Molecular characterization and phylogenetic analysis of pathogenic *Leptospira* serovars based on the *secY* gene

Shadi Afshar[1], Pejvak Khaki[2*], Majid Esmaelizad[3]

[1] Department of Microbiology, Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran. ORCID: https://orcid.org/0009-0007-7378-4095

[2] Department of Microbiology, Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran. ORCID: https://orcid.org/0000-0001-8839-1023

[3] Department of Research and Development, Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran. ORCID: https://orcid.org/0000-0001-8527-784X

*** Corresponding author:**

Pejvak Khaki, Department of Microbiology, Razi Vaccine and Serum Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran. Tel: +98-263 457 0038-46, Fax: +98-2634552194. Email: khakipejvak53@gmail.com

ABSTRACT

Leptospirosis, caused by *Leptospira* spp., is a globally distributed reemerging zoonotic disease, predominantly reported in tropical, subtropical, and temperate regions characterized by high humidity and heavy rainfall. Accurate identification of the causative agent is critical for effective diagnosis and disease control. This study aimed to identify and elucidate the phylogenetic relationships of *Leptospira* serovars isolated in Iran through sequencing of the *secY* gene.

A total of twenty-seven pathogenic and two saprophytic *Leptospira* serovars were obtained from the Reference Laboratory for *Leptospira*, Department of Microbiology, Razi Vaccine and Serum Research Institute (Karaj, Iran). Genomic DNA was extracted, and a 549-bp fragment of the *secY* gene was amplified using specific primers.

PCR amplification successfully identified the *secY* gene in all pathogenic serovars, whereas no amplification was observed in the non-pathogenic *L. biflexa*. Sequence analysis revealed genetic similarity among pathogenic serovars ranging from 69.7% to 100%, highlighting the high discriminatory power of the *secY* gene. Serovars of the same serovar showed >99% sequence identity, whereas greater divergence was observed among different serovars.

This gene effectively differentiated all serovars, with those belonging to the same serovar exhibiting high sequence identity (>99%), while significantly greater divergence was observed among different serovars.

Phylogenetic analysis based on 26 reference serovars from established databases revealed that the majority of the studied serovars clustered with known *Leptospira interrogans* serovars. Additionally, two serovars were grouped with published sequences of L. *borgpetersenii*, L. *alexanderi*, L. *santarosai*, and L. *weilii*.

In conclusion, the *secY* gene exhibited considerable variability among pathogenic *Leptospira* serovars, demonstrating its utility as a molecular marker for accurate identification and phylogenetic classification of *Leptospira* serovars.

Keywords: Leptospirosis, molecular characterization, phylogenetic analysis, *secY* gene.

1. Introduction

Leptospirosis is a bacterial infection caused by spirochetes of the *Leptospira* genus, which includes 71 species and over 320 serovars worldwide (1). This zoonotic disease affects approximately one million individuals annually and results in nearly 60,000 fatalities. *Leptospira* is transmitted either through direct contact with the urine of infected animals or indirectly through exposure to soil and water contaminated with the organism (2).

Once *Leptospira* enters the bloodstream, it can cause both localized and systemic damage in humans. The disease manifests a wide range of clinical presentations, from asymptomatic infections to severe, progressive forms that may result in life-threatening complications(1).

Although leptospirosis is a global concern, its prevalence is highest in subtropical and tropical regions, where warm and humid conditions facilitate the persistence and transmission of *Leptospira* (2-4). North of Iran, with its similar climate, provides ideal conditions for the bacterium to thrive, making leptospirosis endemic in the region. Epidemiological studies indicated that the provinces of Gilan, Mazandaran, and Golestan have the highest incidence rates, making them particularly susceptible to outbreaks(5).

Traditional serological characterization approaches, including the microscopic agglutination test (MAT), are labour-intensive, time-consuming, and often unable to distinguish all serovars effectively. In contrast, molecular techniques, particularly PCR-based methods, provide superior sensitivity and specificity, offering a more precise and efficient alternative for *Leptospira* characterization (6).

Molecular characterization of *Leptospira* serovars relies on PCR amplification and sequence analysis of various genes to accurately identify and differentiate *Leptospira* species and serovars.

Based on pulsed-field gel electrophoresis (PFGE) , Multi-Locus sequence typing (MLST), restriction fragment length polymorphism (RFLP), *16SrRNA* phylogenetic analysis, and DNA-DNA hybridization, the *Leptospira* genus is classified into three major groups: pathogenic, intermediate, and saprophytic. This classification is further supported by the analysis of leptospiral housekeeping genes, which serve as reliable markers for serovar identification. Among these, the *16S rRNA (rrs)* gene sequence is widely employed for the accurate classification of *Leptospira* serovars.

Additionally, several other genetic markers, such as *lipL32*, *lipL41*, *secY* , *gyrB*, *flab*, *rpoB*, *lfb1*, *adk*, and *icdA*, have been employed for serovars differentiation (7).

The *secY* gene, located on the CI chromosome, encodes a pre-protein translocase essential for exporting proteins across the cytoplasmic membrane through the general secretory (Sec) pathway. It has been widely used for both diagnosis and typing of *Leptospira* infections due to its strong discriminatory power. Early diagnostic PCR assays targeted conserved regions of the *secY* gene (8). And it remains a valuable tool for identifying species, strains, and serovars of *Leptospira*. Additionally, *24*provides critical epidemiological insights and serves taxonomic purposes (9).

Given its high discriminatory power, the *secY* gene serves as a valuable molecular marker for the identification and classification of *Leptospira* serovars. Therefore, the present study aims to utilize *secY* gene-based nucleotide sequencing and phylogenetic analysis to discriminate pathogenic *Leptospira* serovars.

