Pakistan Veterinary Journal

ISSN: 0253-8318 (PRINT), 2074-7764 (ONLINE)
DOI: 10.29261/pakvetj/2025.153

RESEARCH ARTICLE

Global Scale Insights into Genetic Variation in Mitochondrial *cox*1 and *nad*1 Genes of *Clonorchis sinensis*

Shahbaz ul Haq[1¶], Talha Javaid[2¶], Majed H. Wakid[3,4], Muhammad Wasim Usmani[5], Muhammad Saqib[2], Muhammad Abdullah Malik[6], Shujaat Hussain[7], Safia Obaidur Rab[8], Mohd Saeed[9], Asif Ali Butt[10], Fenfei Gao[1*], Hong-Bin Yan[11*], Li Li[11] and Mughees Aizaz Alvi[2*]

[1]Department of Pharmacology, Shantou University Medical College, Shantou, Guangdong 51504, China; [2]Department of Clinical Medicine and Surgery, University of Agriculture, Faisalabad, Pakistan; [3]Department of Medical Laboratory Sciences, Faculty of Applied Medical Sciences, King Abdulaziz University, Jeddah 21589, Saudi Arabia; [4]Special Infectious Agents Unit, King Fahd Medical Research Center, King Abdulaziz University, Jeddah, Saudi Arabia; [5]Institute of Drug Discovery Technology, Ningbo University, Ningbo, Zhejiang, 315211, China; [6]Department of Parasitology, University of Agriculture, Faisalabad, Pakistan; [7]Faculty of Veterinary and Animal Sciences, PMAS Arid Agriculture University, Rawalpindi, Pakistan; [8]Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King Khalid University, Abha, Saudi Arabia; [9]Department of Biology, College of Science, University of Hail, Hail, Saudi Arabia; [10]Riphah International University, Faisalabad Campus, Faisalabad, Pakistan; [11]State Key Laboratory for Animal Disease Control and Prevention, College of Veterinary Medicine, Lanzhou University, Gansu Province Research Center for Basic Disciplines of Pathogen Biology, Key Laboratory of Veterinary Parasitology of Gansu Province, Key Laboratory of Veterinary Etiological Biology and Key Laboratory of Ruminant Disease Prevention and Control (West), Ministry of Agricultural and Rural Affairs, National Para-reference Laboratory for Animal Echinococcosis, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Lanzhou 730046, P.R. China.
*Corresponding author: ffgao@stu.edu.cn; yanhongbin@caas.cn; mugheesaizazalvi@gmail.com
¶These are the first co-authors and contributed equally to this study.

ARTICLE HISTORY (25-171)

Received: February 25, 2025
Revised: April 3, 2025
Accepted: April 7, 2025
Published online: April 23, 2025

Key words:
Diversity indices
Epidemiology
Neutrality indices
Population structure
Haplotypes
Clonorchis sinensis

ABSTRACT

Clonorchiasis is a food-borne zoonoses caused by *Clonorchis* (*C.*) *sinensis*. Human infection occurs through consumption of raw or undercooked fish containing metacercaria, an infective stage that can persist for 20 years, leading to chronic diseases. Few studies have been conducted on *C. sinensis* based on population structure, and even fewer have been conducted on molecular characterization. The goal of the current study was to provide epidemiological information on the global population structure of the parasite. After obtaining the *cox*1 and *nad*1 genes sequences from the NCBI GenBank database, median-joining networks were constructed using PopArt software. The neutrality and diversity indexes were computed using the DnaSp software. We also calculated neutrality and diversity indices based on host and geographic location. In the NCBI database, there were 209 *nad*1 and 275 *cox*1 gene sequences. After removing short gene sequences, bioinformatic analysis was performed using 251 *cox*1 and 189 *nad*1 gene sequences. The *cox*1 and *nad*1 genes were shown to have 40 and 42 parsimony informative sites, 87 and 73 haplotypes, and 82 mutations, respectively. We found negative and statistically significant Tajima's D (*cox*1=-2.41336 and *nad*1=-2.50463), Fu's Fs values were negative but did not reach statistical significance (*cox*1=-32.908, *nad*1=-107.788), and high diversity (*cox*1 Hd=0.948, π=0.00317; *nad*1 Hd=0.9192, π=0.00293). South Korean isolates exhibited the highest haplotype diversity for both genes, with Chinese and Russian isolates following closely behind. A statistically significant overall negative Tajima's D value was found for both genes based on host-wise data analysis. According to the study's findings, high diversity and negative neutrality indices indicate that *C. sinensis* population diversity will increase globally. The results of the investigation will contribute to the current understanding of the population structure of *C. sinensis*, for which there is a dearth of data. To have a better understanding of the parasite's epidemiology, we recommend conducting more research using a number of genetic markers.

To Cite This Article: Haq SU, Javaid T, Wakid MH, Usmani MW, Saqib M, Malik MA, Hussain S, Rab SO, Saeed M, Butt AA, Gao F, Yan HB, Li L and Alvi MA, 2025. Global Scale Insights into Genetic Variation in Mitochondrial *cox*1 and *nad*1 Genes of *Clonorchis sinensis*. Pak Vet J. http://dx.doi.org/10.29261/pakvetj/2025.153

INTRODUCTION

The Chinese liver fluke, or *Clonorchis (C.) sinensis*, is a member of the family Opisthorchiidae, order Opisthochiida, and class Trematoda (Qian and Zhou, 2021). This food-borne zoonotic disease affects both humans and animals. The illness is endemic in Southeast Asia and is referred to as clonorchiasis (Zhang *et al*., 2020). The parasite damages the liver and bile ducts, and overall host poisoning can result in a variety of illnesses in humans, including the potentially fatal cholangiocarcinoma (Sota *et al*., 2024).

This parasite has recently been categorized as a member of Group 1 biological agents, which are significant for public health since they are carcinogenic (Kim *et al*., 2022). The life cycle of this involves definitive hosts, such as dogs and cats, and intermediate hosts, including freshwater snails, fish, and humans. Human infections are caused by eating raw or undercooked fish that contains metacercariae, an infectious stage of the parasite that can persist for 20 years and cause chronic morbidities (Qian and Zhou, 2021). Hepatobiliary symptoms, such as cholangiocarcinoma, cholelangitis, cholecystitis, and cholelithiasis, are indicative of the disease (Kinkar *et al*., 2023).

The gold standard for diagnosis is egg counting in feces, while molecular and serological methods are also employed. As an alternative, mebendazole and praziquantel are used to treat the disease chemotherapeutically at both the mass and individual levels (Sota *et al*., 2024). The availability of reservoir hosts, the presence of the parasite's first intermediate host (freshwater snail), eating raw or undercooked fish, travel, population migration from endemic areas to other areas, growing interest in fast food, and the introduction of intermediate hosts by humans and animals are some of the factors that contribute to the parasite's endemicity (Xu *et al*., 2021).

Given the importance and endemicity of this parasite, research into the geographic distribution and population structure of *C. sinensis* is crucial for veterinary and medical health. Very little is known about the global geographic distribution and genetic variation within and across species of *C. sinensis* based on mitochondrial DNA (mtDNA) genes (Qian *et al*., 2020; Kim *et al*., 2021; Qian and Zhou, 2021; Young *et al*., 2021; Kinkar *et al*., 2023; Tantrawatpan *et al*., 2023; Wu *et al*., 2023; Nguyen *et al*., 2024). Generally, mtDNA markers are employed to research population genetics, phylogenetic association, haplotype diversity, and population structure.

