



Regional genetic diversity and COI haplotype characterization of human-derived *Enterobius vermicularis* isolates

Hilal Bedir¹ · Nurullah Çiftçi² · Murat Öztürk³ · Ayşegül Önk¹ · Masma Shahbazova⁴ · İlkey Bahçeci⁵ · Emine Baran Deniz⁶ · Neriman Mor¹ · Mehmet Ezer⁷ · Uğur Arslan⁴ · Salih Maçın⁴ · Fatih Şaban Beriş³ · Muhammet Mustafa Akiner³ · Hülya Çakmur⁸ · Mükremin Özkan Arslan¹

Received: 21 March 2026 / Accepted: 19 June 2026
© The Author(s), under exclusive licence to Springer Nature B.V. 2026

Abstract

Background *Enterobius vermicularis* (pinworm) is a globally distributed intestinal nematode that primarily infects children. Despite its high prevalence, information on the molecular diversity of human-derived *E. vermicularis* populations remains limited, particularly in geographically diverse regions such as Türkiye. This study aimed to characterize genotype and haplotype diversity based on the mitochondrial cytochrome c oxidase subunit 1 (COI) gene in *E. vermicularis* egg isolates collected from humans in different regions of Türkiye.

Methods A total of 71 microscopy-positive samples were analyzed by nested PCR, and 40 COI amplicons (397 bp) yielded high-quality sequences for further evaluation. Sequence analysis revealed 16 distinct haplotypes defined by 20 polymorphic sites. Overall haplotype diversity was high ($H_d=0.871$), whereas nucleotide diversity was low ($\pi=0.00714$).

Results Among the regional populations, the highest haplotype diversity was observed in Kars ($H_d=0.895$), while the highest nucleotide diversity was detected in Balıkesir/Edremit ($\pi=0.00927$). Maximum-likelihood phylogenetic analysis demonstrated that all Türkiye haplotypes belonged to Genotype B and formed three subclusters together with previously reported human-derived reference sequences from Europe and the Middle East. Seven haplotypes showed no 100% match in GenBank and were therefore considered novel, whereas the remaining haplotypes were identical to genotype B sequences previously reported from Greece, Bulgaria, Iran, Iraq, and Türkiye.

Conclusions To our knowledge, this study presents the first COI-based haplotype dataset of *E. vermicularis* egg isolates from multiple regions of Türkiye and demonstrates substantial intraspecific genetic diversity despite the predominance of a single mitochondrial genotype. These findings expand current knowledge of the genetic structure of human-derived *E. vermicularis* in Türkiye and provide a basis for future large-scale molecular epidemiological studies. Further studies including larger sample sizes and additional genetic markers are needed.

Keywords *Enterobius vermicularis* · COI · Haplotype diversity · Phylogenetic analysis

✉ Hilal Bedir
bedirhilal@gmail.com

¹ Faculty of Medicine, Department of Medical Parasitology, Kafkas University, Kars, Türkiye

² Faculty of Medicine, Department of Microbiology, Kafkas University, Kars, Türkiye

³ Faculty of Science, Department of Biology, Recep Tayyip Erdoğan University, Rize, Türkiye

⁴ Faculty of Medicine, Department of Microbiology, Selçuk University, Konya, Türkiye

⁵ Faculty of Medicine, Department of Microbiology, Recep Tayyip Erdoğan University, Rize, Türkiye

⁶ Faculty of Medicine, Department of Public Health, Kafkas University, Kars, Türkiye

⁷ Faculty of Medicine, Department of Urology, Kafkas University, Kars, Türkiye

⁸ Faculty of Medicine, Department of Family Medicine, Kafkas University, Kars, Türkiye

Introduction

Enterobius vermicularis is one of the most common intestinal nematodes infecting humans, with children being particularly susceptible [1–3]. The global prevalence in children has been estimated at approximately 12.9%, with infection rates peaking between 5 and 14 years of age [4]. In Türkiye, the pooled prevalence of *E. vermicularis* infection among school-aged children was reported to be approximately 9% [5]. Low socioeconomic status, poor personal hygiene, inadequate sanitation, overcrowding, and close contact in settings such as households, schools, and daycare centers constitute major risk factors for transmission [6]. Clinically, enterobiasis is most commonly associated with perianal pruritus and irritation, but may also present with recurrent abdominal pain, nasal itching, nocturnal bruxism, appendicitis, eosinophilic enterocolitis, vaginal itching, and sleep disturbances, potentially affecting cognitive performance and child development in chronic infections [2, 7].

E. vermicularis has a direct life cycle, with humans serving as the definitive host [1]. It is transmitted mainly through contaminated fingers, fomites, and close interpersonal contact [6]. As transmission is closely tied to repeated exposure and host contact patterns, local circulation may be influenced not only by hygiene but also by household clustering and human mobility [6, 8]. In addition, *E. vermicularis* is considered an ancient human-associated parasite with a long evolutionary history alongside humans and closely related primates [9–11]. Previous phylogenetic studies suggest that *E. vermicularis* phylogenetic structure broadly reflects host-associated evolutionary history [12, 13]. The combination of these epidemiological and evolutionary features makes *E. vermicularis* a useful model for investigating host-associated parasite diversification and the evolutionary consequences of human migration across geographic space [9–11].

Recent molecular studies have demonstrated that analyses based on the mitochondrial cytochrome c oxidase subunit 1 (COI) gene reveal substantial intraspecific genetic diversity within *E. vermicularis* populations [13–20]. Owing to its relatively high substitution rate, maternal inheritance, and lack of recombination, COI is a widely used marker in population genetic and phylogeographic studies [21–24]. To date, COI-based phylogenetic analyses have identified three major genotypic lineages in *E. vermicularis*, designated as genotypes A, B, and C. Genotypes A and B have been reported in humans and chimpanzees, whereas genotype C has thus far been detected only in chimpanzees [13, 15, 17, 18]. From a geographic perspective, genotypes A and C appear to be predominantly associated with Asian, while genotype B has been consistently reported as the dominant lineage in human isolates originating from Europe and the

Middle East [14–16, 20, 25–28]. This broad phylogeographic pattern suggests that the population structure of *E. vermicularis* may be shaped by host movement and regional connectivity [17].

The Anatolian Peninsula has formed a natural land bridge between Africa, Asia, the Middle East, and Eurasia, enabling the dispersal of terrestrial biota, including humans, over millions of years. Its unique geographic position has also made the region a favorable corridor and settlement area for human populations [29]. These historical and demographic processes may also have shaped the population structure of human-specific parasites. Host mobility, settlement patterns, and regional connectivity represent key determinants of gene flow and phylogeographic structuring, particularly in helminths transmitted through direct contact. Situated at the intersection of Asia, Europe, and the Middle East, Türkiye represents an important geographic transition zone for investigating the genetic structuring of *E. vermicularis* populations. However, although genotype B has been reported as the predominant lineage in human infections in Türkiye, COI-based data at the haplotype level remain limited [20]. Accordingly, the present study aims to characterize the genetic diversity of *E. vermicularis* by analyzing COI sequences obtained from parasite eggs recovered from cellophane tape preparations collected across different geographic regions of Türkiye.