2. Materials and methods

2.1. *Leptospira* Serovars and Culture Conditions

This study utilized 31 *Leptospira* serovars, comprising 29 pathogenic and two non-pathogenic serovars (Table 1). All isolates were obtained from the microbial collection of the Leptospira Reference Laboratory, Department of Microbiology, Razi Vaccine and Serum Research Institute (Karaj, Iran), and had been previously characterized and maintained under standard preservation conditions. The identity of each serovar was confirmed using pulsed-field gel electrophoresis (PFGE) and variable-number tandem repeat (VNTR) analyses, along with gene-based assays targeting the *lipL32, lipL41, ompL1, loa22,* and *16S rRNA* loci. In addition, 26 *secY* gene sequences from various *Leptospira* serovars were selected from the NCBI GenBank database and used for comparative analysis with the sequences obtained from the reference serovars in this study.

All serovars were cultured in EMJH selective medium (Difco, Sparks, USA) enriched with *Leptospira* supplement and 10% rabbit serum, and maintained aerobically at 28°C for 7–10 days. Bacterial growth was confirmed by the dark field microscope (Nikon Eclipse, Japan) observation.

2.2. Genomic DNA Extraction

Genomic DNA from *Leptospira* serovars was extracted using the High Pure PCR Template Preparation Kit (Roche) following the manufacturer's protocol with slight modifications. 50mL of fresh *Leptospira* culture was centrifuged at 17,000 × g for 20 minutes at 4°C. The pellet was resuspended in 200 µL of PBS or TE buffer, lysed with 200 µL of Tissue Lysis Buffer and 40 µL of Proteinase K, and incubated at 56°C for 30 minutes. After adding 100 µL of Binding Buffer, the mixture was transferred to a High Pure Filter Tube for DNA binding by centrifugation at 8,000 × g for 1 minute. Washing was performed using 500 µL each of Inhibitor Removal Buffer, Wash Buffer I, and Wash Buffer II, followed by centrifugation at 8,000 × g for 1 minute, with a final spin at 13,000 × g for 2 minutes. DNA was eluted with 100 µL of preheated Elution Buffer (70°C), incubated at room temperature for 5 minutes, and centrifuged at 8,000 × g for 1 minute. The extracted DNA was stored at -20°C, and its quality and quantity were assessed using a NanoDrop spectrophotometer (Epoch-BioTek, Winooski, VT, USA).

2.3. PCR amplification of the *secY* gene

In the present study, amplification of the *secY* gene (549 bp) in *Leptospira* serovars was performed using specific primers (Table 2). PCR reactions were carried out in a total volume of 16 µL, consisting of 8 µL of 2× Master Mix (Ampliqon, Denmark), 1 µL each of forward and reverse primers(Gene Fanavarn , Iran), (10 pmol/µL), 2 µL of genomic DNA (~100 ng), and 4 µL of sterile deionized water.

The PCR program comprised an initial denaturation at 94°C for 5 minutes, followed by 35 cycles of 94°C denaturation for 60 seconds, 62°C annealing for 60 seconds, and 72°C extension for 60 seconds, concluding with a final 10-minute extension at 72°C. PCR amplicons were analyzed by electrophoresis on a 1% agarose gel at 75 V for 75 minutes and visualized under UV light.

2.4. DNA sequencing and phylogenetic analyses

A total of 29 PCR products were sequenced by the Gene Fanavaran Company, Iran. The resulting sequences were initially assessed for quality using Chromas v2.6, and sequence homologies of *Leptospira* serovars were analyzed using the BLAST tool on the NCBI website. Sequences were then edited and exported in FASTA format using DNASTAR's EditSeq program. For comparison, the sequences obtained in this study were aligned with corresponding sequences from GenBank using the MegAlign software. In addition, the *secY* gene sequences were used to generate a phylogenetic tree and a table of sequence similarity and divergence to explore evolutionary relationships among different *Leptospira* serovars using the software.

Nucleotide sequences were translated into amino acid sequences using MEGA11, which identifies open reading frames and applies the standard genetic code. The resulting protein sequences were then analyzed for sequence diversity to identify potential variations in protein folding and function among the serovars.

Table 1. List of *Leptospira* serovars used in this study.