Phylogeographic analyses are a crucial technique for reconstructing the historical connection of the species and for elucidating the processes that govern the parasite's genetic linages, population structure, and geographic dispersion (Wu *et al*., 2023). Despite the fact that many animals have been the subject of phylogeographic and phylogenetic studies, research on the parasite population, particularly that of helminths, has been reported to be insufficient or nonexistent globally (Wang *et al*., 2018; Kim *et al*., 2019; Shi *et al*., 2020; Xu *et al*., 2021; Li *et al*., 2022).

An international assessment of individual molecular and genetic variation is necessary to comprehend the dynamics of the Chinese liver fluke. Through investigations of mitochondrial *cox*1 and *nad*1 genes sequences submitted to the GenBank from four different nations, the genetic diversity, haplotype diversity, and population structure of *C. sinensis* were examined in the current study. The results of the study will represent the first attempt to look into and comprehend the genetic diversity and population structure of *C. sinensis* on a worldwide scale.

MATERIALS AND METHODS

Information gathering and synchronization: After being downloaded from the National Center for Biotechnology Information, USA, (NCBI), 275 *cox*1 and 209 *nad*1 genes sequences of *C. sinensis* were combined into a dataset. The most recent search was conducted on December 31, 2024. Both genes' sequences (*cox*1 n=275 and *nad*1 n=209) were put together in FASTA format using Molecular Evolutionary Genetics Analysis Version 11 (MEGA 11). The *cox*1 and *nad*1 sequences were trimmed off at both ends and aligned using the whole genome reference sequences (NC012147). Following sequence trimming, 251 *cox*1 and 189 *nad*1 gene sequences were used for bioinformatic analysis.

Analysis of haplotypes and networking: Haplotypes were analyzed using the DnaSP6 application program, and the sequences were examined in FASTA format (Rozas *et al*., 2017). Neutrality indexes, nucleotide and haplotype numbers, and nucleotide change values were acquired in order to determine the genes' genetic structure (Maddison *et al*., 1997). The association between haplotypes was illustrated using a haplotype network made with the Population Analysis with Reticulate Trees program software (Leigh *et al*., 2015).

RESULTS

Our current study examined the gene sequences of 251 *cox*1 and 189 *nad*1 isolates of *C. sinensis* that were retrieved from the NCBI database. The sequences of the *cox*1 and *nad*1 genes are listed in Table 1 and Table 2, respectively.

Research on haplotypes and polymorphism: There were 87 distinct haplotypes and 82 distinct locations where mutations were found in the *cox*1 gene sequences. Hap 05 was the predominant haplotype among them, including 45 gene sequences. Only one sequence was present in each of the fifty-seven haplotypes (Table 3). 73 haplotypes were produced by 82 locations of mutations in the *nad*1 gene. Hap-04, which has 49 sequences, was the predominant haplotype among these (Table 4).

The haplotype network: Hap 05 accounted for 17.92% (45/251) of the haplotype network, making it the most prevalent haplotype. Hap11 came in second with 9.96% (25/251). A distinct single haplotype made up 22.70% (57/251). The predominant haplotype for the *nad*1 gene, Hap-4, accounted for 25.92% (49/189) of the network, with Hap-09 and Hap-14 following closely behind. These haplotypes made up 7.40% (14/189) of the network. A distinct single haplotype made for 25.92% (49/189).

Table 1: Accession numbers for *cox*1 gene fragments from *C. sinensis* isolates used in the research.

Country	Isolates	Accession Numbers
Vietnam	8	OM810331,OM810330, OM810329, OM810328, OM810327, OM810326, OM810325, OM810324
South Korea	2	MT607652, JF729304
Russia	64	MF406176 to MF406206, MT292281 to MT292292, MN116457 to MN116479
China	172	JF729303 to MT292280
Origin information not available	5	FJ381664, MF406175, MN116463, MN116464, NC012147

Table 2: Accession numbers of *nad*1 gene fragments of *C. sinensis* isolates used in the study.

Country	Isolates	Accession Numbers
South Korea	2	MT607652, KY564177
Russia	12	MT292281 to MT292292
China	173	JF729303, JF729304, MT292110 to MT292280
Origin information not available	2	FJ381664, NC012147

Table 3: Haplotypes of *cox*1 sequences of *C. sinensis* and accession numbers of isolates forming groups.

Haplotypes	Isolates	Haplotype No.
1	1	OM810331
2	14	OM810330, OM810329, OM810328, MT292252, MT292247, MT292183, MT292167, MT292162, MT292119, MT292116, MN116477, MN116476, MN116465, MF406202
3	1	OM810327
4	4	OM810326, OM810325, OM810324, MT292238
5	45	NC012147, MT292284, MT292281, MT292263, MT292250, MT292248, MT292245, MT292243, MT292227, MT292225, MT292221, MT292219, MT292214, MT292213, MT292211, MT292206, MT292203, MT292200, MT292198, MT292199, MT292192, MT292185, MT292182, MT292177, MT292170, MT292164, MT292159, MT292128, MT292127, MT292126, MT292113, MT292111, MT292110, MN116475, MN116470, MN116466, MN116463, MN116457, MF406205, MF406203, MF406201, MF406179, MF406178, MF406175, FJ381664
6	3	MT607652, MT292258, MT292254
7	8	MT292292, MT292274, MT292262, MT292260, MN116467, MF406204, MF406183, MF406180
8	4	MT292291, MT292289, MT292288, MT292287
9	2	MT292290, MT292139
10	1	MT292286
11	25	MT292285, MT292265, MT292175, MN116478, MN116464, MN116462, MN116461, MN116460, MN116459, MF406200, MF406199, MF406198, MF406197, MF406196, MF406195, MF406194, MF406193, MF406192, MF406191, MF406190, MF406189, MF406188, MF406185, MF40617, MF406176
12	4	MT292283, MT292282, MF406182, MF406181
13	1	MT292280
14	4	MT292279, MT292267, MT292205, MT292174
15	3	MT292278, MT292271, MF406184
16	2	MT292277, MT292193
17	4	MT292276, MT292268, MT292223, MT292210
18	1	MT292275
19	1	MT292273
20	6	MT292272, MT292216, MT292207, MT292121, MT292120, MT292114
21	6	MT292270, MT292226, MT292208, MT292204, MT292131, MT292112
22	2	MT292269, MT292134
23	1	MT292266
24	1	MT292264
25	7	MT292261, MT292181, MT292152, MT292149, MT292136, MT292135, MT292124
26	1	MT292259
27	1	MT292257
28	1	MT292256
29	1	MT292255
30	2	MT292253, MT292151
31	1	MT292251
32	1	MT292249
33	1	MT292246
34	1	MT292244
35	1	MT292242
36	1	MT292241
37	7	MT292240, MT292237, MT292229, MT292224, MT292176, MN116472, MN116471
38	1	MT292239
39	2	MT292236, MT292217
40	3	MT292235, MT292195, MT292154
41	1	MT292234
42	2	MT292233, MT292157
43	1	MT292232
44	4	MT292231, MT292197, MT292144, MT292140
45	5	MT292230, MT292196, MT292190, MT292118, MT292117
46	1	MT292228
47	7	MT292222, MT292218, MT292215, MT292173, MT292150, MT292138, MT292133
48	1	MT292220
49	1	MT292212
50	1	MT292209
51	1	MT292202
52	1	MT292201
53	1	MT292194

54	1	MT292191
55	11	MT292188, MT292187, MT292184, MT292178, MT292169, MT292160, MT292158, MT292145, MT292130, MT292123, MT292115
56	1	MT292189
57	1	MT292186
58	1	MT292180
59	1	MT292179
60	1	MT292171
61	1	MT292172
62	1	MT292168
63	2	MT292166, MT292165
64	1	MT292163
65	1	MT292161
66	1	MT292156
67	2	MT292155, MT292122
68	1	MT292153
69	1	MT292148
70	1	MT292148
71	1	MT292146
72	1	MT292143
73	1	MT292142
74	1	MT292141
75	1	MT292137
76	1	MT292132
77	1	MT292129
78	1	MT292125
79	1	MN116479
80	1	MN116474
81	1	MN116473
82	1	MN116469
83	1	MN116468
84	2	MN116458, MF406206
85	1	MF406187
86	1	MF406186
87	2	JF729304, JF729303

Table 4: Haplotypes of *nad1* sequences of *C. sinensis* and accession numbers of isolates forming groups.