Materials and methods

Study design, sampling sites and specimen collection

In this study, *E. vermicularis* egg-positive cellophane tape specimens were obtained from both hospital-based and community-based settings in order to capture regional variation in human infections across Türkiye. Between January 2020 and December 2024, a total of 71 cellophane tape preparations were collected from children aged 0–18 years across different regions of Türkiye, including pediatric patients presenting to hospitals with gastrointestinal complaints and elementary school children screened during community-based surveys. Specimens were obtained from four regions (Fig. 1): Eastern Anatolia, Kars province (Kafkas University Health Research and Application Hospital, n=33), Black Sea, Rize province (Recep Tayyip Erdoğan University Training and Research Hospital, n=5), Central Anatolia, Konya province (Selçuk University Faculty of Medicine Hospital, n=11), and Marmara, Balıkesir/Edremit district (elementary school students, n=22) [30]. Only specimens found to be positive for *E. vermicularis* eggs by light microscopy were included in the molecular analyses.



Fig. 1 Map of Türkiye showing the provinces from which the samples were collected. The size of the dots is approximately proportional to the number of samples examined in each location

After microscopic confirmation, egg-positive areas on the adhesive tape were identified, marked, and stored at -20°C until DNA extraction.

Ethical approval

Ethical approval was obtained from the Clinical Research Ethics Committee of Kafkas University (Approval date: 23 April 2025; Approval no.: KAÜ-TFEK 770).

DNA extraction

Genomic DNA was extracted from *E. vermicularis* egg-positive cellophane tape samples using the EcoPURE Genomic DNA Kit (EcoTech Biotechnology, Türkiye), according to the manufacturer's instructions, with minor modifications for cellophane tape materials. Before extraction, all samples were re-examined microscopically, and tape regions showing the highest apparent egg density were selected for processing, and were marked. Egg-positive samples were classified into three categories based on the number of eggs observed per microscopic field of view at $400\times$ total magnification (+; 1–10 eggs), moderate density (++; 11–20 eggs), and high density (+++; >20 eggs). The marked tape sections (approximately 1–2 cm²) were excised and transferred into sterile 1.5 mL microcentrifuge tubes. Each tube received 200 μL of Resuspension Buffer, 200 μL of Tissue Lysis Buffer, and 20 μL of Proteinase K, followed by incubation at 56°C for 24 h. Subsequently, 20 μL RNase A was added, mixed briefly, and incubated at room temperature for 3 min. After the addition of 400 μL Binding Buffer, the lysates were incubated at 70°C for 10 min and then processed according to the manufacturer's column-based purification protocol. DNA was eluted in 50 μL of Elution Buffer and stored at -20°C until downstream molecular analyses.

Nested PCR method

The COI gene of *E. vermicularis* was amplified using a nested PCR approach, following the protocol described by Piperaki et al. [14]. The outer PCR used primers EVM1 (5'-TTTT TGGTCATCCTGAGGTTTATATTC-3') and EVM2 (5'-CC ACATTATCCAAAATAGGATTAGCC-3') (expected product ~390 bp); The inner PCR used primers EVIF (5'-TTGG TCATCCTGAGGTTTATATTC-3') and EVIR (5'-TCCAA AATAGGATTAGCCAACA-3') (~379 bp). The first round PCR was performed in a total volume of 50 μL containing 5 μL of $10\times$ PCR reaction buffer, 2.5 μL of MgCl_2 (50 mM stock MgCl_2), 1 μL of dNTP mix (10 mM stock), 0.25 μL of each primer (100 μM stock), 0.25 μL of Taq DNA polymerase (5 U/ μL stock), 5 μL of the DNA extraction sample, and sterile distilled water to a final volume of 50 μL . Thermal cycling consisted of an initial denaturation at 94°C for 5 min; 40 cycles of 94°C for 1 min, 57°C for 1 min, and 72°C for 1 min; followed by a final extension at 72°C for 5 min. Thermal cycling was performed on a thermocycler (Stratagene, La Jolla, USA) with an initial denaturation at 94°C for 5 min, followed by 45 cycles of denaturation at 94°C for 1 min, annealing at 57°C for 1 min, and extension at 72°C for 1 min, with a final extension at 72°C for 10 min. For the nested PCR reaction, 2 μL of the first round PCR product was used as template, and amplification conditions were identical to those of the outer PCR, except that the annealing temperature was reduced to 53°C and the number of cycles was set to 35. Positive controls consisted of DNA from previously confirmed egg-positive cellophane tape material, while sterile water was used as a negative control. Then, PCR products were separated by electrophoresis on 2% (w/v) agarose gels containing ethidium bromide (0.5 $\mu\text{g}/\text{mL}$) (Invitrogen Life Technologies, USA), run at 100–120 V for 30–45 min, and visualized under UV illumination using a UVP Transilluminator system (Analytik Jena,

USA). Amplicon sizes were assessed using a 100 bp DNA ladder (EcoTech Biotechnology, Türkiye).

DNA sequencing and data analyses

PCR products obtained from the second round of nested PCR were sequenced bidirectionally by MacroGen Inc. (Amsterdam, The Netherlands). Each cellophane tape specimen was processed and sequenced individually, and no pooling was performed across different samples. Raw chromatograms were inspected in both forward and reverse directions, low-quality terminal regions were trimmed, and ambiguous base calls were checked manually during sequence editing. Only high-quality bidirectional reads were retained for downstream analyses. The edited sequences were analyzed in MEGA7, and all sequences were compared with reference sequences using BLASTn at the NCBI database. Sequences showing <97% identity to reference *E. vermicularis* COI sequences were excluded. Resulting sequences were aligned using the ClustalW algorithm implemented in MEGA7. Protein-coding COI sequences were translated into amino acids using the invertebrate mitochondrial genetic code in ExPASy Translate to verify reading-frame integrity and detect premature stop codons. All sequences were trimmed to a uniform length. The final trimmed COI fragment corresponded to 397 bp, yielding a translated alignment of 133 amino acid positions. Sequences exhibiting premature stop codons were retained for comparative analyses and flagged for further evaluation. Sequence length, number of haplotypes (H), number of segregating sites (S), haplotype diversity (Hd), and nucleotide diversity (π) were calculated for each sampling province using DnaSP v5.10. Pairwise genetic differences among isolates from different provinces were calculated using Arlequin v3.5.1.2. Relationships among haplotypes were inferred using median-joining network analysis implemented in NETWORK v10.2 (Bandelt et al., 1999). Phylogenetic relationships among isolates

were inferred using the maximum likelihood (ML) method implemented in MEGA7. The best-fit nucleotide substitution model was selected using the model selection tool implemented in MEGA7. The phylogenetic tree was constructed with 1000 bootstrap replicates to assess nodal support. Bootstrap support values were displayed on the final tree at the corresponding internal nodes. *Enterobius anthropitheci* (AB254450), *Trypanoxyuris atelis* (AB222177), *Trypanoxyuris microon* (AB222176), and *Ascaris suum* (X54253) were included as outgroups. The final ML phylogenetic tree was visualized using FigTree v1.4.4.