No.	Serogroup	Serovar	Country	Accession No.	No.	Accession No.	Serogroup	Serovar/strain	Country
1	Autumnalis	Autumnalis*	Iran	PV251995	32	AKWU02000005	*L.*interrogans	Canicola	Brazil
2	Canicola	Canicola*	Iran	PV251996	33	EU357961	*L.*interrogans	Canicola	Spain
3	Grippotyphosa	Grippotyphosa*	Iran	PV251997	34	CP044014	*L.*interrogans	Icterohaemorrhagiae	Brazil
4	Sejroe	Hardjo-bovis*	Iran	PV251998	35	NZ_CP043891	*L.*interrogans	Icterohaemorrhagiae	Malaysia
5	Icterohaemorrhagiae	Icterohaemorrhagiae*	Iran	PV365287	36	EU357997	*L.*interrogans	Icterohaemorrhagiae	Spain
6	Pomona	Pomona*	Iran	PV365288	37	AHMK02000045	*L.*interrogans	Pomona	USA
7	Sejroe	Sejroe*	Iran	PV365266	38	AFLT02000023	*L.*interrogans	Pomona	Australia
8	Patoc	Patoc	Iran	-	39	NZ_JADDWP010000012	*L.*interrogans	Pomona	NewZealand
9	Sejroe	Hardjo*	Iran	PV365267	40	EU357943	*L.*interrogans	Autumnalis	Spain
10	Pomona	Pomona*	Iran	PV365268	41	MH683048	*L.*interrogans	Autumnalis	Brazil
11	Icterohaemorrhagiae	Icterohaemorrhagiae*	Iran	PV365269	42	NZ_CP097315	*L.*interrogans	Strain N116	Belgium
12	Canicola	Canicola*	Iran	PV365270	43	NZ_CP133195	*L.*interrogans	Strain 22	Kazakhstan
13	Grippotyphosa	Grippotyphosa*	Iran	PV365271	44	NZ_CP072853	*L.*interrogans	Strain UI29382	Laos
14	Patoc	Patoc	Iran	-	45	NZ_CP039283	*L.*interrogans	Strain FMAS_AW1	Sri Lanka
15	Pomona	Pomona*	Iran	PV365272	46	EU358038	*L.*borgpetersenii	Nero	Spain
16	Autumnalis	Autumnalis*	Iran	PV365273	47	MH059525	*L.*borgpetersenii	Hardjo	USA
17	Malaysia	Malaysia*	Iran	PV365274	48	PP372556	*L.*borgpetersenii	Isolate 4280	Spain
18	Celledoni	Celledoni*	Iran	PV365275	49	EU358009	*L.*weilii	Coxi	Spain
19	Lyme	Lyme*	Iran	PV365276	50	EU358065	*L.*weilii	Langati	Spain
20	Djasiman	Djasiman*	Iran	PV365277	51	JN683942	*L.* kirschneri	Strain 201001687	Mayotte
21	Pyrogenes	Pyrogenes*	Iran	PV365278	52	EU358068	*L.* noguchii	Carimagua	Spain
22	Canicola	Canicola*	Iran	PV365279	53	KY113324	*L.* noguchii	Strain U386/2016	Brazil
23	Icterohaemorrhagiae	Icterohaemorrhagiae*	Iran	PV365280	54	NZ_AHMT02000039	*L.* alexanderi	Manhao	China
24	Ballum	Ballum*	Iran	PV365281	55	MT233030	*L.* alexanderi	Banna	India
25	Javanica	Javanica*	Iran	PV365282	56	EU358067	*L.* santarosai	Navet	Spain
26	Australis	Australis*	Iran	PV365283	57	EU358063	*L.* santarosai	Rama	Spain
27	Lai	Laitype lanylokowii*	Iran	PV365284					
28	Bataviae	Bataviae*	Iran	PV365285					
29	Sejroe	hardjobovis*	Iran	PV365286					
30	Canicola	Canicola*	Iran	PX122093					
31	Icterohaemorrhagiae	Icterohaemorrhagiae*	Iran	PX122094					

*The *secY* gene from these serovars was characterized and sequenced.

Table2. Oligonucleotide primers targeting the *secY* gene used in this study.

Primers	Sequence (5 '-3 ')	Fragment length (bp)	Tm (˚C)	GC	Reference
Forward primer	ATGCCGATCATTTTTGCTTC	549	53.20	40	(10, 11)
Reverse primer	CCGTCCCTTAATTTTAGACTTCTTC	549	59.70	40	(10)

3. Results

3.1 *SecY* Amplification

The results of PCR products observed a 549-bp fragment that was exclusively in pathogenic *Leptospira* serovars and was absent in non-pathogenic serovars.

3.2 Nucleotide Sequencing and Phylogenetic Analysis

In this study, we analyzed the *secY* gene sequences of 29 pathogenic *Leptospira* serovars (Table 3, Figure 1).

As shown in Table 3, a comparative analysis of *secY* gene sequences revealed similarities ranging from 69.7% to 100%, highlighting the gene's discriminatory power in differentiating serovars. Identical serovars exhibited very high similarity (≥99.5%).

Among the analyzed sequences, 100% identity was observed in four *L.* Icterohaemorrhagiae serovars, three *L.* Pomona serovars, two *L.* Autumnalis serovars, and two *L.* Grippotyphosa serovars. High sequence similarity was also found among the four *L.* Canicola serovars (>99.5%) and the two *L.* Hardjo bovis serovars (99.8%). In contrast, *L.* Celledoni and *L.* Javanica exhibited the lowest similarity to the other serovars. Notably, *L.* Celledoni showed the greatest divergence, with only 69.7% similarity to *L.* Malaysia among the 29 sequenced serovars.

Phylogenetic analysis based on *secY* gene sequences (Figure 1) classified the 29 *Leptospira* serovars into two main clusters. Cluster I included 27 serovars and was further divided into the two subclusters. Subcluster 1 comprised the majority of pathogenic serovars—such as *L.* Icterohaemorrhagiae, *L.* Canicola, *L.* Pomona, *L.* Grippotyphosa, and *L.* Sejroe (Hardjo-bovis)—which were genetically closely related. Other serovars, including *L.* Autumnalis, *L.* Bataviae, and *L.* Sejroe (Sejroe), were also grouped within this subcluster. Subcluster 2 contained only *L.* Malaysia, which shared over 80% similarity with subcluster 1 but was genetically distinct enough to form a separate branch.
Cluster II consisted of *L.* Celledoni and *L.* Javanica, both of which displayed substantial divergence from the remaining serovars.
Overall, this classification highlights the genetic diversity among pathogenic *Leptospira* serovars, with *L.* Celledoni and *L.* Javanica forming a distinct lineage separate from the majority of the analyzed serovars.

Table 3. Similarity and Divergence Percentages among 29 Pathogenic Serovars based on *secY* Gene Sequence Analysis Using the Clustal V Method in MegAlign Software.