Haplotypes	Isolates	Haplotype No.
1	3	MT607652, JF729304, MT292258
2	1	JF729303
3	1	KY564177
4	49	FJ381664, NC012147, MT292286, MT292113, MT292126, MT292127, MT292128, MT292129, MT292131, MT292134, MT292136, MT292139, MT292148, MT292154, MT292159, MT292162, MT292179, MT292182, MT292185, MT292189, MT292195, MT292200, MT292201, MT292203, MT292204, MT292206, MT292208, MT292209, MT292220, MT292226, MT292234, MT292235, MT292239, MT292244, MT292246, MT292249, MT292250, MT292259, MT292267, MT292268, MT292269, MT292270, MT292271, MT292276, MT292278, MT292279g, MT292290, MT292178, MT292261
5	1	MT292223
6	1	MT292110
7	3	MT292111, MT292168, MT292219
8	1	MT292112
9	14	MT292114, MT292120, MT292138, MT292150, MT292173, MT292207, MT292215, MT292216, MT292218 ,MT292222, MT292272, MT292275, MT292281, MT292284
10	1	MT292115
11	4	MT292116, MT292119, MT292183, MT292247
12	3	MT292117, MT292118, MT292190
13	1	MT292121
14	14	MT292122, MT292124, MT292130, MT292152, MT292155, MT292172, MT292181, MT292187, MT292188, MT292212, MT292287, MT292288, MT292289, MT292291
15	1	MT292123
16	1	MT292125
17	2	MT292132, MT292228
18	1	MT292133
19	1	MT292137
20	3	MT292140, MT292147, MT292156
21	1	MT292141
22	1	MT292142
23	1	MT292143
24	1	MT292144
25	1	MT292145
26	1	MT292146
27	1	MT292149
28	2	MT292151, MT292253
29	3	MT292153, MT292169, MT292237
30	1	MT292157
31	1	MT292158
32	1	MT292160
33	1	MT292161
34	3	MT292163, MT292199, MT292248
35	7	MT292164, MT292171, MT292193, MT292211, MT292225, MT292227, MT292277

36	2	MT292165, MT292166
37	1	MT292167
38	1	MT292170
39	2	MT292174, MT292205
40	3	MT292175, MT292265, MT292285
41	4	MT292176, MT292224, MT292229, MT292240
42	1	MT292177
43	1	MT292180
44	1	MT292184
45	1	MT292186
46	1	MT292191
47	1	MT292192
48	1	MT292194
49	2	MT292196, MT292230
50	2	MT292197, MT292231
51	4	MT292198, MT292213, MT292243, MT292263
52	1	MT292202
53	1	MT292210
54	1	MT292214
55	2	MT292217, MT292236
56	1	MT292221
57	1	MT292232
58	1	MT292233
59	1	MT292238
60	1	MT292241
61	1	MT292242
62	1	MT292245
63	1	MT292251
64	1	MT292252
65	1	MT292254
66	3	MT292255, MT292256, MT292257
67	4	MT292260, MT292262, MT292274, MT292292
68	1	MT292264
69	1	MT292266
70	1	MT292273
71	1	MT292280
72	2	MT292282, MT292283
73	1	MT292135

Table 5: Diversity and neutrality indices obtained using nucleotide data of the *cox*1 (843 bp) gene sequences of *C. sinensis*

Indices	*nad1 (*902 base pairs*)*	
Total sequences	251	189
Total mutations	82	82
Parsimony sites	40	42
Haplotypes	87	73
Hd	0.948	0.9192
Π	0.00317	0.00293
Tajima's *D*	-2.41336*	-2.50463*
Fu's Fs	-32.908	-107.788
Fu and Li's D value	-5.34428	-4.49238
Fu and Li's F value	-4.65087	-4.18730

*indicates P<0.05.

Analysis of gene flow, diversity and neutrality: Table 5 displays the *cox*1 and *nad*1 gene sequences' diversity and neutrality markers. The values of Fu's FS (Fu, 1997) and Tajima's D (Tajima, 1989) were calculated to determine whether or not populations were influenced by selection. Tajima's D, Fu's Fs, and Fu's LD values for the *cox*1 gene were highly negative, indicating a large number of alleles. Tajima's D value for the *nad*1 gene was negative and statistically significant (-2.50463*), similar to that of the *cox*1 gene, but Fu's Fs was negative but not significant (-107.788). The results of the D and F test statistics for Fu and Li were likewise negative, just as the *cox*1 gene.

The neutrality and diversity indices in various nations were also ascertained by analyzing the sequences of the *cox*1 and *nad*1 genes that were part of the study. The polymorphism study pertaining to the future population increase of *C. sinensis* in various nations is provided in Table 6. The sequences deposited in the NCBI repository lacked country information when the median joining network was being constructed. Consequently, such sequences were eliminated throughout the median-joining network construction process, and a new file was created. Fig. 1 and 2 display the median-joining network of the *cox*1 and *nad*1 haplotype data by nation.

Tajima's D values were positive for the Vitenamese isolates and negative and significant for the Chinese *cox*1 isolates (Table 6). Since there were less than four isolates in South Korea, Tajima's D was not calculated, although it was negative but negligible for Russia. As may be predicted from genetic hitchhiking or the recent population growth, Fu's Fs values were negative for Chinese and Russian isolates but positive for South Korean and Vietnamese isolates.

Tajima's D values for the Chinese *nad*1 isolates were substantial and negative (Table 7). Since there were less than four isolates in South Korea, Tajima's D was not calculated, although it was negative but negligible for Russia. China and Russia have negative Fu's FS values, which was to be predicted given genetic hitchhiking or the recent population growth.

185 *nad*1 and 244 *cox*1 gene sequences were used for host-wise analysis of diversity and neutrality indices. The analyses may be found in Tables 8 and 9. Overall, the Tajima's D value was negative and statistically significant. Analysis by host revealed that isolates from dogs and cats had considerably negative Tajima's D values for both genes. In all hosts, the values of Fu's Fs for the diversity and neutrality indices of the *cox*1 and *nad*1 gene isolates were found to be largely negative but negligible. The host-wise *cox*1 and *nad*1 haplotype data's median-joining network is shown in Fig. 3 and 4, respectively.