Results

Molecular and sequence analysis

A total of 71 microscopy-positive cellophane tape specimens containing *E. vermicularis* eggs was successfully amplified using a nested PCR assay targeting the COI gene. A clearly visible COI band was obtained after nested PCR in 40 of the 71 cellophane tape samples (56.3%), whereas no visible band was detected in the remaining 31 samples (43.7%) (Fig. 2). 40 PCR amplicons yielded high-quality bidirectional sequence data and were included in subsequent analyses. These sequences originated from four provinces in Türkiye: Kars (n=15), Konya (n=10), Balıkesir/Edremit (n=10), and Rize (n=5). All validated COI sequences generated in this study were deposited in the GenBank database under accession numbers: PX931745–PX931760.

Genetic variation and haplotype diversity

Analysis of 40 COI sequences revealed 20 polymorphic sites, all of which corresponded to transition substitutions, resulting in the identification of 16 distinct haplotypes. At the provincial level, Kars exhibited the highest haplotype

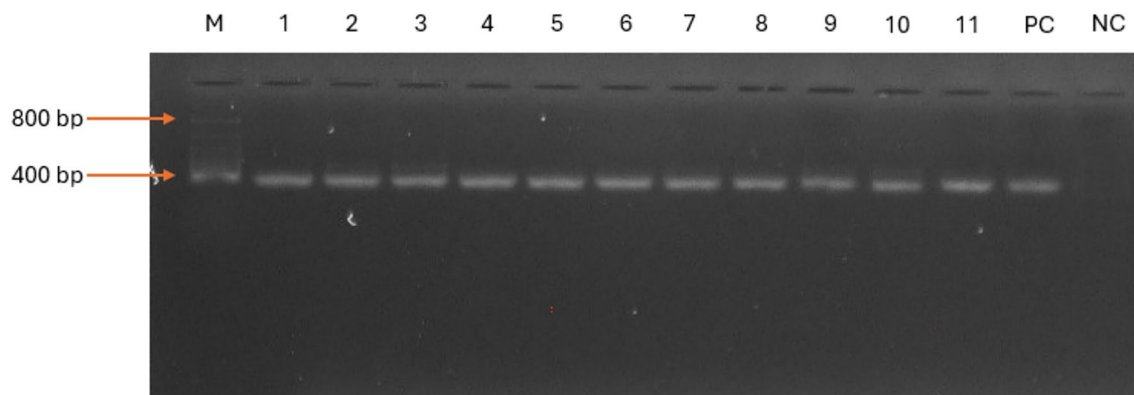


Fig. 2 Agarose gel electrophoresis of nested PCR products targeting the mitochondrial COI gene of *E. vermicularis* obtained from cellophane tape samples. The expected amplicon size was 397 bp. M:

100 bp DNA ladder; lanes 1–11: representative *E. vermicularis* egg-positive samples; PC: positive control; NC: no-template negative control

Table 1 Genetic diversity indices and neutrality test results of *E. vermicularis* COI sequences by province in Türkiye

Isolates	Genetic diversity						Neutrality analyses	
	n	S	H	Hd	π	k	Tajima's D	Fu's Fs
Rize	5	3	3	0.700	0.00302	1.20	-1.0485	-0.1859
Balıkesir (Edremit)	10	8	4	0.733	0.00927	3.689	1.3136*	2.0926
Konya	10	10	6	0.778	0.00581	2.311	-1.5345	-1.4587
Kars	15	13	10	0.895	0.00603	2.4	-1.5717	-5.2031*
Total	40	20	16	0.871	0.00714	2.844	-0.7103	-1.1888

n, number of sequences; S, number of segregating sites; H, number of haplotypes; Hd, haplotype diversity; π , nucleotide diversity; k, average number of nucleotide differences. Tajima's D and Fu's Fs were used to assess deviations from neutrality. Statistically significant values ($P < 0.05$) are indicated in bold and marked with an asterisk (*)

Hap_1	FWFFGHPEV	YILILPAFGI	VSHRILCLTG	KKEVFGLHGM	IYAIISIGLI	GRVVGHHMF	TIGFDIRTL	YFMVATIIIA	VPTGVKVFWS	LLTLIGGRLV	VQPLLFWVLG	FIFMFTTGGI	TGIMLANPIL	DNV
Hap_10														
Hap_11														
Hap_12														
Hap_13														
Hap_14														
Hap_15														
Hap_16														
Hap_2														
Hap_3														
Hap_4														
Hap_5														
Hap_6														
Hap_7														
Hap_8														
Hap_9														

Fig. 3 Amino acid sequence alignment of the translated mitochondrial COI fragment (133 aa) for *E. vermicularis* haplotypes (Hap_1–Hap_16) identified from isolates collected in Kars, Konya, Balıkesir (Edremit), and Rize, Türkiye. Hap_1 is shown as the reference

diversity ($Hd=0.895$) together with the highest number of haplotypes ($H=10$), indicating substantial within-population haplotypic variation. By contrast, Balıkesir/Edremit showed the highest nucleotide diversity ($\pi=0.00927$), suggesting comparatively greater sequence divergence among the haplotypes detected in that population. Konya displayed intermediate values for both haplotype diversity ($Hd=0.778$) and nucleotide diversity ($\pi=0.00581$), whereas Rize showed the lowest nucleotide diversity ($\pi=0.00302$) and relatively limited haplotypic variation ($H=3$; $Hd=0.700$). Overall, the combination of high haplotype diversity and low nucleotide diversity across the full dataset suggests the presence of multiple closely related haplotypes rather than deeply divergent mitochondrial lineages. Across the full dataset, total haplotype diversity was $Hd=0.871$ and the mean number of nucleotide differences was $k=2.844$ (Table 1).

To evaluate demographic patterns at the provincial level, Tajima's D and Fu's Fs neutrality tests were performed (Table 1). Negative Tajima's D and Fu's Fs values were observed in all populations except Balıkesir/Edremit. The Kars exhibited a statistically significant negative Fu's Fs value ($Fs=-5.20$; $P<0.05$), whereas the Balıkesir/Edremit showed a positive and statistically significant Tajima's D value ($D=1.3136$; $P<0.05$) (Table 1).

Amino acid alignment inferred from the mitochondrial COI gene region comprised 16 haplotypes across 133 amino acid positions (Fig. 3). Using Haplotype 1 as the reference sequence, the majority of positions were conserved across

sequence; dots (.) indicate amino acids identical to the reference, and letters denote amino acid substitutions. An asterisk (*) indicates a premature stop codon in the translated fragment

haplotypes. Protein-level variation was limited and primarily consisted of single-site amino acid substitutions. Specifically, a cysteine (C) substitution was observed in Haplotype 11; Haplotypes 3 and 12 shared a serine (S) substitution within the same alignment block; a glycine (G) substitution was detected in Haplotype 14; and an alanine (A) substitution was observed in Haplotype 4. Haplotype 7 exhibited two substitutions (A and V), whereas Haplotype 8 displayed a single valine (V) substitution. Haplotypes 2, 9, 10, 13, and 16 were identical to Haplotype 1 across all 133 amino acid positions. Premature stop codons (*) were detected at the start of the translated fragment in Haplotypes 5, 6, and 15, and these sequences were therefore flagged for further verification.