Percent Identity

Divergence

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	
1		99.5	99.5	100.0	98.9	98.9	98.9	98.9	98.6	98.6	99.8	100.0	99.8	98.6	98.6	90.5	90.5	99.3	99.3	99.1	99.3	98.9	96.4	99.1	99.3	99.1	84.4	81.0	80.1	1
2	0.5		99.8	99.5	99.1	99.1	99.1	99.1	98.9	98.9	99.3	99.5	99.3	98.9	98.9	90.3	90.3	99.1	98.9	98.9	98.9	99.1	96.6	98.9	99.1	99.3	84.0	80.6	79.7	2
3	0.5	0.2		99.5	99.3	99.3	99.3	99.3	99.1	99.1	99.3	99.5	99.3	99.1	99.1	90.5	90.5	99.3	98.9	99.1	98.9	99.3	96.8	99.1	99.3	99.5	84.0	80.6	79.7	3
4	0.0	0.5	0.5		98.9	98.9	98.9	98.9	98.6	98.6	99.8	100.0	99.8	98.6	98.6	90.5	90.5	99.3	99.3	99.1	99.3	98.9	96.4	99.1	99.3	99.1	84.4	81.0	80.1	4
5	1.1	0.9	0.7	1.1		100.0	100.0	100.0	99.8	99.8	99.0	98.4	99.0	99.8	99.8	99.6	99.3	98.4	99.4	99.6	99.4	98.9	97.9	99.2	98.2	96.5	88.8	79.9	78.1	5
6	1.1	0.9	0.7	1.1	0.0		100.0	100.0	99.8	99.8	99.0	98.4	99.0	99.8	99.8	99.6	99.3	98.4	99.4	99.6	99.4	98.9	97.9	99.2	97.8	96.5	88.8	79.9	78.1	6
7	1.1	0.9	0.7	1.1	0.0	0.0		100.0	99.8	99.8	99.0	98.4	99.0	99.8	99.8	99.6	99.3	98.4	99.4	99.6	99.4	98.9	97.9	99.2	97.8	96.5	88.8	79.9	78.1	7
8	1.1	0.9	0.7	1.1	0.0	0.0	0.0		99.8	99.8	99.0	98.4	99.0	99.8	99.8	99.6	99.3	98.4	99.4	99.6	99.4	98.9	97.9	99.2	98.2	96.5	88.8	79.9	78.1	8
9	0.9	0.7	0.5	0.9	0.2	0.2	0.2	0.2		100.0	99.0	99.0	99.0	99.8	99.8	97.1	97.4	99.8	99.0	99.6	99.0	97.6	98.1	99.6	99.8	100.0	86.7	79.0	81.1	9
10	0.9	0.7	0.5	0.9	0.2	0.2	0.2	0.2	0.0		99.0	99.0	99.0	99.8	99.8	97.1	97.4	99.8	99.0	99.6	99.0	97.6	98.1	99.6	99.8	100.0	86.7	79.0	81.1	10
11	0.0	0.5	0.5	0.0	1.1	1.1	1.1	1.1	0.8	0.8		100.0	100.0	99.0	99.0	96.5	96.7	99.4	99.6	99.4	99.6	96.9	97.3	99.2	99.4	99.2	87.3	79.6	81.5	11
12	0.0	0.5	0.5	0.0	1.7	1.7	1.7	1.7	0.8	0.8	0.0		100.0	97.1	97.1	96.5	96.7	98.0	99.6	99.4	99.6	98.9	97.3	98.8	98.4	98.2	87.3	79.9	80.6	12
13	0.0	0.5	0.5	0.0	1.1	1.1	1.1	1.1	0.8	0.8	0.0	0.0		99.0	99.0	96.5	96.7	99.4	99.6	99.4	99.6	96.9	97.3	99.2	99.4	99.2	87.3	79.6	81.5	13
14	1.1	0.9	0.7	1.1	0.0	0.0	0.0	0.0	0.2	0.2	1.1	1.1	1.1		100.0	99.3	99.6	99.6	99.4	99.6	99.4	97.6	97.9	99.8	99.6	99.4	87.7	79.1	80.4	14
15	1.1	0.9	0.7	1.1	0.0	0.0	0.0	0.0	0.2	0.2	1.1	1.1	1.1	0.0		99.3	99.6	99.6	99.4	99.6	99.4	97.6	97.9	99.8	99.6	99.4	87.7	79.1	80.4	15
16	1.0	1.2	1.0	1.0	0.4	0.4	0.4	0.4	0.5	0.5	0.9	0.9	0.9	0.2	0.2		99.8	99.3	96.9	96.5	96.7	89.1	95.2	100.0	99.3	99.1	87.1	77.9	81.1	16
17	1.0	1.2	1.0	1.0	0.4	0.4	0.4	0.4	0.5	0.5	0.9	0.9	0.9	0.2	0.2	0.0		99.8	97.1	96.7	96.9	89.1	95.4	99.8	99.8	98.9	86.8	78.0	80.9	17
18	0.7	0.9	0.7	0.7	0.8	0.8	0.8	0.8	0.2	0.2	0.6	0.6	0.6	0.4	0.4	0.4	0.4		99.4	100.0	99.4	98.0	97.9	98.4	100.0	99.2	88.2	80.3	80.2	18
19	0.5	0.9	0.9	0.5	0.6	0.6	0.6	0.6	0.8	0.8	0.4	0.4	0.4	0.6	0.6	0.5	0.5	0.6		99.4	100.0	96.9	97.3	99.6	99.4	99.2	86.9	79.6	81.5	19
20	0.7	0.9	0.7	0.7	0.4	0.4	0.4	0.4	0.2	0.2	0.6	0.6	0.6	0.4	0.4	0.2	0.2	0.0	0.6		99.4	97.6	97.7	99.8	100.0	99.8	86.0	79.9	81.2	20
21	0.5	0.9	0.9	0.5	0.6	0.6	0.6	0.6	0.8	0.8	0.4	0.4	0.4	0.6	0.6	0.5	0.5	0.6	0.0	0.6		96.9	97.3	99.6	99.4	99.2	86.7	79.8	81.5	21
22	0.9	0.7	0.5	0.9	0.2	0.2	0.2	0.2	0.0	0.0	0.9	0.9	0.9	0.2	0.2	0.5	0.5	0.2	0.9	0.2	0.9		95.3	98.9	98.7	98.9	82.0	79.6	78.7	22
23	0.9	0.7	0.5	0.9	0.2	0.2	0.2	0.2	0.0	0.0	0.9	0.9	0.9	0.2	0.2	0.5	0.5	0.2	0.9	0.2	0.9	0.0		97.7	97.9	98.1	86.5	78.0	79.5	23
24	0.9	1.1	0.9	0.9	0.8	0.8	0.8	0.8	0.4	0.4	0.8	0.8	0.8	0.2	0.2	0.0	0.0	0.4	0.4	0.2	0.4	0.5	0.4		98.4	98.2	88.4	79.5	79.8	24
25	0.7	0.9	0.7	0.7	1.0	1.0	1.0	1.0	0.2	0.2	0.6	0.6	0.6	0.4	0.4	0.4	0.4	0.0	0.6	0.0	0.6	0.2	0.2	0.4		99.2	88.2	81.0	80.3	25
26	0.9	0.7	0.5	0.9	2.0	1.8	1.8	2.0	0.0	0.0	0.8	0.8	0.8	0.6	0.6	0.9	0.9	0.2	0.8	0.2	0.8	0.7	0.0	0.8	0.2		88.8	81.6	80.1	26
27	6.9	7.4	7.4	6.9	7.7	7.7	7.7	7.7	7.2	7.2	6.3	6.3	6.3	7.8	7.8	8.0	8.0	7.1	6.7	7.0	6.8	8.0	7.3	7.8	7.1	7.5		69.7	74.0	27
28	20.9	21.5	21.6	20.9	21.3	21.3	21.3	21.3	20.6	20.6	19.9	19.8	19.9	20.3	20.3	19.2	19.2	20.4	19.9	20.2	19.9	22.1	20.5	19.9	20.3	20.4	25.7		99.0	28
29	21.1	21.7	21.8	21.1	21.3	21.1	21.1	21.3	20.0	20.0	19.4	19.3	19.4	20.4	20.4	19.3	19.4	20.2	19.4	19.9	19.5	22.5	20.1	19.9	20.1	20.7	25.1	0.4		29
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	