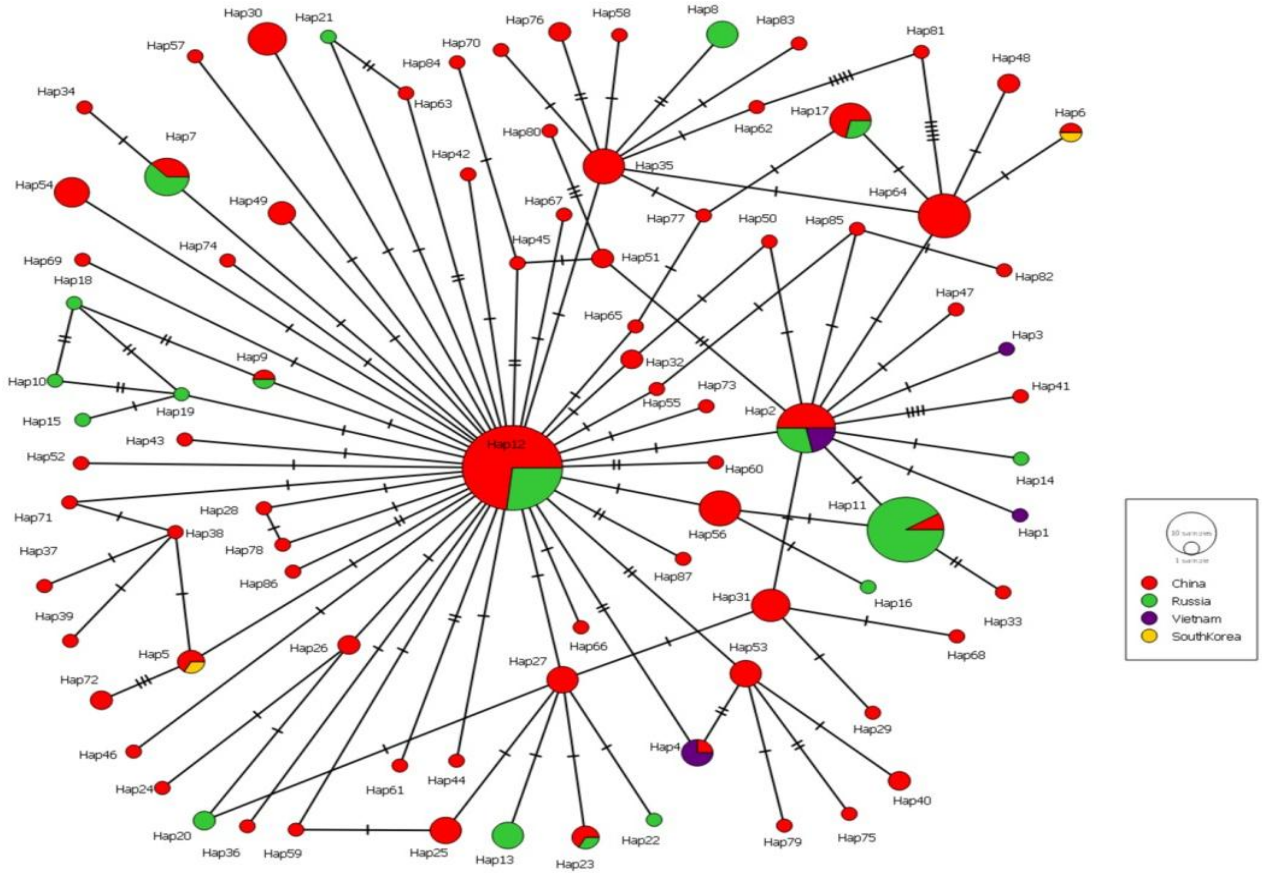

Fig. 1: *Clonorchis sinensis* haplotypes as a region-wise median-joining network based on *cox*1 gene sequences (843 bp). The haplotype frequencies are proportionate to the circle sizes. Mutations are indicated by hatch marks.

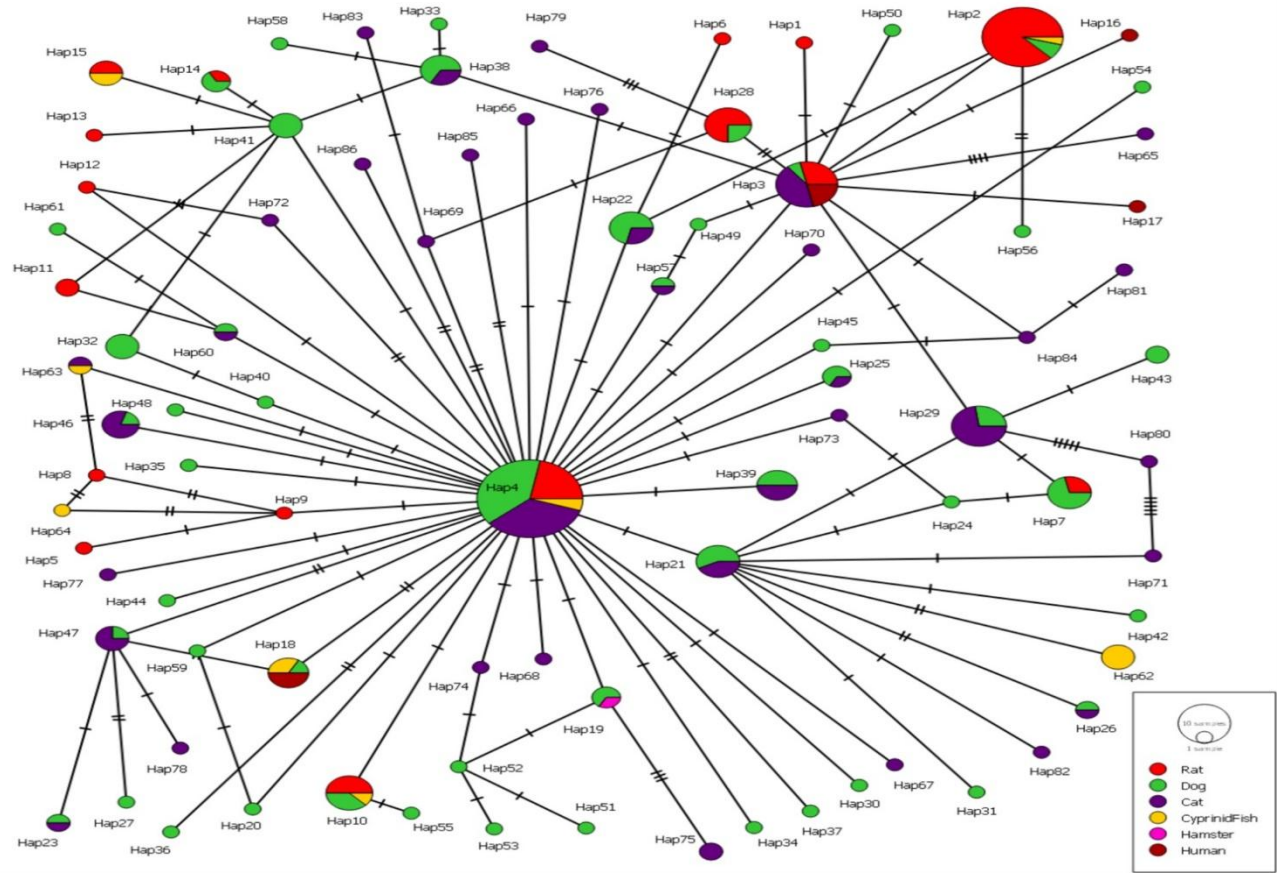

Fig. 2: *Clonorchis sinensis* haplotypes as a host-wise median-joining network based on *cox*1 gene sequences (843 bp). The haplotype frequencies are proportionate to the circle sizes. Mutations are indicated by hatch marks.