Pairwise genetic distance analysis

Pairwise genetic distances among provincial populations were estimated using the pairwise difference method (Table 2). The highest and statistically significant genetic distance was observed between the Balıkesir/Edremit and Konya (0.2611; $P<0.05$). A significant level of genetic differentiation was also detected between Balıkesir/Edremit and Kars (0.1908; $P<0.05$), whereas the genetic distance between Balıkesir/Edremit and Rize was lower and not statistically significant (0.1754; $P>0.05$). In contrast, comparisons among the Konya, Kars, and Rize yielded very low pairwise genetic distance values. Slightly negative values,

Table 2 Pairwise genetic distances among *E. vermicularis* COI populations by province in Türkiye

	Balıkesir (Edremit)	Konya	Kars	Rize
Balıkesir (Edremit)	0.0000			
Konya	0.2611*	0.0000		
Kars	0.1908*	-0.0159	0.0000	
Rize	0.1754	-0.0738	-0.0762	0.0000

Statistically significant values ($P < 0.05$) are indicated in bold and marked with an asterisk (*)

which are generally interpreted as zero, were obtained in some comparisons (Konya–Kars: -0.016 ; Konya–Rize: -0.074 ; Kars–Rize: -0.076 ; all $P > 0.05$).

Haplotype network and geographic distribution

The distribution of COI sequences obtained from 40 *E. vermicularis* isolates across 16 haplotypes is illustrated in a median-joining haplotype network (Fig. 4), and the provincial distribution of haplotypes is presented in Table 3. Haplotype 10 (PX931746) was the most frequent and was shared by 13 isolates obtained from Kars ($n=5$), Konya ($n=5$), and Rize ($n=3$). Haplotype 2 (PX931753) included four isolates from Balıkesir/Edremit ($n=2$), Kars ($n=1$), and Konya ($n=1$). Haplotype 4 (PX931755) comprised three isolates from Balıkesir/Edremit ($n=2$) and Kars ($n=1$), whereas Haplotype 6 (PX931757) was shared by one isolate each from Kars ($n=1$), Konya ($n=1$), and Rize ($n=1$). Among the Balıkesir/Edremit isolates, five clustered in Haplotype 1 (PX931745) and one in Haplotype 3 (PX931754). The

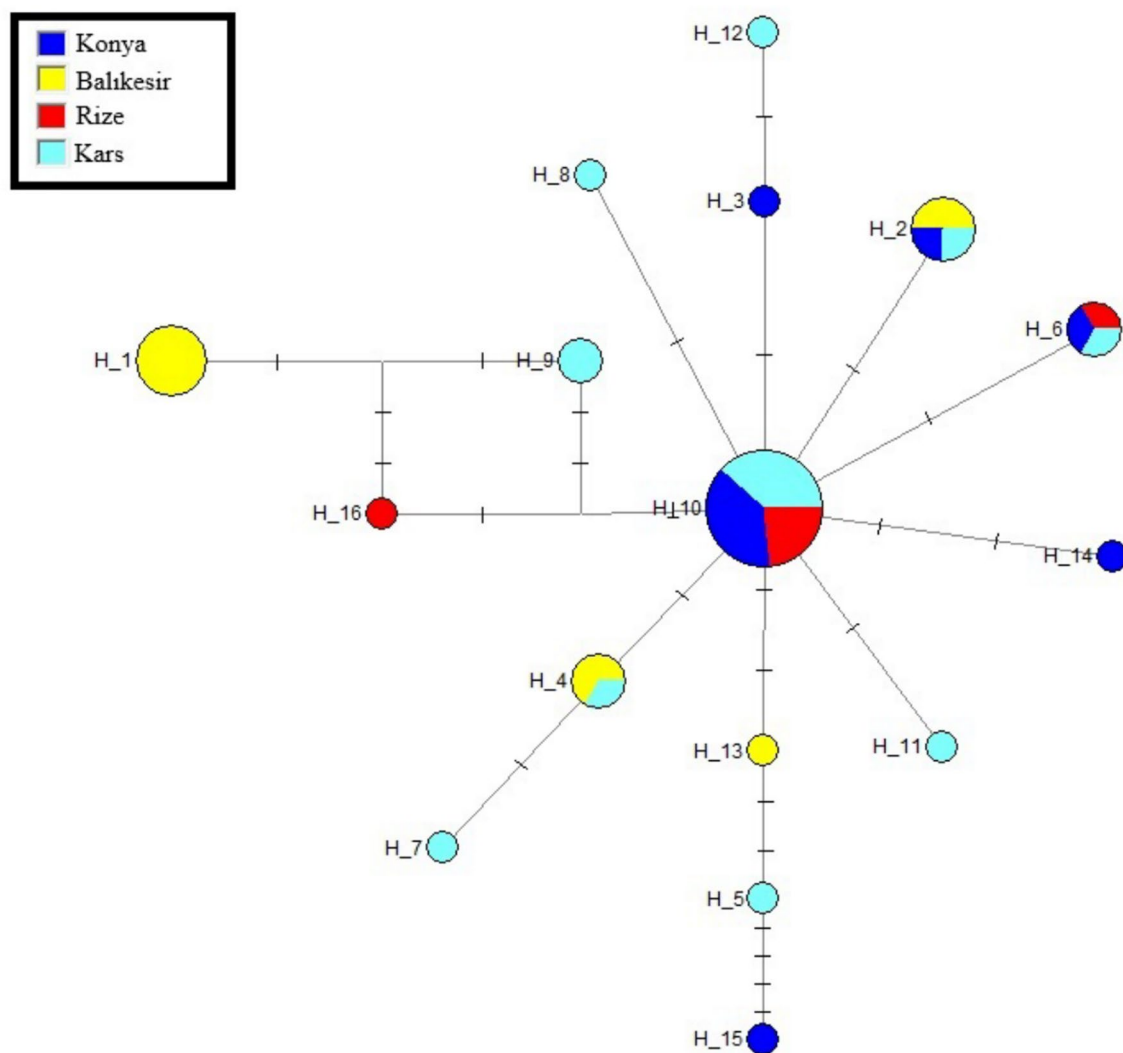


Fig. 4 Median-joining haplotype network of *E. vermicularis* COI sequences constructed using PopART v1.7. The analysis was based on a 397 bp fragment of the COI gene obtained from 40 isolates, resulting in the identification of 16 haplotypes. Each circle represents a distinct

haplotype, with circle size proportional to the number of isolates. Colored segments within circles indicate the province of origin of the isolates. Short tick marks along the connecting lines denote the number of mutational steps separating haplotypes

Table 3 Provincial distribution of COI-based haplotypes identified in *E. vermicularis* isolates

