Molecular identification and phylogenetic analysis of *Pulex irritans* in different regions of Iran

Shahin Seidy, Mousa Tavassoli, Farnaz Malekifard

Department of Pathobiology, Faculty of Veterinary Medicine, Urmia University, Urmia, West Azerbaijan, Iran.

**ABSTRACT**

The present study was conducted to perform a molecular comparison of *Pulex irritans* based on the mitochondrial genome in four different climatic regions including the Caspian Sea region, a mountainous region, Persian Gulf region, and the Central Desert region, and based on nuclear ribosome genome in the west and northwestern Iran. A total of 1937 adult flea samples were collected including 1019 *P. irritans* (52.61%) and 918 *Ctenocephalides canis* (47.39%) from various hosts including humans (14.1%), sheep (22%), goats (33.5%), dogs (25.6%) and houses (6.7%) between April 2018 and May 2019. The samples collected from different hosts had similar morphological characteristics. However, there were slight differences based on mitochondrial markers and nuclear ribosomal markers in the study populations. The results from the phylogenetic tree based on three nuclear ribosome and mitochondrial markers showed that despite the slight differences in this sequence of different hosts and cities, all samples from different regions are in the same phylogeny. The results of ribosomal and mitochondrial genome analysis showed that these pieces are useful for demonstrating intraspecific similarity, and differentiation at species level and genus of *P. irritans*.

**Keywords**

Flea, Iran, ITS1, ITS2, Cox1, Pulex irritans

**Abbreviations**

- COX1: Cytochrome oxidase 1
- PCR: Polymerase chain reaction
- ITS1: Internal transcribed spacer I
- ITS2: Internal transcribed spacer II
- mtDNA: Mitochondrial DNA
- rDNA: ribosomal DNA
- MBST: Molecular Biological System Transfer
- µl: microliter

https://IJVST.um.ac.ir
**Introduction**

*Pulex irritans* belongs to *Siphonaptera* as a relatively-small order of wingless holometabolous insects, known as ectoparasites in their adult stage [1].

Certain flea species can transmit pathogens such as *Yersinia pestis*, *Rickettsia typhi*, and *Bartonella henselae* through their bites, feces and saliva. Biting can also cause severe itching and skin infections [2, 3, 4]. *P. irritans* can cause serious health complications and transmit diseases in many parts of the world [5]. Global warming, international trade, travel, and population growth influence the epidemiology of these infections [6, 7, 8].

Cladistic analysis using morphological features including head, chest, and abdomen have been conducted to classify various fleas, nevertheless, accurately identifying certain flea species is difficult owing to variations in their morphological characteristics. Molecular markers are therefore used to accurately identify different flea species and investigate their phylogenetic relationships [9, 10, 11, 12].

Molecular markers can be used to effectively characterize and specify an insect owing to their high stability and conserved nature; for instance, ribosomal DNA (rDNA) with its subunits such as 28s, 18s, 5.8s, ITS1 and ITS2 is commonly used to identify insect species [9, 10, 11]. As the largest protein-encoding gene in the mitochondria of metazoan organisms, the cytochrome oxidase subunit I (COXI) gene has been extensively used to perform phylogenetic research and identify species and their differences [13, 14].

Accurately identifying flea species is crucial given that they constitute a vector of dangerous pathogens and thus a serious threat to public health. This study employed ITS1, ITS2, and COX1 as molecular markers to investigate the molecular characteristics of different populations of *P. irritans* and their phylogenetic relationships in four areas in Iran.