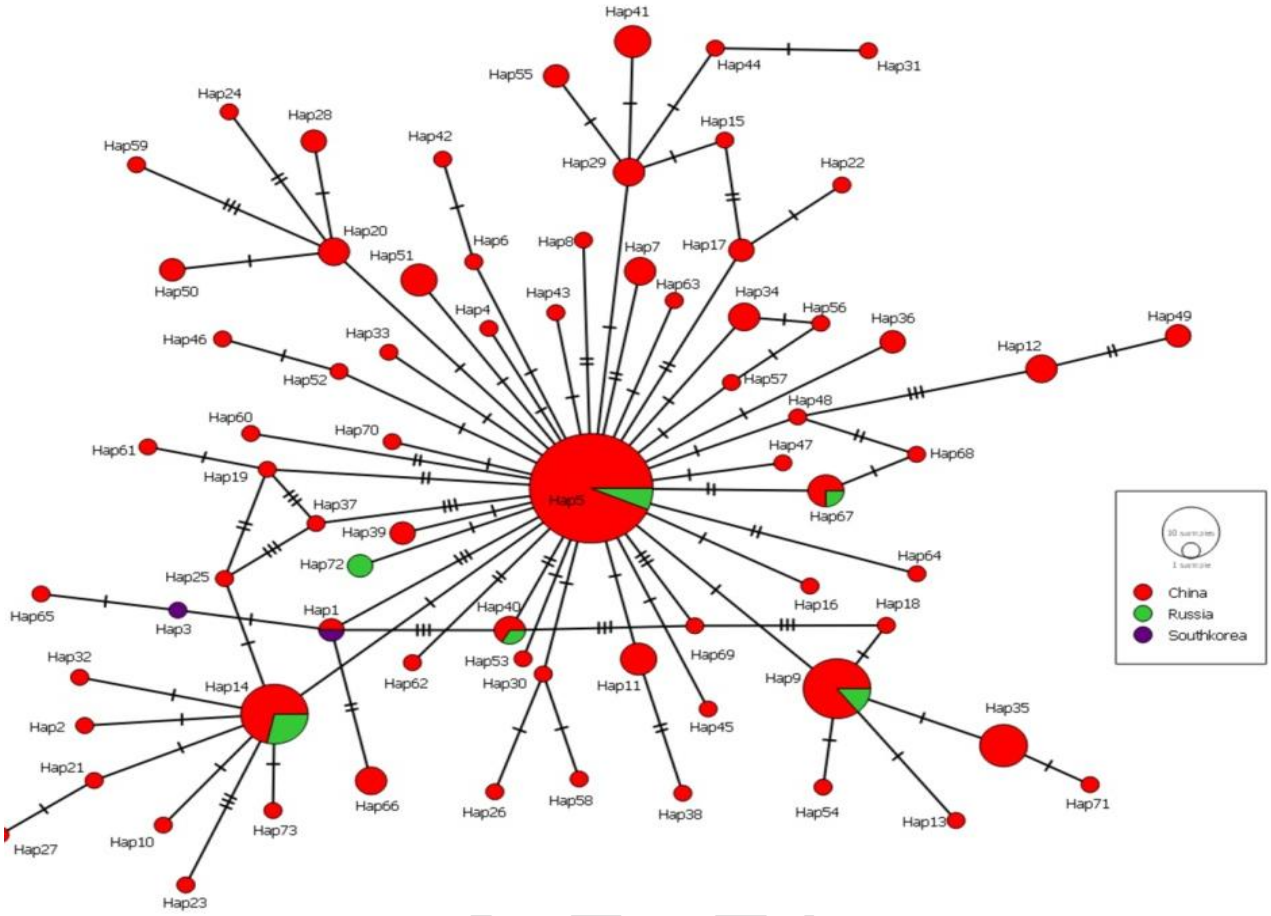

Fig. 3: *Clonorchis sinensis* haplotypes as a region-wise median-joining network based on *nad*1 gene sequences (902 bp). The haplotype frequencies are proportionate to the circle sizes. Mutations are indicated by hatch marks.

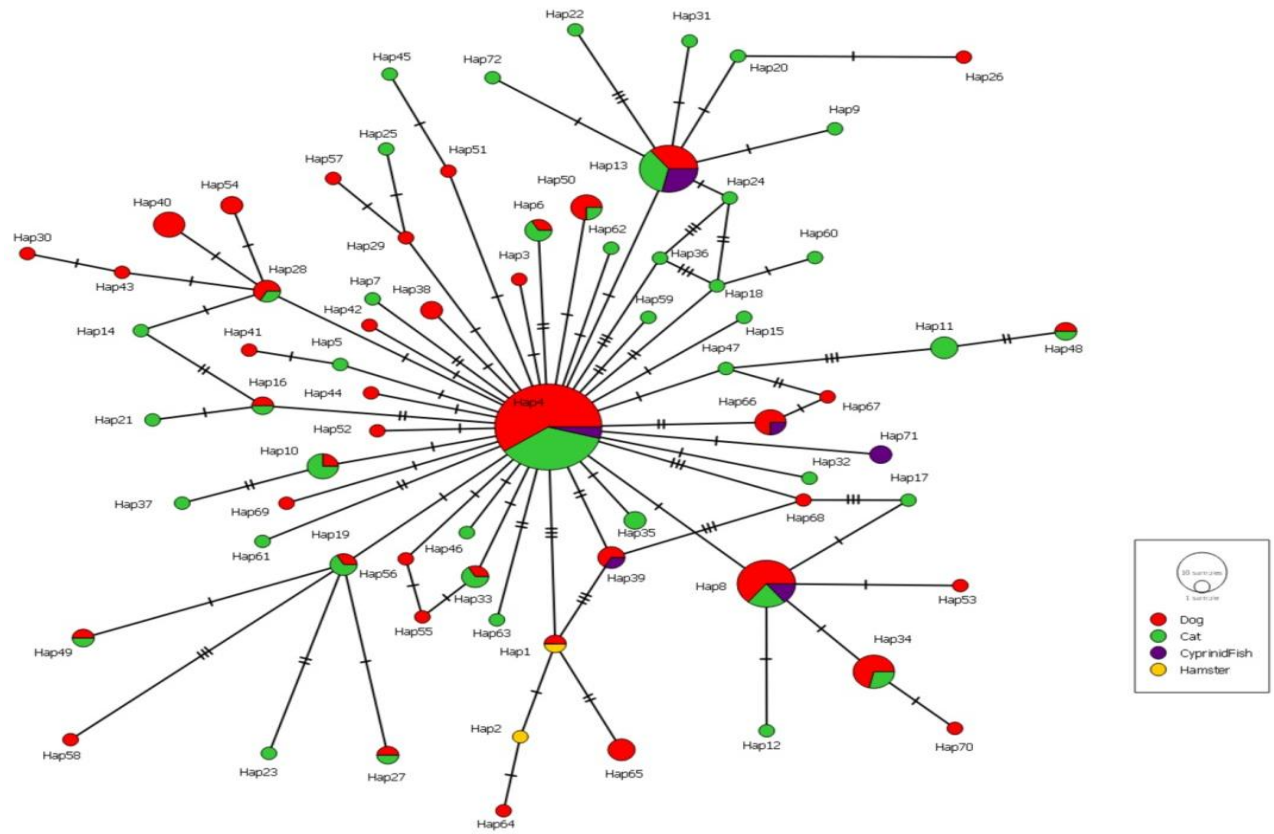

Fig. 4: *Clonorchis sinensis* haplotypes as a host-wise median-joining network based on *nad*1 gene sequences (902 bp). The haplotype frequencies are proportionate to the circle sizes. Mutations are indicated by hatch marks.

Table 6: Region-wise diversity and neutrality indices of *cox*1 gene sequences from the NCBI GenBank database

	Total sequences	Total mutations	Parsimony Sites	Haplotypes	Hd	π	Tajima's D	Fu's FS	Fu and Li's D value	Fu and Li's F value
China	172	72	32	74	0.957	0.00323	-2.41727*	-112.278	-5.36173	-4.71787
Russia	64	19	14	18	0.843	0.00291	-1.18963	-7.857	-0.37925	-0.75395
South Korea	2	4	0	2	1.000	0.00474	---	1.386	----	---
Vietnam	8	5	3	4	0.786	0.00250	0.42038	0.261	0.12651	0.19719

* indicates P<0.05.