Sample information					Samples with 100% sequence identity
Haplotype	Sample N	Sample no	City	Genbank accession no	
Haplotype 1	5	BE1	Balıkesir	PX931745	PX210567 (96), PV131525 (93)
		BE6	Balıkesir	PX931745	
		BE7	Balıkesir	PX931745	
		BE9	Balıkesir	PX931745	
		BE10	Balıkesir	PX931745	
Haplotype 2	2	BE2	Balıkesir	PX931753	PX210624 (96), PX210606 (96), PX210602 (96), PX210584 (96), PX210556 (96), MH802610 (96), MH802609 (96), MH802608 (96), MH802607 (96), MH802606 (96), MH802601 (96), MH802600 (96), MH802596 (96), LC819594 (96), PX210601 (96), PX210571 (96), PX210745 (96), PX210572 (96), MH802602 (96), MH802597 (96), PX210552 (95), MH802611 (95), MZ360958 (95), PX210709 (95), MH802595 (95), OL773362 (95), OL773361 (95), OL773359 (95), MH802593 (95)
		BE4	Balıkesir	PX931753	
	1	KR13	Kars	PX931753	
	1	KN33	Konya	PX931753	
Haplotype 3	1	BE3	Balıkesir	PX931754	
Haplotype 4	2	BE5	Balıkesir	PX931755	
		BE8	Balıkesir	PX931755	
Haplotype 5	1	KR11	Kars	PX931755	
	1	KR12	Kars	PX931756	KJ780776 (93), OL773357 (93), OL773356 (93), OL773355 (93)
Haplotype 6	1	KR14	Kars	PX931757	PX210624 (96), PX210606 (96), PX210602 (96), PX210584 (96), PX210556 (96), MH802610 (96), MH802609 (96), MH802608 (96), MH802607 (96), MH802606 (96), MH802601 (96), MH802600 (96), MH802596 (96), LC819594 (96), PX210601 (96), PX210571 (96), PX210745 (96), PX210572 (96), MH802602 (96), MH802597 (96)
	1	KN35	Konya	PX931757	
	1	RZ39	Rize	PX931757	
Haplotype 7	1	KR15	Kars	PX931758	-
Haplotype 8	1	KR16	Kars	PX931759	PX210698 (96), PV131533 (93)
Haplotype 9	2	KR17	Kars	PX931760	PV131526 (93)
		KR20	Kars	PX931760	
Haplotype 10	5	KR18	Kars	PX931746	PX210624 (96), PX210606 (96), PX210602 (96), PX210584 (96), PX210556 (96), MH802610 (96), MH802609 (96), MH802608 (96), MH802607 (96), MH802606 (96), MH802601 (96), MH802600 (96), MH802596 (96), LC819594 (96), PX210601 (96), PX210571 (96), PX210745 (96), PX210572 (96), MH802602 (96), MH802597 (96), PX210552 (95), MH802611 (95), MZ360958 (95), PX210709 (95), MH802595 (95), OL773362 (95), OL773361 (95), OL773359 (95), MH802593 (95)
		KR19	Kars	PX931746	
		KR21	Kars	PX931746	
		KR22	Kars	PX931746	
		KR24	Kars	PX931746	
		KN28	Konya	PX931746	
		KN29	Konya	PX931746	
	5	KN30	Konya	PX931746	
		KN32	Konya	PX931746	
		KN34	Konya	PX931746	
		RZ36	Rize	PX931746	
		RZ38	Rize	PX931746	
		RZ40	Rize	PX931746	
Haplotype 11	1	KR23	Kars	PX931747	
Haplotype 12	1	KR25	Kars	PX931748	
Haplotype 13	1	KN26	Konya	PX931749	
Haplotype 14	1	KN27	Konya	PX931750	
Haplotype 15	1	KN31	Konya	PX931751	
Haplotype 16	1	RZ37	Rize	PX931752	

Province codes in sample IDs indicate origin: BE=Balıkesir (Edremit), KR=Kars, KN=Konya, RZ=Rize

two remaining Kars isolates grouped within Haplotype 9 (PX931760). All other haplotypes were represented by single isolates, including Haplotype 5 (PX931756), Haplotype 7 (PX931758), Haplotype 8 (PX931759), Haplotype 11 (PX931747), and Haplotype 12 (PX931748). In Konya, three isolates each corresponded to a unique haplotype Haplotype 13 (PX931749), Haplotype 14 (PX931750), and Haplotype 15 (PX931751), whereas the remaining Rize isolate was assigned to Haplotype 16 (PX931752).

Phylogenetic analysis

Maximum-likelihood (ML) phylogenetic analysis based on COI sequences revealed three major genotypes, designated as A, B, and C, within *E. vermicularis* (Fig. 5). Bootstrap support values based on 1000 replicates are shown at the corresponding internal nodes in Fig. 5. The basal separation of the Genotype B cluster from the Genotype A/C lineages was strongly supported (bootstrap=99%), whereas several internal branches within Genotype B showed low-to-moderate support. Accordingly, the phylogenetic interpretation was focused primarily on genotype-level clustering and the broad placement of Türkiye haplotypes rather than weakly supported internal relationships. All 16 haplotypes identified in the present study clustered exclusively within Genotype B. Within Genotype B, the Türkiye haplotypes were divided into three subclades together with human-derived reference sequences. The largest subclade included Haplotype 1–14 and Haplotype 16 and grouped predominantly with human isolates reported from Bulgaria, along with additional references from Greece, Iran, Iraq, the Czech Republic, Denmark, and Poland. Haplotype 15 formed a separate subclade with reference sequences from Türkiye, Bulgaria, and Greece. A third subclade included additional reference sequences from Iran, Greece, and Bulgaria and also contained Türkiye sequences. In contrast, reference sequences from Thailand, South Korea, and Japan, including non-human-primate-associated isolates, formed distinct and more distant clades corresponding to Genotypes A and C, clearly separated from the Genotype B cluster.

Comparison of the 16 haplotypes identified in this study with GenBank reference sequences showed that several haplotypes exhibited 100% nucleotide identity with previously published isolates from Iran, Bulgaria, Iraq, and Türkiye (Table 3). In contrast, seven haplotypes (Haplotype 3: PX931754, Haplotype 4: PX931755, Haplotype 7: PX931758, Haplotype 11: PX931747, Haplotype 12: PX931748, Haplotype 13: PX931749, and Haplotype 14: PX931750) showed no complete nucleotide identity with any GenBank reference sequence and were therefore classified as novel haplotypes. Nevertheless, these haplotypes generally differed from their closest reference sequences by

Fig. 5 Maximum-likelihood (ML) phylogenetic tree based on COI sequences (397 bp) of *E. vermicularis* haplotypes obtained from egg-derived DNA from cellophane tape samples collected across provinces of Türkiye. The tree was inferred in MEGA under the best-fit substitution model (HKY+G). Bootstrap support values were calculated using 1000 replicates and are shown at internal nodes; numbers at nodes indicate bootstrap percentages. *Enterobius anthropopitheci* (AB254450), *Trypanoxyuris atelis* (AB222177), *Trypanoxyuris microon* (AB222176), and *Ascaris suum* (X54253) were used as outgroups. The scale bar indicates substitutions per site, and 100%-identical GenBank references are annotated with accession number, host, and country