**Results**

After morphological examinations using reliable identification keys for Iranian fleas, 1019 (52.61%) out of 1937 samples were identified as *P. irritans* (Table 1). Furthermore, the amplicons obtained from PCR for ITS1, ITS2 (rDNA), and COX1 (mtDNA) were approximately 1000 bp, 500 bp and 700 bp, respectively. The Clustal Omega based comparison of the ITS1 sequence in *P. irritans* showed a similarity of 99.59% and a molecular diversity of 0.41% between the sequences in ten different regions (Figure 1). A 100% similarity was observed among the sequences of Sanandaj (MN684797.1), Khorramabad (MN684787.1), Urmia (MN684784.1), Bahar (MN684777.1), and Kermanshah (MN 684776.1). These sequences were also similar (99.79%) to those of Hamadan (MN684779.1), Kuhdasht (MN684790.1), Kamyaran (MN684786.1), Gilangharb(MN684775.1), and Mahabad (MN684778.1) with four nucleotide differences. Four nucleotide polymorphisms were observed at positions 152, 165, 193 and 279 within the ITS1 spacer in the ten populations. Sequencing the ITS1 fragment showed three 99-bp replicas in the first half of the sequence beginning at positions 146, 246, and 352 and ending at position 431. Comparing the ITS2 gene sequences in the ten regions showed a 100% similarity between the sequences (Figure 2). Comparing the COX1 sequence of the nuclear genome (mtDNA) in the four different geographical zones (accession ID: MN173748.761) found all the species to belong to *P. irritans* and showed a 99.86% similarity between the sequences (Figure 3).

Multiple alignments of amino acids in COX1 showed a 99.86% similarity and only one difference at position 54 of the methionine isolate and other isolates in Hamadan (Figure 4). Two subclades of the phylogenetic tree based on the similarities between COX1 sequences in the present study and registered sequences in GenBank comprised subclade A1, including *P. irritans* of the present and previous studies conducted in Iran and that isolated in China, Turkey, Jerusalem, Spain and Croatia. Subclade A2 included *Ctenocephalides felis* in Iran and *C. canis* in Turkey. Subclades A3 and A4 included *C. orientis* and *C. canis*, respectively. Iranian subclade B also included *Nosopsyllus fasciatus* and *Xenopsylla cheopis* (Figure 5).

Three subclades of the phylogenetic tree based on the similarities between ITS1 sequences in the present study and registered sequences in GenBank comprised subclade A, including the sequences of the present and previous studies conducted in Iran. Subclades B and C included *C. canis* and *N. fasciatus*, respectively (Figure 6).

Four subclades of the phylogenetic tree based on the similarities between ITS2 sequences in the present study and registered sequences in GenBank comprised subclade C, including the sequences of the present study. Subclades A and B included *C. canis* and *X. cheopis*, respectively. Subclade D also included *N. fasciatus* (Figure 7).

The morphological characteristics of *P. irritans* were consistent with the molecular findings. Analyzing the partial COX1 gene can help identify *P. irritans* species and assess their intraspecific similarity. ITS1 and ITS2 sequences could also be used as useful molecular identification and phylogenetic analysis of *Pulex irritans*.
Table 2. Nucleotide sequence and specificity of the primers used.

<table>
<thead>
<tr>
<th>Accession number</th>
<th>Primers</th>
<th>Host</th>
<th>Location</th>
<th>Isolation source</th>
<th>Longitude, latitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>MN684792.1</td>
<td>ITS2</td>
<td>Dog</td>
<td>Kermanshah</td>
<td>Ker</td>
<td>34.1397° N, 45.9206° E</td>
</tr>
<tr>
<td>MN684782.1</td>
<td>ITS2</td>
<td>Human</td>
<td>Kermanshah</td>
<td>Ker</td>
<td>34.3277° N, 47.0778° E</td>
</tr>
<tr>
<td>MN684801.1</td>
<td>ITS2</td>
<td>Goat</td>
<td>Sanandaj</td>
<td>San</td>
<td>35.3219° N, 46.9862° E</td>
</tr>
<tr>
<td>MN684781.1</td>
<td>ITS2</td>
<td>Sheep</td>
<td>Kurdestan</td>
<td>Kur</td>
<td>34.7956° N, 46.9368° E</td>
</tr>
<tr>
<td>MN684780.1</td>
<td>ITS2</td>
<td>Dog</td>
<td>Mahabad</td>
<td>Mah</td>
<td>34.9083° N, 48.4393° E</td>
</tr>
<tr>
<td>MN684793.1</td>
<td>ITS2</td>
<td>Dog</td>
<td>Lorestan</td>
<td>Kho</td>
<td>33.5275° N, 47.6111° E</td>
</tr>
<tr>
<td>MN684794.1</td>
<td>ITS2</td>
<td>Dog</td>
<td>Ahvaz</td>
<td>Ahv</td>
<td>36.5659° N, 53.0586° E</td>
</tr>
<tr>
<td>MN684798.1</td>
<td>ITS2</td>
<td>Dog</td>
<td>Sari</td>
<td>Sar</td>
<td>30.2839° N, 57.0834° E</td>
</tr>
<tr>
<td>MN684795.1</td>
<td>ITS2</td>
<td>Dog</td>
<td>Esfarayen</td>
<td>Esf</td>
<td>37.4710° N, 57.1013° E</td>
</tr>
</tbody>
</table>
molecular markers for species-level differentiation and intra-specific similarity determination. The COX1 gene is, however, more efficient than the ITS1 and ITS2 genes in detecting genus, species and intra-species similarities. The diversity at the genus level and even between members of the same genus caused by high levels of replacement in the Internal transcribed spacer counteracted their beneficial effects. Given this diversity, the sequences of the heterogenic versions of ITS fragments are determined in cloning plasmids and then specifically, which requires laborious and costly studies[15].