Table 7: Region-wise diversity and neutrality indices obtained using nucleotide data of the *nad*1 gene sequences of *C. sinensis* obtained from the NCBI GenBank database.

	Total sequences	Total mutations	Parsimony sites	Haplotypes	Hd	π	Tajima's D	Fu's FS	Fu and Li's D value	Fu and Li's F value
China	173	81	40	71	0.921	0.00297	-2.51487*	-104.507	-4.59287	-4.26214
Russia	12	7	3	6	0.864	0.00195	-0.94609	-1.676	-0.78603	-0.85445
South Korea	2	1	0	2	1.000	0.00111	-	0.000	-	-

* indicates statistical significance with P<0.05, 185 *nad*1 and 244 *cox*1 gene sequences were used for host-wise analysis of diversity and neutrality indices. The analyses may be found in Tables 8 and 9. Overall, the Tajima's D value was negative and statistically significant. Analysis by host revealed that isolates from dogs and cats had considerably negative Tajima's D values for both genes. In all hosts, the values of Fu's Fs for the diversity and neutrality indices of the *cox*1 and *nad*1 gene isolates were found to be largely negative but negligible. The host-wise *cox*1 and *nad*1 haplotype data's median-joining network is shown in Fig. 3 and 4, respectively.

Table 8: Host-wise diversity and neutrality indices obtained using nucleotide data of the *cox*1 gene sequences of *C. sinensis* obtained from the NCBI GenBank database.

Indices	Cat	Cyprinid fish	Dog	Hamster	Homo sapiens	Rat	Overall
Total sequences	74	12	97	1	8	52	244
Total mutations	46	12	46	-	5	16	81
Parsimony sites	18	5	20	-	3	10	39
Haplotypes	37	7	50	-	4	15	86
Hd	0.944	0.879	0.962	-	0.786	0.803	0.952
Π	0.00323	0.00383	0.00319	-	0.00250	0.00246	0.00320
Tajima's *D*	-2.32416*	-0.78908	-2.21512*	-	0.42038	-1.27252	-2.4068*
Fu's Fs	-34.172	-1.163	-65.816	-	0.261	-6.635	-33.336
Fu and Li's D value	-3.99663	-0.90334	-4.00439	-	0.12651	-1.05927	-5.38114
Fu and Li's F value	-3.81084	-0.90379	-3.76983	-	0.19719	-1.26580	-4.67453

* indicates P<0.05.

Table 9: Host-wise diversity and neutrality indices obtained using nucleotide data of the *nad*1 gene sequences of *C. sinensis* obtained from the NCBI GenBank database

Indices	Cat	Cyprinid fish	Dog	Hamster	Overall
Total sequences	74	12	97	2	185
Total mutations	53	7	53	1	81
Parsimony sites	20	3	22	0	42
Haplotypes	43	6	41	2	72
Hd	0.941	0.864	0.903	1.000	0.921
π	0.00305	0.00195	0.00283	0.00111	0.00293
Tajima's *D*	-2.46211*	-0.94609	-2.40768*	-	-2.50123*
Fu's Fs	-34.252	-1.676	-33.904	0.000	-105.626
Fu and Li's D value	-4.21188	-0.78603	-4.12466	-	-4.34520
Fu and Li's F value	-4.02456	-0.85445	-3.93552	-	-4.09359

* indicates P<0.05.

DISCUSSION

The ongoing worldwide climate shift and significant human migratory movements present opportunities for human parasites to spread to new regions. This makes knowing their genetic diversity and phylogeography more important. Genetic data is essential for understanding the biological history, current condition, and basis for treatment, control, and prognostication of infestations of parasite species (Chelomina *et al*. 2014). Parasitic diseases significantly impact the livestock industry, leading to economic losses (Mahmood *et al*., 2022; Mohammed *et al*., 2022; Qamar *et al*., 2022; Rafique *et al*., 2022;). These could include zoonotic manifestations and production losses, which bring about financial losses to livestock, agriculture, and eventually the national GDP.

A key defining characteristic is genetic variety, which governs the longevity of parasite diversity and survival. The ITS1-5.8S-ITS2 region has traditionally been the primary genetic marker for studies on this parasite (Tatonova *et al*., 2012). Park and Yong (2001), Lee and Huh (2004), Park (2007), Liu *et al*. (2012), and Tatonova *et al*. (2013) employed the partial ITS sections of the nuclear DNA (nDNA) and mtDNA gene sequences to study the intraspecific variation of *C. sinensis*. In addition to providing new insights into the genetic diversity of the Chinese liver fluke, this study is the first global population genetic analysis of *C. sinensis* based on *cox*1 and *nad*1 gene sequences that we are aware of.

One of the most recognized indications for identifying the genetic variety of parasites is the *cox*1 gene sequences. According to Le *et al*. (2000), mitochondrial DNAs are one of the reliable and instructive molecular markers for helminth parasite categorization and genetic characterization. In order to ascertain the genetic diversity and intraspecific variation among *C. sinensis* isolates from various hosts, 251 *cox*1 gene sequences that were previously released in the GenBank were examined. The results of this study's population structure analysis of *C. sinensis* could not be compared to previously published findings due to insufficient data. However, a comparison was done with the population pattern of other helminths, such as *Fasciola* (*F*.) *hepatica* and *Echinococcus (E.) granulosus*. To find out how much polymorphism there was in the population, nucleotide diversity was estimated. In the present investigation, we observed mean nucleotide diversities in *C. sinensis cox*1 (0.00317) and *nad*1 (0.00293) genes were low, although Ohiolei *et al*. (2019a) reported higher nucleotide mean diversity in *E. granulosus* sensu stricto (s.s.). Similar findings have been made with *cox*1 mean nucleotide diversity indices of *E. canadensis* (0.00255) and *E. granulosus* s.s. (0.00026) (Selcuk *et al*., 2022).

Mean nucleotide diversity was reported to be 0.00264 and 0.00628 by Ohiolei *et al*. (2019b) and 0.00322 and 0.0014 by Alvi *et al*. (2020) in two further full-length *cox*1 and *nad*1 gene investigations on *Taenia* (*T*.) *hydatigena* and *E. granulosus*, respectively. *Moniezia* (*M*.) *expansa*, another tapeworm species that affects the population of sheep and goats, showed increased mean nucleotide diversity in the genes for *cox*1 (0.03787) and *nad*1 (0.04402) (Alshammari *et al*., 2024). But there have also been reports of even more nucleotide variety in *F. hepatica* (*cox*1=0.17426, *nad*1=0.21029) (Alvi *et al*., 2023).

Through analysis of mt-CO1 sequences submitted to GenBank, the population structure and genetic diversity of *C. sinensis* were examined in this work. Our findings offer insights into the global population dynamics and gene flow. The parasite's genetic diversity and population structure were examined using 189 *nad*1 and 251 *cox*1 gene sequences that were uploaded to the NCBI database. In order to develop a suitable therapy and management program, the biological evolution of the disease has been examined (Zhang *et al*., 2019).