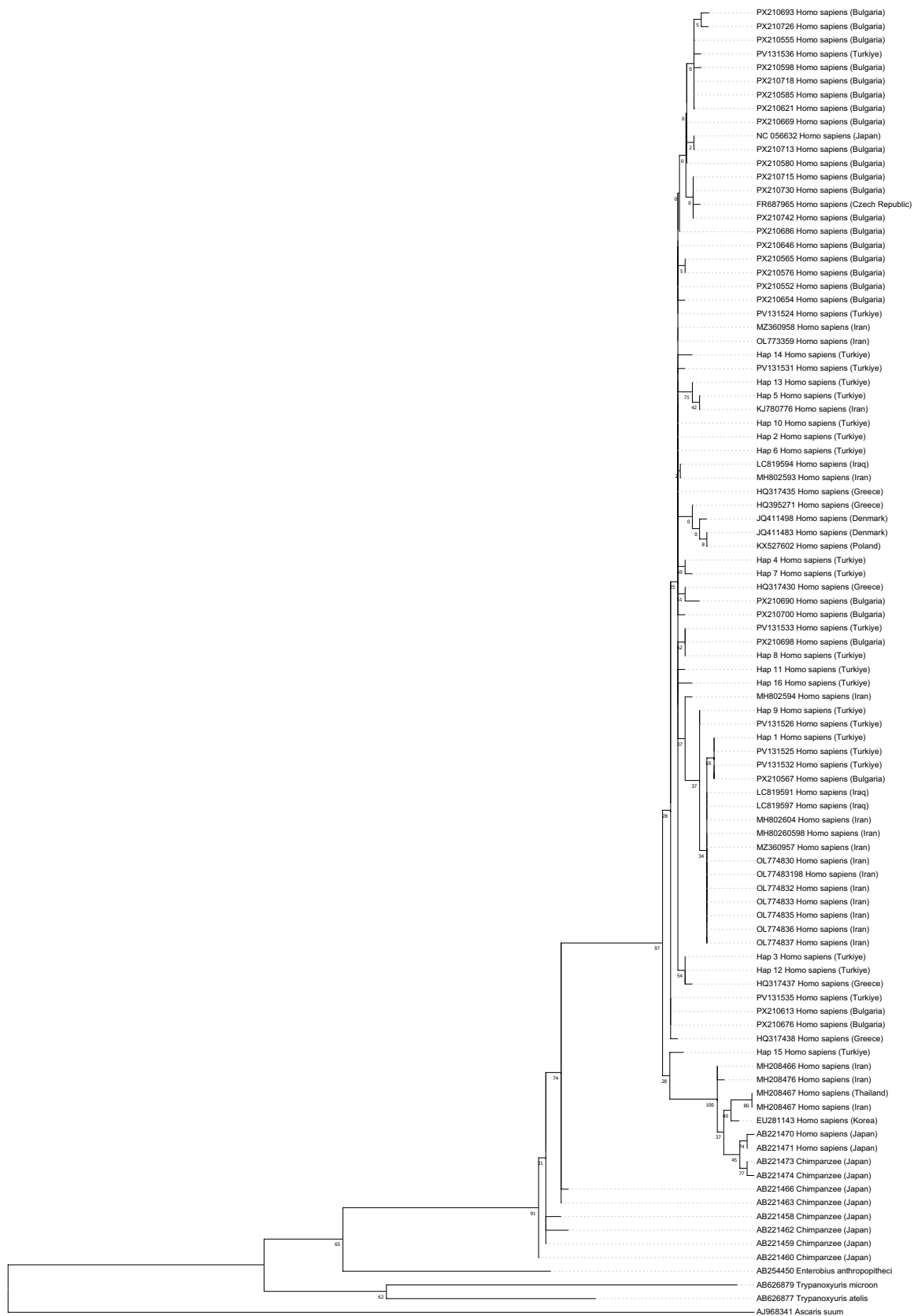
a single nucleotide substitution, exhibiting $\geq 99.7\%$ nucleotide identity. Based on their phylogenetic placement, most novel haplotypes were positioned within genotype B close proximity to reference sequences originating from Iran and Greece, whereas Haplotype 4 was embedded within a Bulgaria-dominated reference assemblage. In contrast, Haplotype 11 was located on a branch adjacent to a small reference cluster comprising sequences from Bulgaria (PX210742), the Czech Republic (FR687965), and Japan (NC056632) (Fig. 5).

Discussion

In this study, the genetic diversity of *E. vermicularis* DNA isolated from eggs obtained from cellophane tape specimens collected across different geographic regions of Türkiye was analyzed using the COI gene. PCR amplification of the COI gene was successfully amplified from cellophane tape specimens with high egg density, yielding high-quality amplicons. However, as previously reported, the reduced amplification success in samples with low egg density may be attributable to low DNA concentration, DNA degradation, prolonged storage under unsuitable conditions, and PCR inhibitors associated with the cellophane tape matrix, which may contain fecal residues [14, 17, 20].

Analysis of 40 COI sequences identified 16 haplotypes based on 20 polymorphic sites, all of which were transition substitutions. The combination of high haplotype diversity ($H_d=0.871$) and low nucleotide diversity ($\pi=0.00714$) indicates substantial mitochondrial haplotype variation with limited sequence divergence among haplotypes. Similarly, high haplotype diversity ($H_d=0.89167$) has been reported from northeastern Iran despite a lower number of haplotypes [26, 28]. This pattern may reflect demographic processes consistent with host mobility; additionally, the relatively high substitution rate of mitochondrial COI may enhance haplotype resolution and thereby contribute to the observed diversity [21].

Regarding the distribution of the identified haplotypes, the data showed that a substantial proportion of the samples clustered into a limited number of predominant haplotypes,



whereas most haplotypes were represented by single isolates. The occurrence of shared haplotypes across provinces suggests that some *E. vermicularis* mitochondrial lineages may circulate beyond a single local population. In particular, the detection of Haplotype 10 in Kars, Konya, and Rize, and of Haplotypes 2, 4, and 6 in more than one province, indicates that these haplotypes are not geographically restricted within the present dataset. This pattern may reflect regional population connectivity shaped by host movement, household- or community-level transmission, and repeated circulation of common lineages. However, because the present analysis is based on a limited number of samples and a single mitochondrial marker, these findings should be interpreted cautiously and cannot on their own be taken as definitive evidence of direct gene flow or specific historical dispersal events. Larger-scale studies incorporating broader sampling and additional nuclear and mitochondrial markers will be necessary to clarify the processes underlying haplotype sharing. In particular, the inter-provincial sharing of Haplotype 10 (Kars $n=5$; Konya $n=5$; Rize $n=3$), Haplotype 2 (Balıkesir $n=2$; Kars $n=1$; Konya $n=1$), Haplotype 4 (Balıkesir $n=2$; Kars $n=1$), and Haplotype 6 (Kars $n=1$; Konya $n=1$; Rize $n=1$) suggests the possibility of a common infection source associated with host mobility (Table 3). By contrast, the predominance of singleton haplotypes (excluding Haplotypes 1 and 9) indicates that low-frequency local variants are common and that haplotype composition is heterogeneous among provinces (Table 3). Comparable patterns have also been reported from Greece, Iran and Türkiye; the sharing of the same haplotype among human isolates supports the possibility of a common infection source [14, 19, 20]. Whereas more localized populations with limited mobility may exhibit lower genetic variation [19]. Moreover, it has been reported that a COI sequence obtained from a child in Greece is identical to that of an infected chimpanzee isolate from Japan [14]. Nevertheless, the occurrence of multiple haplotypes within each province and the sharing of predominant haplotypes across provinces imply weak geographic structuring at the COI level. Such patterns may reflect host movement, but could also be influenced by the evolutionary dynamics of COI, as previously discussed [14, 21]. Taken together, these findings suggest that predominant haplotypes may spread locally and regionally via host mobility, and that larger-scale sampling and the inclusion of additional mitochondrial and/or nuclear markers are needed to more reliably explain geography-independent haplotype sharing.