Figure 1.
Alignments of partial ribosomal DNA (ITS1) sequences in *Pulex irritans* isolated from sheep, human, goats, and dogs in the west and northwest of Iran (Kho(sheep), Bah(sheep), Ker(human), San(goat), Urm(sheep), Kuh(human), Kam(human), Ham(sheep), Ghi(goat), and Mah(dog) with accession nos. MN684787.1, MN684777.1, MN684776.1, MN684797.1, MN684784.1, MN684790.1, MN684786.1, MN684779.1, MN684775.1, MN684778.1) Nucleotides in horizontal boxes display repeated units, and vertical boxes indicates the polymorphic site. Bah: Bahar; Ghi: Gilan-e Gharb; Ham: Hamedan; Kam: Kamiyaran; Ker: Kermanshah; Kho: Khorramabad; Kuh: Kuhdasht; Mah: Mahabad; San:Sanandaj; Urm: Urmia.
Molecular identification and phylogenetic analysis of Pulex irritans Seydi et al., IJVST 2021; Vol. 13, No.2
DOI: 10.22067/ijvst.2021.70946.1055

Discussion

Molecular methods have enabled the identification of different ectoparasite species with high morphological similarities [16]. The present study employed phylogenetic and molecular approaches to compare P. irritans in four different geographical zones of Iran. The morphological characteristics of P. irritans were found to be consistent with those obtained in previous studies conducted in Iran to determine the flea fauna in different hosts and cities [17, 4, 18].

The host specificity can affect the intraspecific genetic diversity given the high levels of intraspecific genetic variations in general parasite species, which cause their infestation of broad ranges of hosts [19]. Research suggests no specific hosts for fleas, as they prefer diverse types of the host, which demonstrates intraspecific genetic diversity. Hornok et al. (2018) reported diversity in synanthropic flea species such as C. canis and P. irritans by investigating mitochondrial sequences [20]. The present morphological data showed no differences between the P. irritans specimens collected from different geographical zones. The results of the morphological investigation were also consistent with those of the molecular exploration. The present study reported a 100% similarity between the nucleotide sequence of COX 1 gene in the P. irritans and the sequence of this gene in the P. irritans in similar studies conducted in Iran [MF380390.1, MF380391.1 (NCBI)]. Similarities to the sequences in Spain (LT797470.1) and China (MF006666.1) were also reported as 99.85% and 99.55%, respectively. Furthermore, the present findings were consistent with those obtained by Seyyedzadeh et al. (2018), who reported no differences between the isolates of C. canis in different areas of West Azerbaijan, Iran [21] and with those obtained by Zurita et al. (2019), who observed no significant differences in morphological data between P. irritans in Spain and Argentina. In addition, Zurita et al. (2019) reported a significant intraspecific similarity between the two populations based on mitochondrial genes [22]. Hornok et al. (2018) observed no morphological differences between human and wild carnivorous P. irritans specimens in Hungary and Croatia [20], which is consistent with the present results. In contrast, Krasnov et al. (2015) found morphological differences between the flea species isolated from different hosts in different geographic areas to suggest high levels of genetic diversity [23].