1 L.Canicola RTCC2805
2 L.Canicola RTCC2824
3 L.Canicola RTCC2836
4 L.Canicola LRDG51
5 L.Icterohaemorrhagiae RTCC2812
6 L.Icterohaemorrhagiae RTCC2823
7 L.Icterohaemorrhagiae RTCC2837
8 L.Icterohaemorrhagiae LRDMo6
9 L.Grippotyphosa RTCC2808
10 L.Grippotyphosa RTCC2825
11 L.Pomona RTCC2815
12 L.Pomona RTCC2822
13 L.Pomona RTCC2829
14 L.Autumnalis RTCC2802
15 L.Autumnalis RTCC2830
16 L.Sejroe hardjo bovis RTCC2810
17 L.Sejro hardjo bovis RTCC2843
18 L.Sejroe Hardjoprajitno RTCC2821
19 L.Sejroe sejroe RTCC2817
20 L.Lyme RTCC2833
21 L.Djasiman RTCC2834
22 L.Pyrogenes RTCC2835
23 L.Ballum RTCC2838
24 L.Australis RTCC2840
25 L.Lai type lanylokowii RTCC2841
26 L.Bataviae RTCC2842
27 L.Malaysia RTCC2831
28 L.Celledoni RTCC2832
29 L.Javanica RTCC2839

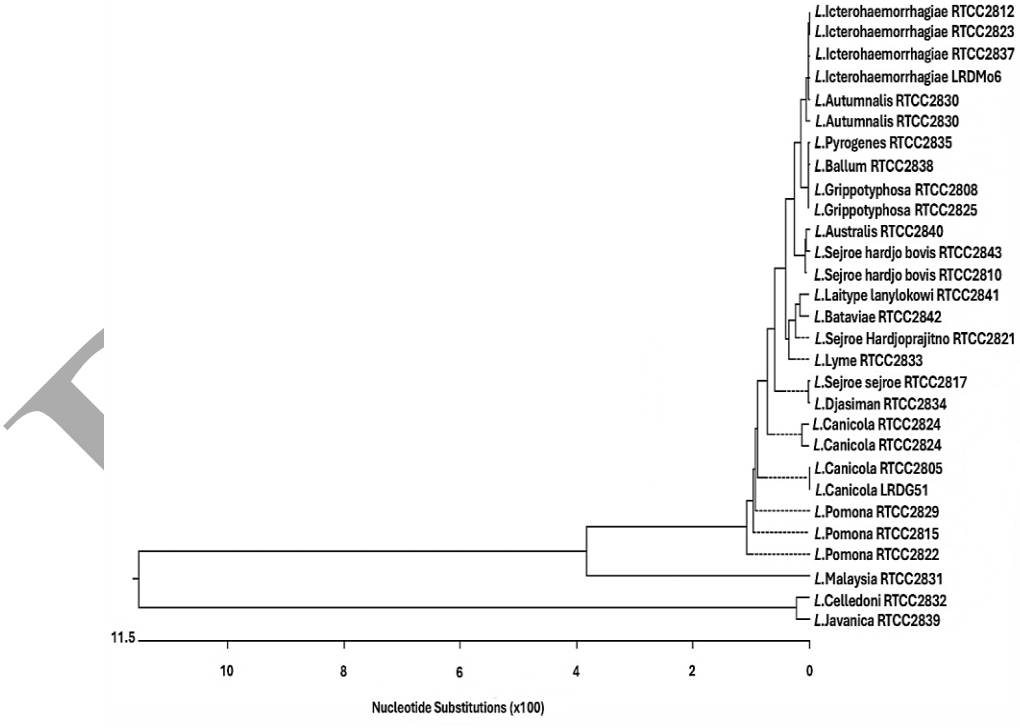

Figure1. Phylogenetic tree of 29 pathogenic serovars based on *secY* gene sequences constructed using the Clustal V algorithm in MegAlign software

3.3 Meta-Analysis of *Leptospira* Serovars with NCBI Genetic Resources

A comprehensive phylogenetic analysis of the twenty-nine *Leptospira* serovars, combined with twenty-six serovars from the NCBI gene bank, revealed a distinct separation into two major clades, as shown in Figure 2.