All 16 haplotypes identified in this study clustered within Genotype B. The presence of Genotype B in human isolates was first reported in Greece and has subsequently been documented in studies from Europe and the Middle East [14–16, 20, 25–28]. To date, COI-based phylogenetic studies have

identified three major genotypic clusters in *E. vermicularis*, designated as genotypes A, B, and C. Genotypes A and B have been reported from both human and chimpanzee isolates, whereas genotype C has so far been detected exclusively in chimpanzees [13, 15, 17, 18]. Nakano et al. [13] suggested that Genotype B likely represents a human-associated lineage and may have been introduced into captive chimpanzees through human-to-chimpanzee transmission. Previous data from Türkiye similarly indicated that local isolates belong to this widespread human-associated lineage [20], and our results further support the predominance of Genotype B in Türkiye. However, the 16 Türkiye haplotypes were resolved into three subclusters within Genotype B and grouped with previously reported human-derived reference sequences from Bulgaria, Greece, Iran, Iraq, the Czech Republic, Denmark, and Poland (Fig. 5). Importantly, our findings suggest that the Türkiye isolates show close phylogenetic affinity with human-derived Genotype B sequences reported from Europe and the Middle East, and that the observed geographic sharing of haplotypes may reflect host movement. This pattern indicates close phylogenetic relatedness between the Türkiye isolates and Genotype B lineages circulating in neighboring regions. The detection of identical or closely related haplotypes in Türkiye and adjacent countries (Iran, Iraq, and Greece) is consistent with regional connectivity that may be facilitated by the Anatolian Peninsula's biogeographic position as a natural land bridge between Africa, Asia, the Middle East, and Eurasia [29]. In this context, the population structure of *E. vermicularis* in Türkiye may have been influenced by circulation processes linked to human mobility and migration.

This study has some limitations. The analyses were conducted using a limited number of cellophane tape samples obtained from four geographically distinct provinces of Türkiye (Kars, Konya, Balıkesir/Edremit, and Rize), therefore the observed haplotype richness and diversity should not be directly generalized to the national population. When the provincial patterns are compared in more detail, Kars appears to harbor the richest haplotypic composition, with the highest Hd value and the largest number of haplotypes, whereas Balıkesir/Edremit is distinguished by the highest π value, reflecting greater nucleotide-level divergence among fewer haplotypes. In contrast, Rize showed lower haplotype richness and the lowest nucleotide diversity, suggesting a comparatively more homogeneous local population. These findings indicate that different provinces contribute differently to the overall genetic structure: some populations are characterized mainly by haplotype richness, whereas others show relatively greater sequence divergence among haplotypes. Nevertheless, given that COI-based data from Türkiye have previously been restricted to a single local focus, this study provides the first multi-regional COI haplotype

dataset and contributes a baseline framework for assessing the scope of genetic diversity at the country scale. Future studies incorporating broader geographic coverage and more balanced sample sizes across provinces, together with additional mitochondrial and nuclear markers, are needed to more reliably evaluate genotyping patterns in *E. vermicularis*.

Conclusion

In conclusion, this study provides the first multi-regional mitochondrial COI-based haplotype dataset of human-derived *Enterobius vermicularis* isolates from Türkiye. Analysis of egg-derived DNA obtained from cellophane tape samples revealed substantial intraspecific genetic diversity, with 16 haplotypes identified among 40 sequenced isolates despite the exclusive predominance of Genotype B. The coexistence of shared and province-specific haplotypes, together with the clustering of all Türkiye isolates within three subclades of Genotype B, suggests a pattern of regional circulation shaped by host mobility and local transmission dynamics. The close phylogenetic relationship between the Türkiye isolates and previously reported human-derived sequences from Europe and the Middle East further supports the geographic connectivity of this lineage. Although the sample size and geographic coverage remain limited, the present findings expand the current molecular epidemiological knowledge of *E. vermicularis* in Türkiye and provide a useful baseline for future studies. Broader sampling from additional provinces and the inclusion of complementary mitochondrial and nuclear markers will be important to better resolve population structure, transmission patterns, and evolutionary relationships in this human-associated parasite.

Acknowledgements All procedures involving human participants were conducted in accordance with the ethical standards of the institutional and/or national research committee and with the principles of the Declaration of Helsinki. Freely given, informed consent to participate in the study was obtained from all individual participants included in the research. For participants under 16 years of age, informed consent was obtained from a parent or legal guardian.

Author contribution H. B., N. Ç., M. Ö., U. A., S. M.,: methodology, sample collection, data assessment, writing original draft, supervision, conceptualization A. Ö., M. S., İ. B., N. M., M. E., H. Ç., M. Ö. A.: sample collection E. B. D., N. M., H. Ç.,: writing original draft.

Data availability The mitochondrial COI sequences generated in this study have been deposited in the GenBank database under accession numbers PX931745–PX931760. Additional data supporting the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interests Authors have no conflict of interests. All procedures involving human participants were conducted in accordance with the ethical standards of the institutional and/or national research committee and with the principles of the Declaration of Helsinki. Freely given, informed consent to participate in the study was obtained from all individual participants included in the research. For participants under 16 years of age, informed consent was obtained from a parent or legal guardian. Ethical approval was obtained from the Clinical Research Ethics Committee of Kafkas University (KAÜ-TFEK 770).