Evidence suggests ITS1 and ITS2 constitute appropriate molecular markers for analyzing phylogenetic relationships in fleas at their species level [24, 11]. The present research observed a 99.59% similarity between the nucleotide sequences of the ITS1 spacer, in five provinces in Iran. Analyzing the nucleotide sequence of the ITS1 spacer, also showed three 99-bp replicas in its first half, which is consistent with the results of Ghawami et al. (2018), who compared the
Figure 3.
Alignments of partial mitochondrial DNA (COX1) sequences of Pulex irritans among populations from this study (Urm(human), San(sheep), Sar(sheep), Mah(goat), Kuh(sheep), Kho(human), Esf(human), Kam(sheep), Ghi(sheep), Ahv(goat), Ker(human), Kerm(sheep), Ham(human), and Bah(human) with accession nos. MN173761.1, MN173760.1, MN173759.1, MN173758.1, MN173757.1, MN173756.1, MN173755.1, MN173754.1, MN173752.1, MN173750.1, MN173749.1, MN173748.1, MN173753.1, MN173751.1). The vertical boxes indicate the polymorphic sites. Bah: Bahar; Ghi: Gilan-e Gharb; Ham: Hamedan; Kam: Kamiyaran; Ker: Kermanshah; Kho: Khorramabad; Kuh: Kuhdasht; San: Sanandaj; Urm: Urmia.

nucleotide sequence of the ITS1 spacer, in *P. irritans* in two geographic areas and reported only one nucleotide difference with a 99.85% similarity [25]. In line with the present results, those obtained by Vobis et al. (2004) suggested a relatively-constant nucleotide se-

sequence of the ITS1 spacer, in different *C. felis* populations [11].

A 100% similarity was observed between the ITS2 nucleotide sequence of different samples. Zurita et al. [11] reported a 100% similarity between the ITS2 sequence of *C. felis* populations from different geographical areas.

Seidy et al., IJVST 2021; Vol.13, No.2 DOI: 10.22067/ijvst.2021.70946.1055
RESEARCH ARTICLE

IRANIAN JOURNAL OF VETERINARY SCIENCE AND TECHNOLOGY

Molecular identification and phylogenetic analysis of Pulex irritans
Seydi et al., IJVST 2021; Vol. 13, No.2
DOI: 10.22067/ijvst.2021.70946.1055

Figure 4.
Sequence alignment of COX1 amino acids for Pulex irritans in different geographical areas of Iran (Kerm (sheep), Bah (human), Ham (human), Ker(human), Ahv(goat), Ghi(sheep), Kam(sheep), Efs(human), Kho(human), Kuh(sheep), Mah(goat), Sar(sheep), San(sheep), Urm(human), with accession nos MN173748.1, MN173751.1, MN173753.1, MN173749.1, MN173750.1, MN173752.1, MN173754.1, MN173755.1, MN173756.1, MN173757.1, MN173758.1, MN173759.1, MN173760.1, MN173761.1). The vertical boxes indicate the polymorphic sites. Ahv: Ahvaz; Bah: Bahar; ESF: Esfarayen; Ghi: Gilan-e Gharb; Ham: Hamedan; Kam: Kamyaran; Ker: Kermanshah; Ker: Kerman; Kho: Khorramabad; Kuh: Kuhdasht; San: Sanandaj; Urm: Urmia.

(2019) reported an intraspecific variation of 90.1%-100% in four C. canis clones based on the ITS2 spacer, which is consistent with the present results. Vobis et al. (2004) also used the ITS2 sequence to identify different flea species and determine their intra-species differences [11]. P. irritans and C. canis belonging to the Pulicidae family constitute the main flea genera infesting human and cattle in the west and northwest of Iran. No significant differences were observed among the samples collected from different hosts. Insignificant differences were found between the study populations in terms of mitochondrial markers and the nuclear ribosome. The results of plotting the phylogenetic tree showed all the samples collected from different regions in the same branch despite their negligible sequence differences. Analyzing mitochondrial genome also showed that this sequence is more useful than the nuclear ribosomal genome in illustrating intra-species similarity and differentiation at genus and species levels.