According to the phylogenetic tree, Clade I includes the majority of the pathogenic serovars analyzed in this study. These are grouped into well-defined subclades, showing close genetic relationships with several GenBank reference strains of *Leptospira interrogans*, *Leptospira noguchii*, and *Leptospira kirschneri*. The GenBank-derived serovars within this clade are geographically diverse, originating from Brazil, Spain, Malaysia, the United States, Australia, New Zealand, Belgium, Kazakhstan, Laos, and Sri Lanka.

Conversely, Clade II is characterized by the serovars *L.* Celledoni and *L.* Javanica, which exhibit the highest genetic divergence from the other serovars in the dataset. Despite this divergence, both serovars displayed phylogenetic relationships with distinct strains of *Leptospira borgpetersenii*, as well as with *L. santarosai*, *L. weilii*, and *L. alexanderi* included in the dataset.

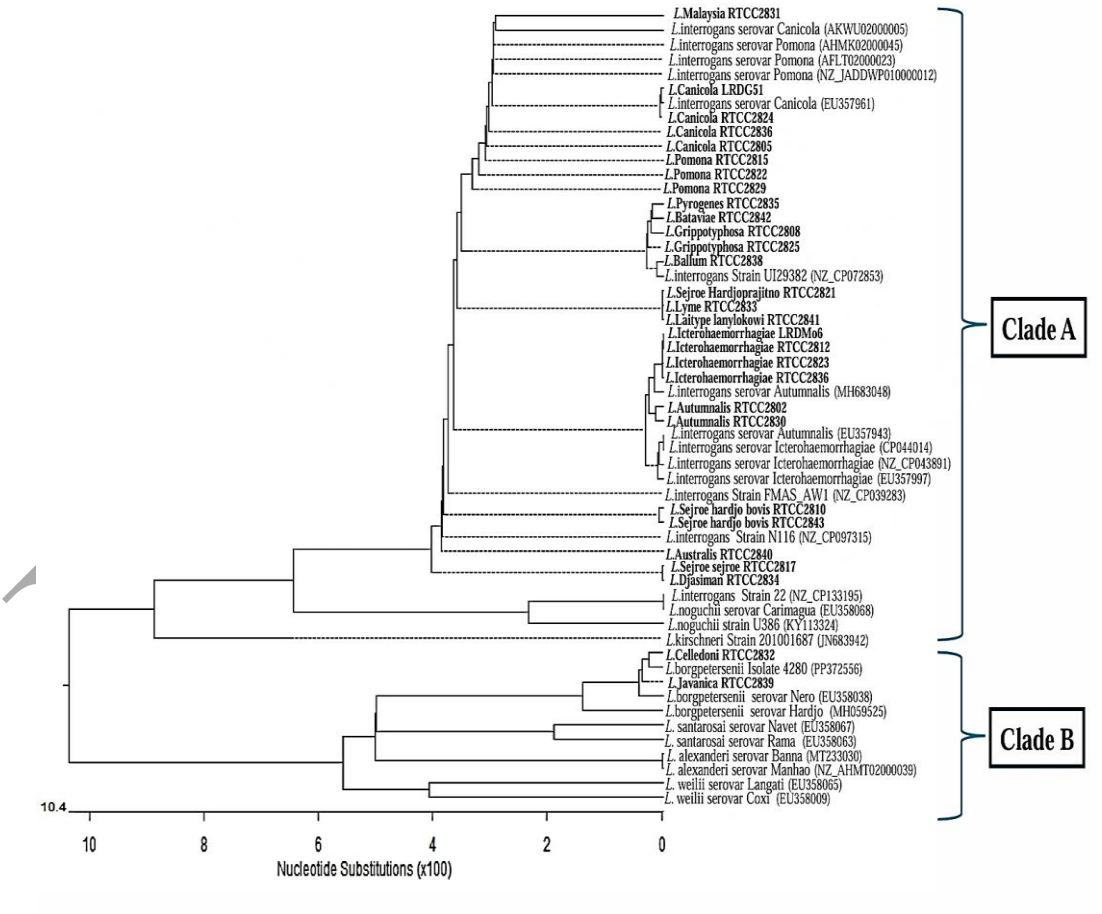

Figure2. Dendrogram illustrating the genetic relationships between 29 clinical serovars and 26 reference serovars from the NCBI database, based on *secY* gene sequence alignment using the Clustal V algorithm in MegAlign software.

3.4 Prediction of amino acid sequences

In total, 159 amino acids were identified in each of the 29 sequences studied. Analysis of the dendrogram shown in Figure 3 indicated that most of the sequenced serovars, except for three, have highly similar sequences. This similarity indicated that the protein-coding sequences and possibly the secondary and tertiary structures of these proteins are identical or very close to each other. Based on the amino acid sequences, three serovars—*L.* Celledoni, *L.* Javanica, and *L.* Malaysia —differ by up to 0.026% from the other studied serovars, placing them in a separate clade. These differences are primarily due to variations in amino acids at specific positions.

