolates in Southwest Turkey with multilocus sequence typing and genetic structure analysis in relation to different countries. *Infection, Genetics and Evolution*, *84*, 104459.

Mas-Coma, S., Valero, M., & Bargues, M. (2019). Fascioliasis. *Advances in Experimental Medicine and Biology*, *1154*, 71–103.

Mirzaei, M., Ghahvei, Y., Lefoulon, E., Lia, R. P., Otranto, D., Martin, C., & Sazmand, A. (2018). Morphological and molecular characterization of *Onchocerca fasciata* (Nematoda, Onchocercidae) from dromedary camels (*Camelus dromedarius*) in Iran. *Parasite*, **25**.

Moreira-de-Sousa, C., de Souza, R. B., & Fontanetti, C. S. (2018). HSP70 as a biomarker: An excellent tool in environmental contamination analysis—A review. *Water, Air, & Soil Pollution*, *229*(8), 1–12.

Nemati, S., Fazaeli, A., Hajjaran, H., Khamesipour, A., Anbaran, M. F., Bozorgomid, A., & Zarei, F. (2017). Genetic diversity and phylogenetic analysis of the Iranian Leishmania parasites based on HSP70 gene PCR-RFLP and sequence analysis. *The Korean Journal of Parasitology*, *55*(4), 367.

Rokni, M. B., Bashiri, H., Raeghi, S., Teimouri, A., Shojaeimotlagh, V., Shiee, M. R., & Bozorgomid, A. (2020). Molecular phylogenetic and genetic variability of *Fasciola gigantica* in Kermanshah province, western Iran with an overview to understand haplotypes distribution in Asia and Africa. *Veterinary Research Forum: An International Quarterly Journal*, *11*(3), 265–271.

Rokni, M. B., Bozorgomid, A., Heydarian, P., & Aryaeipour, M. (2018). Molecular evidence of human fasciolosis due to *Fasciola gigantica* in Iran: A case report. *Iranian Journal of Public Health*, *47*(5), 750.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, *34*(12), 3299–302.

Saadatnia, A., Solhjoo, K., Davami, M. H., Raeghi, S., & Abolghazi, A. (2022). Molecular identification of *Fasciola* isolated from the liver of meat animals in Fars Province, Iran. *Journal of Parasitology Research*, *2022*, 4291230.

Shafiei, R., Sarkari, B., Sadjjadi, S. M., & Mowlavi, G. R., & Moshfe, A. (2014). Molecular and morphological characterization of *Fasciola* spp. isolated from different host species in a newly emerging focus of human fascioliasis in Iran. *Veterinary Medicine International*, *2014*, 405740.

Squire, S. A., & Ryan, U. (2017). Cryptosporidium and Giardia in Africa: Current and future challenges. *Parasites & Vectors*, *10*(1), 1–32.

Thang, T. N., Vázquez-Prieto, S., Vilas, R., Paniagua, E., Ubeira, F. M., & Ichikawa-Seki, M. (2020). Genetic diversity of *Fasciola hepatica* in Spain and Peru. *Parasitology International*, *76*, 102100.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Nazari, N., Rokni, M. B., Ichikawa-Seki, M., Raeghi, S., Hajjaran, H., Falahi, S., Hamzavi, Y., Heydarian, P., Davari, A., Ghadiri, K., & Bozorgomid, A. (2023). Assessment of genetic markers for multilocus sequence typing (MLST) of *Fasciola* isolates from Iran. *Veterinary Medicine and Science*, *9*, 924–933. https://doi.org/10.1002/vms3.995