To find out how distinct each of the population's haplotypes is, haplotype diversity was computed. The present study's *cox*1 and *nad*1 haplotype diversity values (*cox*1=0.948, *nad*1=0.9192) differed significantly from Selcuk *et al*. (2022)'s Hd values for the *cox*1 gene of *E. granulosus* s.s. (Hd=0.640). Nonetheless, the results were consistent with those of *E. granulosus* s.s. (*cox*1=0.925, *nad*1=0.834) and *E. ortleppi* (*cox*1=0.857) published by Alvi *et al*. (2020). However, the results are significantly greater than those found by Alvi *et al*. (2020) about the diversity of *E. ortleppi*'s *nad*1 gene haplotypes (0.25). Ohiolei *et al*. (2019b) reported similar results for *T. hydatigena* (*cox*1=0.906, *nad*1=0.867). Conversely, the results are in conflict with those published by Wakid and Alsulami (2022).

The results of this investigation also conflict with those of Ohiolei *et al*. (2019a), who examined the complete *cox*1gene sequences of *E. canadensis* isolates from Nigeria (*cox*1=0.3935, *nad*1=0.181). Alvi *et al*. (2023) and Alshammari *et al*. (2024) found similar patterns for *F. hepatica* (*cox*1=0.869, *nad*1=0.860) and *M. expansa* (*cox*1=0.950, *nad*1=0.944), respectively. Analyses of the *cox*1 and *nad*1 gene sequences conducted during the current investigation have identified 87 and 73 haplotypes, respectively. According to Oryan *et al*. (2015), impartiality values, such as population growth and nucleotide diversity, were tested using neutrality tests (Tajima's D and Fu's LD).

Tajima's D test is used to identify deviations from neutrality standards. Positive values show heterozygosity (having selection advantage), whereas negative values show a rapid rise in the population with one allele having a selected advantage over the other (Stephens *et al*., 2001; Vamathevan *et al*., 2008). The sequences evaluated using Tajima's D in this study had low and negative values (*cox*1=-2.41336, *nad*1=-2.50463), which is suggestive of a population growth in the near future. The current study's results are consistent with those of Wakid and Alsulami (2022) and Ohiolei *et al*. (2019b) for *T. hydatigena* (*cox*1= -1.23884). Tajima's D value for the *nad*1 gene in this investigation, however, was different from what Ohiolei *et al*. (2019b) reported. Similar negative results have been reported for *E. granulosus* s.s. (*cox*1=-1.45764, *nad*1=-0.26602), *E. canadensis* (*cox*1=-1.00957), and *M. expansa* (-0.64744) by Alvi *et al*. (2020), Ohiolei *et al*. (2019a), and Alshammari *et al*. (2024). Conversely, additional investigations have also revealed positive values for *F. hepatica* (*cox*1=3.40314) and *E. ortleppi* (*cox*1=0.59845) (Alvi *et al*., 2020; Alvi *et al*., 2023).

Fu's FS can be used as a sensitivity marker to examine the parasite population's expansion. The current investigation has reported negative values of Fu's FS (*cox*1=-32.908, *nad*1=-107.788), which suggests that diverse populations of parasites belonging to the same gene pool, particularly *C. sinensis*, have a common growth pattern (Fu *et al*., 1997; Li *et al*., 2009). Alvi *et al*. (2020b) and Selcuk *et al*. (2022) have reported similar *cox*1 Fu's FS values for *E. granulosus* s.s. (-11.823; -2.47269). However, the current study's *nad*1 Fu's Fs values are statistically insignificant despite being very negative. The Fu's Fs trend is unaffected by the findings that Alvi *et al*. (2020), Alvi *et al*. (2023), and Alshammari *et al*. (2024) previously published for *E. ortleppi*, *F. hepatica*, and *M. expansa*, respectively.

Consequently, we report new data on genetic diversity in populations of *C. sinensis* within and between Russia and Vietnam. We believe that this information could have important implications for evolution, medicine, and epidemiology. The full-size *cox*1 and *nad*1 gene and the anticipated structure of the *cox*1 protein have demonstrated potential as molecular markers for trematode genetic research. Additional samples and additional molecular markers are needed to investigate the population genetic structure of this liver fluke.

The mtDNA *cox*1 gene sequences for the zoonotic trematode *C. sinensis* provide useful genetic markers for phylogeographic reconstructions and intraspecific variability assessment. To some extent, however, the phylogeographic data can be regarded as preliminary. Increasing the geographic range of this parasite and the number of DNA regions analyzed will allow for a more accurate reconstruction of its worldwide history (Tatonova *et al*. 2013).

Conclusions: The results of this investigation may shed light on the genetic diversity and population organization of *C. sinensis*. Higher diversity index values suggest greater genetic variation in the *cox*1 gene. The low and negative neutrality indices suggest potential population expansion. The length of the gene sequences has an impact on these characteristics. In order to determine the degree of variation in various genes, the current study recommended evaluating full-length genetic markers. As a neglected parasite, this has limited molecular characterization and population structure studies globally. Further studies in unexplored regions are essential for a more comprehensive understanding of the parasite's epidemiology

Authors contribution: The study was conceptualized by MAA, MS, AAB and HBY. MAA, MWU, and MHW created the methodology, while SUH, MWU, LL, and MAM implemented it. TJ, SUH, FG, LL and MAA prepared the original draft, while SH, SOR, MHW, and MS handled the review and editing. The submitted version

of the manuscript has been read and approved by all authors.

Funding: The authors thankful to the Deanship of Research and Graduate Studies, King Khalid University, Abha, Saudi Arabia for financially supporting this work through the Large Research Group Project under Grant No. R.G.P.2/153/46.

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability: All data during study are included in this manuscript.

REFERENCES

Alshammari A, Ali U, Kabli AM, *et al*., 2024. Global scenario of genetic diversity in *cox1* and *nad1* genes of *Moniezia expansa*. Parasite Epidemiol Control 24:e00333.

Alvi MA, Ohiolei JA, Saqib M, *et al*., 2020. *Echinococcus granulosus* (sensu stricto) (G1, G3) and *E. ortleppi* (G5) in Pakistan: phylogeny, genetic diversity and population structural analysis based on mitochondrial DNA. Parasite Vector 13:1-10.

Alvi MA, Khalid A, Ali RM, *et al*., 2023. Genetic variation and population structure of *Fasciola hepatica*: an in silico analysis. Parasitol Res 122:2155-2173.

Chelomina GN, Tatonova YV, Hung NM, *et al*., 2014. Genetic diversity of the Chinese liver fluke *Clonorchis sinensis* from Russia and Vietnam. Int J Parasitol 44:795-810.

Fu YX, 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147: 915-25.

Kim HS, Nam HW, Ahn HJ, *et al*., 2022. Relationship between *Clonorchis sinensis* infection and cholangiocarcinoma in Korea. Korean J Parasitol 60:261-271.

Kim JG, Ahn CS, Sripa B, *et al*., 2019. *Clonorchis sinensis* omega-class glutathione transferases are reliable biomarkers for serodiagnosis of clonorchiasis and opisthorchiasis. Clin Microbiol Infect 25:109-e1.

Kim JW, Yi J, Park J, *et al*., 2021.Transcriptomic profiling of three-dimensional cholangiocyte spheroids long term exposed to repetitive *Clonorchis sinensis* excretory-secretory products. Parasite Vector 14:1-6.

Kinkar L, Korhonen PK, Saarma U, *et al*., 2023. Genome-wide exploration reveals distinctive northern and southern variants of *Clonorchis sinensis* in the Far East. Mol Ecol Resour 23:833-843.

Knudsen B, Knudsen T, Flensborg M, *et al*., 2007. CLC Main Workbench, Version 5.5; CLC bio: Aarhus, Denmark.

Lee SU and Huh S, 2004. Variation of nuclear and mitochondrial DNAs in Korean and Chinese isolates of *Clonorchis sinensis*. Korean J Parasitol 42:145-148.