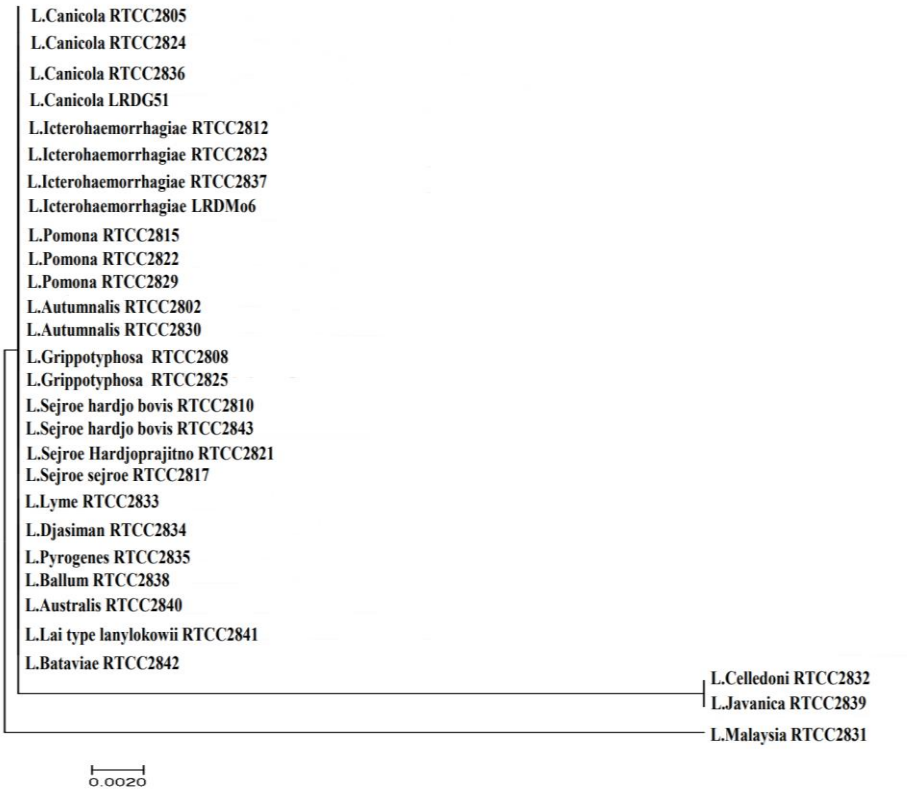

Figure 3. Dendrogram showing amino acid sequence similarity among 29 *Leptospira* serovars, generated using MEGA 11 software

4. Discussion

Leptospirosis, a neglected zoonotic infection, occurs primarily in temperate and tropical regions, especially in areas characterized by high rainfall and neutral to mildly alkaline soils, like north of Iran. The challenges associated with routine diagnosis and disease management in these regions contribute to potential underreporting, thereby limiting effective disease surveillance and control measures(5, 12). Consequently, the prompt and accurate identification of leptospirosis, as well as the differentiation between pathogenic and non-pathogenic serovars, is crucial for effective prevention, control, and treatment strategies(3).

Over the past years, several serological and molecular studies have been conducted in Iran to enhance the understanding and diagnosis of *Leptospira*. Notable advancements include the characterization of genetic patterns in *Leptospira* serovars used for vaccine production in Iran through VNTR (MLVA) analysis, the application of pulsed-field gel electrophoresis for the rapid and reliable molecular identification of pathogenic *Leptospira* species, and the cloning of genes encoding key proteins such as LipL32, OmpL1, Loa22, and LigB in *Leptospira interrogans*. Furthermore, molecular characterization of genes such as *LipL32, Loa22, LipL41, OmpL1,* and *OmpL37* has been undertaken at the Razi Vaccine and Serum Research Institute in Karaj, contributing to a deeper understanding of the genetic diversity and pathogenicity of *Leptospira* species (13, 14).

Despite these advancements, a standardized approach for accurate identification of leptospirosis in medical diagnostic laboratories and healthcare centers is still lacking.

Our study focused on the molecular characterization and phylogenetic analysis of the *secY* gene in *Leptospira* serovars

Previous researches have proposed the *secY* gene as a potential marker for the molecular identification of *Leptospira* serovars due to its strong phylogenetic discrimination among pathogenic *Leptospira* serovars(15, 16).

In this study, the *secY* gene was found exclusively in pathogenic *Leptospira* serovars and was absent in saprophytic *Leptospira* serovars. This finding aligns with earlier researchs, which also demonstrated that this gene is present only in pathogenic serovars(15, 17).

In this study, sequencing of 29 pathogenic *Leptospira* serovars revealed that different serovars of *L. interrogans* were grouped into distinct clades, while identical serovars clustered within the same subclades. The phylogenetic analysis demonstrated clear differentiation among the different serovars, indicating the strong discriminatory power of the *secY* gene.

Md. Mahtab et al. (15) investigated the molecular profiling and phylogenetic assessment of *Leptospira* species circulating in India using nucleotide sequences of the *secY* gene. Phylogenetic evaluation of four isolates demonstrated that these regional strains grouped closely with several previously reported pathogenic *Leptospira interrogans* strains.

The results of this study, aligned with the present study, support the *secY* gene as a reliable marker for both the identification and phylogenetic analysis of *Leptospira* species in leptospirosis. Similar to our findings, their study demonstrates that *Leptospira interrogans* serovars can be effectively characterized and differentiated using this marker.

Identical serovars in the present study exhibited high genetic similarities, exceeding 99.5%. For instance, three serovars of *L*. Icterohaemorrhagiae, two serovars of *L*. Grippotyphosa, three serovars of *L*. Pomona, and two serovars of *L*. Autumnalis demonstrated 100% genetic similarity. Additionally, three serovars of *L*. Canicola showed 99.8% similarity, while two

serovars of *L*. Sejroe hardjo bovis exhibited 99.7% similarity. These findings highlight the close genetic relationships among these serovars and confirm the effectiveness of the *secY* gene in categorizing them into genetically related clusters.

Nathan E. Stone et al.(18) investigated the environmental presence of pathogenic *Leptospira* spp. in Puerto Rico after the 2017 hurricanes. From 2018 to 2020, they collected soil and water samples from 22 sites and used *secY* and *lipL32* genes for detection and phylogenetic analysis. Of 86 positive samples, *secY* sequences were obtained from 32, all clustering within the pathogenic clade.