Materials & Methods

Location

The Iranian plateau is located in the northern hemisphere at 32.4279 °N, 53.6880 °E. The four study geographical zones in Iran included region 1: Caspian Sea (temperature: 8.00-26.00 °C, annual precipitation: 400-1500 mm), region 2: Mountainous area (temperature: -5.00-29.00 °C, annual precipitation: 200-500 mm), region 3: Persian Gulf (temperature: 12.60-35.00 °C, annual precipitation: 200-300 mm) and region 4: The Central Desert (temperature: -4.00-44.00 °C, annual rainfall: below 100 mm). According to meteorological data, Iran is climatically divided into three zones, i.e. warm-dry, cold, temperate-humid, and warm-humid [26]. Four different geographical zones in this country from which adult fleas were collected included Kermanshah, Kordes-tan, Hamadan, Lorestan, West Azerbaijan, Khuzestan, Khorasan, Mazandaran, and Kerman (Figure 8).

Sampling

A total of 1937 flea samples were directly isolated from the host body and houses using light traps and human-baited traps. They were transferred in 70% ethanol to the parasitology laboratory of the Faculty of Veterinary Medicine. They were then cleaned with distilled water several times, immersed in 5% potassium hydroxide for 24 hours, re-washed with distilled water, and dehydrated with a graded series of alcohol (50%-96%). The fleas were ultimately identified by observing the slides with an optical microscope according to diagnostic keys [27, 28].

Morphological identification

All the P. irritans samples collected from different provinces and hosts were morphologically characterized by the following
Figure 5.
The phylogenetic tree of *Pulex irritans* based on COX1. The molecular phylogenetic analysis was performed and the evolutionary history was inferred by using the Maximum Likelihood method. The log likelihood (-1898.1614) of the phylogenetic tree was the highest. The percentage of the tree including a cluster of the relevant taxa is shown beside the branches whose length was measured as the number of substitutions per site on this scale drawing of the tree. The analysis involved 44 nucleotide sequences. All the positions containing gaps and missing data were eliminated. The final dataset contained 448 positions. Evolutionary analyses were conducted in MEGA6.
Molecular identification and phylogenetic analysis of *Pulex irritans* Seydi et al., IJVST 2021; Vol. 13, No.2

**Figure 6.**
The phylogenetic tree of *Pulex irritans* based on ITS1. The molecular phylogenetic analysis was performed and the evolutionary history obtained using maximum likelihood. The log likelihood (-1703.2418) of the phylogenetic tree was the highest. The percentage of the tree including a cluster of the relevant taxa is shown beside the branches whose length was measured as the number of substitutions per site on this scale drawing of the tree. The analysis involved 19 nucleotide sequences. All the positions containing gaps and missing data were eliminated. The final dataset contained 436 positions. Evolutionary analyses were conducted in MEGA6.

**Figure 7.**
The phylogenetic tree of *Pulex irritans* based on ITS2. The molecular phylogenetic analysis was performed and the evolutionary history obtained using maximum likelihood. The log likelihood (-1272.22468) of the phylogenetic tree was the highest. The percentage of the tree including a cluster of the relevant taxa is shown beside the branches whose length was measured as the number of substitutions per site on this scale drawing of the tree. The analysis involved 26 nucleotide sequences. All the positions containing gaps and missing data were eliminated. The final dataset contained 291 positions. Evolutionary analyses were conducted in MEGA6.
Molecular identification and phylogenetic analysis of *Pulex irritans*

Seidy et al., IJVST 2021; Vol. 13, No.2

DOI: 10.22067/ijvst.2021.70946.1055

**Table 2.** Nucleotide sequence and specificity of the primers used.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer nucleotide sequence</th>
<th>Annealing temperature (°C)</th>
<th>Fragment length (base pair; bp)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITS1</td>
<td>For: GTA CAC ACC GCC GGT GCG TAC T &lt;br&gt;Rev: GCT GCG TTC TTC ATC GAC CC</td>
<td>54 °C</td>
<td>1000 bp</td>
<td>[29]</td>
</tr>
<tr>
<td>ITS2</td>
<td>For: GGG TCG ATG AAG AAC GCA GC &lt;br&gt;Rev: GGG CAC ATG CTA GAC TCC GTGGTT CAA G</td>
<td>54 °C</td>
<td>500 bp</td>
<td>[29]</td>
</tr>
<tr>
<td>COX1</td>
<td>For: GGT CAA CAA ATC ATA AAGATA TTG G &lt;br&gt;Rev: GAA GGG TCA AAG AAT GAT GT</td>
<td>55 °C</td>
<td>700 bp</td>
<td>[30]</td>
</tr>
</tbody>
</table>