Li YL, Kong XY, Yu ZN, *et al*., 2009. Genetic diversity and historical demography of Chinese shrimp *Feneropenaeus chinensis* in Yellow Sea and Bohai Sea based on mitochondrial DNA analysis. Afr J Biotechnol 8:7.

Li WH, Jia WZ, Qu ZG, *et al*., 2013. Molecular characterization of *Taenia multiceps* isolates from Gansu Province, China by sequencing of mitochondrial cytochrome C oxidase subunit 1. Korean J Parasitol 51:197-201.

Li Z, Xin H, Qian MB, *et al*., 2022. *Clonorchis sinensis* reinfection rate and reinfection determinants: a prospective cohort study in Hengxian county, Guangxi, China. J Infect Dis 225:481-491.

Liu GH, Li B, Li JY, *et al*., 2012. Genetic variation among *Clonorchis sinensis* isolates from different geographic regions in China revealed by sequence analyses of four mitochondrial genes. J Helminthol 86:479-484.

Le TH, Blair D and McManus DP, 2000. Mitochondrial DNA sequences of human schistosomes: the current status. Int J Parasitol. 30:283-290.

Leigh JW, Bryant D and Nakagawa S, 2015. POPART: full-feature software for haplotype network construction. Methods Ecol Evol 6:1110-6.

Maddison DR, Swofford DL and Maddison WP, 1997. NEXUS: an extensible file format for systematic information. Syst Biol 46:590-621.

Mahmood Q, Younus M, Sadiq S, *et al*., 2022. Prevalence and Associated Risk Factors of Cystic Echinococcosis in Food Animals--A Neglected and Prevailing Zoonosis. Pak Vet J 42: 59-64.

Mohammed AB, Mero WM and Nerway CA, 2021. Molecular identification of *Fasciola* spp. isolated from domestic animals based on DNA sequencing of the nuclear ribosomal ITS1-ITS2 markers, Kurdistan Region, Iraq. Pak Vet J 42: 246-250.

Nguyen TT, Bui DT, Losson B, *et al*., 2024. Effectiveness of health education in improving knowledge, attitude and practice related to foodborne zoonotic trematodes in Vietnam, with a particular focus on *Clonorchis sinensis*. Trop Med Int Health 29:280-291.

Ohiolei JA, Yan HB, Li L, *et al*., 2019. Cystic echinococcosis in Nigeria: first insight into the genotypes of *Echinococcus granulosus* in animals. Parasite Vector 12:1-10.

Ohiolei JA, Luka J, Zhu GQ, *et al*., 2019. First molecular description, phylogeny and genetic variation of *Taenia hydatigena* from Nigerian sheep and goats based on three mitochondrial genes. Parasite Vector 12:1-11.

Oryan A, Moazeni M, Amrabadi O, *et al*., 2015. Comparison of distribution pattern, pathogenesis and molecular characteristics of larval stages of *Taenia multiceps* in sheep and goats. Small Rumin Res 132:44-49.

Park GM and Yong TS, 2001. Geographical variation of the liver fluke, *Clonorchis sinensis*, from Korea and China based on the karyotypes, zymodeme and DNA sequences. Southeast Asian J Trop Med Public Health 32:12-16.

Park GM, 2007. Genetic comparison of liver flukes, *Clonorchis sinensis* and *Opisthorchis viverrini*, based on rDNA and mtDNA gene sequences. Parasitol Res 100:351-357.

Qamar W, Zaman MA, Faheem M, *et al*., 2022. Molecular confirmation and genetic characterization of *Haemonchus contortus* isolates at the nuclear ribosomal ITS2 region: first update from Jhang region of Pakistan. Pak Vet J 42:251-255.

Qian MB and Zhou XN, 2021. *Clonorchis sinensis*. Trends Parasitol 37:1014-1015.

Qian MB, Jiang ZH, Ge T, *et al*., 2020. Rapid screening of *Clonorchis sinensis* infection: performance of a method based on raw-freshwater fish-eating practice. Acta Trop 207:105380.

Rafique A, Nasir S, Ashraf A, *et al*., 2022. Sero-surveillance and risk factors analysis of caprine toxoplasmosis in Faisalabad Punjab, Pakistan. Pak Vet J 42:102-106.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, *et al*., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol 34:3299-3302.

Selcuk MA, Celik F, Kesik HK, *et al*., 2022. In Silico evaluation of the haplotype diversity, phylogenetic variation and population structure of human *E. granulosus* sensu stricto (G1 Genotype) sequences. Pathogens 11:1346.

Shi Y, Yu K, Liang A, *et al*., 2020. Identification and analysis of the tegument protein and excretory-secretory products of the carcinogenic liver fluke *Clonorchis sinensis*. Front Microbiol 11: 555730.

Sota P, Anditya M, Kotepui M, *et al*., 2024. Prevalence estimates of *Opisthorchis viverrini* and *Clonorchis sinensis* infection in the Greater Mekong subregion: a systematic review and meta-analysis. Infect Dis Poverty 13:1-15.

Stephens JC, Schneider JA, Tanguay DA, *et al*., 2001. Haplotype variation and linkage disequilibrium in 313 human genes. Science 293:489-93.

Tantrawatpan C, Maleewong W, Thanchomnang T, *et al*., 2023. Intron Regions as Genetic Markers for Population Genetic Investigations of *Opisthorchis viverrini* sensu lato and *Clonorchis sinensis*. Animals 13:3200.

Tatonova YV, Chelomina GN and Besprosvannykh VV, 2012. Genetic diversity of nuclear ITS1–5.8 S–ITS2 rDNA sequence in *Clonorchis sinensis* Cobbold, 1875 (Trematoda: Opisthorchidae) from the Russian Far East. Parasitol Int 61:664-674.

Tatonova YV, Chelomina GN and Besprozvannykh VV, 2013. Genetic diversity of *Clonorchis sinensis* (Trematoda: Opisthorchiidae) in the Russian southern Far East based on mtDNA *cox1* sequence variation. Folia Parasitol 60:155-62.

Vamathevan JJ, Hasan S, Emes RD, *et al*., 2008. The role of positive selection in determining the molecular cause of species differences in disease. BMC Evol Biol 8:1-4.

Wakid MH and Alsulami MN, 2022. Genotyping of *Taenia hydatigena* isolated from sheep and goats in KSA based on Cox1 gene. Saudi J Biol Sci 29:1270-1275.

Wang D, Korhonen PK, Gasser RB, *et al*., 2018. Improved genomic resources and new bioinformatic workflow for the carcinogenic parasite *Clonorchis sinensis*: Biotechnological implications. Biotechnol Adv 36:894-904.

Wu Y, Deng X, Wu Z, *et al*., 2023. Multilayer omics reveals the molecular mechanism of early infection of *Clonorchis sinensis* juvenile. Parasite Vector 16:285.

Xu M, Jiang Y, Yin J, *et al*., 2021. Risk factors for *clonorchis sinensis* infection in residents of Binyang, Guangxi: a cross-sectional and logistic analysis study. Front Public Health 9:588325.

Young ND, Stroehlein AJ, Kinkar L, *et al*., 2021. High-quality reference genome for *Clonorchis sinensis*. Genomics 113:1605-1615.

Zhang X, Sun B, Tang Q, *et al*., 2019. Molecular identification and phylogenetic analysis of nuclear rDNA sequences of *Clonorchis sinensis* isolates from human fecal samples in Heilongjiang Province, China. Front Microbiol 10:26.

Zhang Y, Gong QL, Lv QB, *et al*., 2020. Prevalence of *Clonorchis sinensis* infection in fish in South-East Asia: A systematic review and meta-analysis. J Fish Dis 43:1409-1418.