Until now, no studies have reported the molecular characterization of *Leptospira* serovars in Iran using the *secY* gene. However, further research with larger sample sizes is needed to validate these findings, particularly regarding the pathogenic *Leptospira* species found in Iran. In the future, this gene could be utilized in combination with other genes in the MLST approach for the identification of *Leptospira*.

Overall, our findings highlighted the utility of the *secY* gene as a robust molecular marker for the identification and phylogenetic differentiation of pathogenic *Leptospira* serovars in Iran.

Acknowledgments:

We would like to thank the staff of the *Leptospira* Reference Laboratory, Department of Microbiology, Razi Vaccine and Serum Research Institute, Karaj, Iran, for their technical assistance and support throughout the study.

Authors' Contributions:

Laboratory experiments, data analysis, and manuscript drafting: S.A.
Supervision; Conceptualization; Investigation; Resources; Project administration; Funding acquisition, Methodology; Validation; Writing - review & editing: P.K
Bioinformatics analysis: M.E.

Conflict of Interests:

The authors declare no conflicts of interest.

Ethics Approval:

This study did not involve human or animal subjects. All experimental procedures were conducted in accordance with the approved biosafety protocols of the Razi Vaccine and Serum Research Institute.

Funding:

This study was financially supported by Razi accine & Serum Research Institute, Karaj, Iran.

Data Availability Statement:

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

5. References

1. Wilkinson D, Edwards M, Shum C, Moinet M, Anderson N, Benschop J, et al. Molecular typing of Leptospira spp. in farmed and wild mammals reveals new host-serovar associations in New Zealand. N Z Vet J. 2024;72(1):1–9.
2. Haake DA, Levett PN. Leptospirosis in humans. Leptospira and leptospirosis. 2015:65–97.
3. Karpagam KB, Ganesh B. Leptospirosis: a neglected tropical zoonotic infection of public health importance—an updated review. Eur J Clin Microbiol Infect Dis. 2020;39(5):835–46.
4. Azevedo IR, Amamura TA, Isaac L. Human leptospirosis: In search for a better vaccine. Scand J Immunol. 2023;98(5):e13316.
5. Khalili M, Sakhaee E, Amiri FB, Safat AA, Afshar D, Esmaeili S. Serological evidence of leptospirosis in Iran; A systematic review and meta-analysis. Microb Pathog. 2020;138:103833.
6. Di Azevedo M, Lilenbaum W. An overview on the molecular diagnosis of animal leptospirosis. Lett Appl Microbiol. 2021;72(5):496–508.
7. Arent Z, Pardyak L, Dubniewicz K, Plachno B, Kotula-Balak M. Leptospira taxonomy: then and now. 2022.
8. Gravekamp C, Van de Kemp H, Franzen M, Carrington D, Schoone G, Van Eys G, et al. Detection of seven species of pathogenic leptospires by PCR using two sets of primers. Microbiology. 1993;139(8):1691–700.
9. Hamond C, Pestana CP, Medeiros MA, Lilenbaum W. Genotyping of Leptospira directly in urine samples of cattle demonstrates a diversity of species and strains in Brazil. Epidemiol Infect. 2016;144(1):72–5.
10. Ahmed N, Devi SM, De los Á Valverde M, Vijayachari P, Machang'u RS, Ellis WA, et al. Multilocus sequence typing method for identification and genotypic classification of pathogenic Leptospira species. Ann Clin Microbiol Antimicrob. 2006;5:1–10.
11. Ahmed A, Ferreira AS, Hartskeerl RA. Multilocus sequence typing (MLST): markers for the traceability of pathogenic Leptospira strains. Veterinary Infection Biology: Molecular Diagnostics and High-Throughput Strategies: Springer; 2014. p. 349–59.
12. Yasouri SR, Ghane M, Doudi M, Rezaei A, Naghavi NS. A Study of Leptospirosis Epidemiology in Iran and Diagnostic Techniques for Human, Livestock and Environment Samples. Medical Laboratory Journal. 2020;14(6).
13. Khaki P, Bagherpour M, Gharakhani M, Soltani MS, Shahcheraghi F, Nikbin VS. Application of pulsed-field gel electrophoresis for molecular identification of pathogenic Leptospira species in Iran: a rapid and reliable method. Med Lab J. 2024;16(3):323.
14. Malek Mohammadi Y, Khaki P, Gharakhani M. Molecular Characterization of Gene Encoding Outer Membrane Protein loa22 in Pathogenic Leptospira Serovars in Iran. J Trop Med. 2024;2024(1):3900663.
15. Mahtab M, Khan F, Azam M, Rizvi M, Sultan A, Shukla I, et al. Molecular characterization and phylogenetic analysis of human pathogenic Leptospira species circulating in a tertiary care hospital of Western Uttar Pradesh in India. Pathog Glob Health. 2019;113(6):275–81.
16. Dogonyaro BB, van Heerden H, Potts AD, Fasina FO, Casanovas-Massana A, Kolo FB, et al. Molecular characterization of Leptospira species detected in the kidneys of slaughtered livestock in Abattoirs in Gauteng Province, South Africa. Pathogens. 2023;12(5):666.

17. Victoria B, Ahmed A, Zuerner RL, Ahmed N, Bulach DM, Quinteiro J, et al. Conservation of the S10-spc-α locus within otherwise highly plastic genomes provides phylogenetic insight into the genus Leptospira. PLoS One. 2008;3(7):e2752.
18. Stone NE, Hall CM, Ortiz M, Hutton S, Santana-Propper E, Celona KR, et al. Divergent lineages of pathogenic Leptospira species are widespread and persisting in the environment in Puerto Rico, USA. medRxiv. 2021:2021.11. 09.21265981.