**Figure 8.**
Provinces of the study in the four geographical areas included Kermanshah, Kordestan, Hamadan, Lorestan, West Azerbaijan, Khuzestan, Khorasan, Mazandaran and Kerman.

features: the round-headed anterior margin with a pair of eyes and no protrusion, absence of genal and pronotal ctenidia, asymmetrical club of antenna, strong setae under the eyes, row of small spines mass on the inner surface of the posterior side of the coxa in females 7-10 and in males 8-12, and no pleural rod on the mesopleural segment [27]. Comparing the samples collected from different provinces and hosts based on morphological characteristics identified 52.61% to belong to *P. irritans* species despite their negligible differences.

**Molecular analysis**

The samples were individually crushed with a sterile scalpel, poured into a sterile microtube and their total DNA was extracted using DNA extraction kits (MBST, Tehran, Iran) according to the manufacturer’s instructions. The extracted DNA samples were stored in sterile microtubes at -20 °C until use. DNA was extracted from the samples isolated from different hosts in different areas. The specific primers used by Vobis et al (2004) were employed to identify *P. irritans* and 1000-bp ITS1 and 500-bp ITS2 fragments in the rDNA of *P. irritans*. Primers used by Folmer et al. (1994) were used to amplify the mitochondrial COX1 gene fleas *P. irritans* (Table 2). The result of this proliferation was observed as a 700 bp band [29]. To determine molecular differences between the species isolated from different hosts and areas, PCR was performed in a total volume of 50 µl, containing a 5-µl DNA template, a 5-µl 10x PCR buffer, 1-µl Taq Polymerase, 1 µl of each primer (20 µM), 1 µl of dNTP (100 µM) and 4 µl of MgCl2 (50 mM). All these materials were procured from SinaClon, Iran.

The thermal cycle of the PCR on fragments ITS1 and ITS2 was as follows: incubation at 94 °C for 5 minutes to denature double-stranded DNA followed by 30 cycles of one-minute denaturation cycles at 94 °C, annealing at 54 °C for 1 minute, extension at 72 °C for 1 minute and an additional extension at 72 °C for 5 minutes. The COX1 gene also underwent replication at 95 °C for 5 minutes followed by 35 cycles of denaturation cycles at 95 °C for 1 minute, annealing at 55 °C for 45 seconds, extension at 72 °C for 45 seconds, and final extension at 72 °C for 10 minutes. The PCR products were electrophoresed on 1.5% agarose gel, stained with a safe stain, visualized under UV light and then purified for sequencing. After purifying the PCR products, 34 samples, including 10 positive samples for ITS1, 10 positive for ITS2, and 14 positive for COX1 sequence, collected from different hosts were transferred to Takapouzist Company for sequencing.

**Data analysis**

A total of 34 PCR products were transferred to Takapouzist for sequencing. An accession number was obtained by online recording the sequencing results on NCBI by the host and geographical area. The data were analyzed using blasting sequences on NCBI and plotting the phylogenetic tree in Mega 6 using maximum likelihood and bootstrapping (1000 replicates). The clustering method proposed by Zurita et al.(2019) was also conducted. The required sequences were ultimately aligned and compared using EMBOSS Needle and Clustal Omega to evaluate the similarity percentage of the nucleotide sequences of mitochondrial (mtDNA) and ribosomal genomes (rDNA) in different provinces and hosts.
Molecular identification and phylogenetic analysis of *Pulex irritans*

Seydi et al., IJVST 2021; Vol. 13, No.2
DOI: 10.22067/ijvst.2021.70946.1055

**Authors’ Contributions**

Conceived and designed the experiments and revised the manuscript draft: S.S., M.T., performed the experiments, analysed the data and drafted the manuscript: S.S. All authors approved the final version of the manuscript.

**Acknowledgements**

The authors would like to express their gratitude to the authorities of the Faculty of Veterinary Medicine and Research Council of Urmia University for their approval and financial support of this research.

**Competing Interests**

The authors declare that there is no conflict of interest.

**References**