References

1. Cook GC (1994) *Enterobius vermicularis* infection. Gut 35:1159–1162. <https://doi.org/10.1136/GUT.35.9.1159>
2. Dudlová A, Juriš P, Jarčuška P et al (2018) The incidence of pinworm (*Enterobius vermicularis*) in pre-school and school aged children in the Eastern Slovakia. Helminthol 55:275–280. <https://doi.org/10.2478/HELM-2018-0030>
3. Jin J, Ruan W, Xu W et al (2025) *Enterobius vermicularis* infection and its risk factors among rural schoolchildren in Zhejiang Province: a cross-sectional study. Parasitol Res. <https://doi.org/10.1007/S00436-025-08525-W>
4. Lashaki EK, Mizani A, Hosseini SA et al (2023) Global prevalence of enterobiasis in young children over the past 20 years: a systematic review and meta-analysis. Osong Public Health Res Perspect. <https://doi.org/10.24171/J.PHRP.2023.0204>
5. Halidi AG, Yaran K, Aydemir S et al (2025) Prevalence of intestinal parasites in school-age children in Turkey: a systematic review and meta-analysis. PLoS Negl Trop Dis 19:e0013186. <https://doi.org/10.1371/JOURNAL.PNTD.0013186>
6. Burkhart CN, Burkhart CG (2005) Assessment of frequency, transmission, and genitourinary complications of enterobiasis (pinworms). Int J Dermatol 44:837–840. <https://doi.org/10.1111/J.1365-4632.2004.02332.X;PAGEGROUP=STRING:PUBLICAT>
7. Wendt S, Trawinski H, Schubert S et al (2019) Diagnostik und Therapie des Madenwurmbefalls. Dtsch Arztebl Int 116:213–219. <https://doi.org/10.3238/ARZTEBL.2019.0213>
8. Lee CY, Cheng PC, Ho CH et al (2019) Immunologic evaluation of extracted intestinal proteins from *Angiostrongylus cantonensis* adult worms. J Microbiol Immunol Infect 52:620–629. <https://doi.org/10.1016/j.jmii.2017.07.008>
9. Hugot JP, Reinhard KJ, Gardner SL, Morand S (1999) Human enterobiasis in evolution: origin, specificity and transmission. Parasite 6:201–208. <https://doi.org/10.1051/PARASITE/1999063201>
10. Araújo A, Reinhard K, Ferreira LF et al (2013) Paleoparasitology: the origin of human parasites. Arq Neuropsiquiatr 71:722–726. <https://doi.org/10.1590/0004-282X20130159>
11. Ferreira, Luiz Fernando, Karl Jan Reinhard, and Adauto Araújo (2014) Foundations of Paleoparasitology. <https://doi.org/10.7476/9788575415986>
12. Brooks DR, Ferrao AL (2005) The historical biogeography of co-evolution: Emerging infectious diseases are evolutionary accidents waiting to happen. J Biogeogr 32:1291–1299. <https://doi.org/10.1111/J.1365-2699.2005.01315.X;PAGE=STRING:ARTICLE/CHAPTER>
13. Nakano T, Okamoto M, Ikeda Y, Hasegawa H (2006) Mitochondrial cytochrome c oxidase subunit I gene and nuclear rDNA regions of *enterobius vermicularis* parasitic in captive chimpanzees with special reference to its relationship with pinworms in

- humans. *Parasitol Res.* <https://doi.org/10.1007/S00436-006-0238-4>
14. Piperaki ET, Spanakos G, Patsantara G et al (2011) Characterization of *Enterobius vermicularis* in a human population, employing a molecular-based method from adhesive tape samples. *Mol Cell Probes* 25:121–125. <https://doi.org/10.1016/J.MCP.2011.03.005>
 15. Hasegawa H, Sato H, Torii H (2012) Redescription of *Enterobius* (*Enterobius*) *macaci* Yen, 1973 (Nematoda: Oxyuridae: Enterobiinae) based on material collected from wild Japanese Macaque, *Macaca fuscata* (Primates: Cercopithecidae). *J parasitol.* <https://doi.org/10.1645/GE-28671>
 16. Kubiak K, Dzika E, Pauksztó (2017) Enterobiasis epidemiology and molecular characterization of *Enterobius vermicularis* in healthy children in north-eastern Poland. *Helminthologia* 54:284–291. <https://doi.org/10.1515/HELM-2017-0042>
 17. Janthu P, Dumidae A, Subkrasae C et al (2022) Prevalence and genetic analysis of *Enterobius vermicularis* in schoolchildren in lower northern Thailand. *Parasitol Res.* <https://doi.org/10.1007/S00436-022-07626-0>
 18. Tomanakan K, Sanpool O, Chamavit P et al (2020) Genetic variation of *Enterobius vermicularis* among schoolchildren in Thailand. *J Helminthol* 94:e7. <https://doi.org/10.1017/S0022149X18000962>
 19. Haghshenas M, Koosha M, Latifi A et al (2023) Detection of *Enterobius vermicularis* in archived formalin-fixed paraffin-embedded (FFPE) appendectomy blocks: It's potential to compare genetic variations based on mitochondrial DNA (cox1) gene. *PLoS ONE* 18:e0281622. <https://doi.org/10.1371/JOURNAL.PONE.0281622>
 20. Gülbudak H, Ülger ST, Bozok T, Aslan G (2025) Epidemiological and molecular characterization of *Enterobius vermicularis* cases in Mersin, Turkey: An analysis based on mitochondrial DNA cox1 gene sequences. *One Health* 20:101079. <https://doi.org/10.1016/j.onehlt.2025.101079>
 21. Blouin MS (2002) Molecular prospecting for cryptic species of nematodes: mitochondrial DNA versus internal transcribed spacer. *Int J Parasitol* 32:527–531. [https://doi.org/10.1016/S0020-7519\(01\)00357-5](https://doi.org/10.1016/S0020-7519(01)00357-5)
 22. Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *Proc Royal Society B Biol Sci* 270:313–321. <https://doi.org/10.1098/RSPB.2002.2218>
 23. Holterman M, Van Der Wurff A, Van Den Elsen S et al (2006) Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown clades. *Mol Biol Evol* 23:1792–1800. <https://doi.org/10.1093/MOLBEV/MSL044>
 24. Traversa D, Kuzmina T, Kharchenko VA et al (2008) Haplotypic variability within the mitochondrial gene encoding for the cytochrome c oxidase 1 (cox1) of *Cylicocyclus nassatus* (Nematoda, Strongylida): Evidence for an affiliation between parasitic populations and domestic and wild equid hosts. *Vet Parasitol* 156:241–247. <https://doi.org/10.1016/J.VETPAR.2008.05.031>
 25. Ferrero MR, Röser D, Nielsen HV et al (2013) Genetic variation in mitochondrial DNA among *Enterobius vermicularis* in Denmark. *Parasitol* 140:109–114. <https://doi.org/10.1017/S0031182012001308>
 26. Tavan A, Mikaeili F, Sadjjadi SM et al (2020) Prevalence and genotype distribution of *Enterobius vermicularis* among kindergarteners in Shiraz and Khorramabad cities. *Iran Asian Pac J Trop Med* 13:308–313. <https://doi.org/10.4103/1995-7645.280229>
 27. Zelck UE, Bialek R, Weiß M (2011) Molecular phylogenetic analysis of *Enterobius vermicularis* and development of an 18s ribosomal DNA-targeted diagnostic PCR. *J Clin Microbiol* 49:1602–1604. <https://doi.org/10.1128/JCM.02454-10>
 28. Shilanabadi KF, Derakhshan FK, Raeghi S (2023) Genotype distribution of *Enterobius vermicularis* isolates from northwest provinces of Iran. *Ethiop J Health Sci* 33:433–440. <https://doi.org/10.4314/EJHS.V33I3.6>
 29. Ayteke Aİ, Harvati K (2016) The human fossil record from Turkey. *Vertebrate Paleobiol Paleanthropol.* https://doi.org/10.1007/978-94-024-0874-4_6
 30. Türk Mikrobiyoloji Cemiyeti Dergisi. <https://tmc.dergisi.org/text.php3?id=850>. Accessed 16 Apr 2026

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.