Analysis of COI and ITS2 regions of DNA obtained from *Paragonimus westermani* eggs in ancient coprolites on Joseon dynasty mummies

Jong Ha Hong¹, Chang Seok Oh¹,², Min Seo³, Dong Hoon Shin¹,²/⁴

¹Seoul National University College of Medicine, Laboratory of Bioanthropology, Paleopathology and History of Diseases, Seoul, South Korea
²Seoul National University College of Medicine, Institute of Forensic Science, Seoul, South Korea
³Dankook University College of Medicine, Department of Parasitology, Cheonan, South Korea

The genetic information of ancient *Paragonimus westermani*, the oriental lung fluke infecting over 20 million people worldwide, has not been thoroughly investigated thus far. We analysed genetic markers (COI and ITS2) of *P. westermani* from coprolite specimens (n = 6) obtained from 15th to 18th century Korean mummies. Our results indicated that all *P. westermani* sequences were generally distinct from the other species of the genus *Paragonimus*. The sequences were clustered into three groups: Group I for East Asia; Group II for South and Southeast Asia; and Group III for India and Sri Lanka. In this study, we found that ancient *P. westermani* sequences in Korea belong to Group I, adding invaluable information to the existing knowledge of *Paragonimus* paleogenetics.

Key words: ancient parasite - *Paragonimus westermani* - ancient DNA

Paragonimiasis occurs in definitive hosts after ingestion of infected intermediate hosts. In general, *Paragonimus* eggs hatch before entering snails (the first intermediate hosts). When the cercariae penetrate crustaceans such as crayfishes or crabs (the second intermediate hosts), they encyst in gills, muscles, and viscera, developing into metacercariae. After definitive hosts (e.g. humans) eat the freshwater crustaceans raw, the metacercariae excyst in the small intestine, penetrating the intestinal wall. They then traverse the diaphragm, entering the peribronchiolar tissues of the lung where they mature into adults within 8-12 weeks.²

The genus *Paragonimus* constitutes a species-rich group from the tropical regions in South and Southeast Asia, Africa, and Central and South America as well as the temperate zone of North America and East Asia.³ Almost 50 nominal species and subspecies of *Paragonimus* have been reported so far,¹ and approximately 16 species have been revealed to cause human diseases. Paragonimiasis mainly affects the lung. The adult worms stimulate the formation of worm capsules which then ulcerate, potentially leading to clinical symptoms such as cough, blood-tinged sputum, and pulmonary pain. Paragonimiasis could also be induced by the migration of young adult worms to the ectopic organs.²,⁴,⁵,⁶,⁷ Among *Paragonimus*, the most common species is *P. westermani*, the oriental lung fluke mainly reported in Korea, China, Taiwan, Japan, and the Philippines.¹,⁸,⁹ Scholars have estimated that over 20 million people are currently infected with *P. westermani* worldwide.¹,¹⁰

Recently, researchers have attempted to reveal the genetic characteristics of *P. westermani* through DNA analysis.¹,³,⁸,¹¹,¹²,¹³,¹⁴,¹⁵ Phylogenetic analyses of cytochrome c oxidase subunit I (COI) and internal transcribed spacer 2 (ITS2) DNA regions have revealed that *P. westermani* are clustered into at least two groups (East Asia and South/Southeast Asia), in association with geographically distinct distributions.¹,³,⁸,¹¹,¹²,¹³ However, although ancient eggs have been detected in archaeological samples, very few parasitological reports regarding paleogenetics of ancient *Paragonimus* spp. have been published so far. The only report was our previous study on ITS2 DNA sequence of ancient *P. westermani* eggs obtained from a 17th century Korean mummy.¹⁷ In that study, the ancient DNA sequence of *P. westermani* was very similar to that of modern *P. westermani* reported in East Asia, but was genetically distinct from the *P. westermani* sequences of Southeast Asia.¹⁷

Nevertheless, our previous study was performed with only one sample and a single genetic marker (ITS2). Over the past several years, the genetic information of another *P. westermani* genetic marker (COI) has been continuously accumulated by multiple researchers.¹¹,¹²,¹³,¹⁶,¹⁷ Our paleoparasitological studies have also microscopically detected ancient *Paragonimus* eggs in many more coprolite specimens (n = 6) from the 15th to 18th century Korean mummies.¹⁸,¹⁹ In this regard, we attempted to examine multiple genetic markers (COI and ITS2) in newly collected ancient *P. westermani* eggs in order to obtain more comprehensive information about the evolutionary history of *P. westermani*.

The samples used in this study were acquired from the 15th to 18th century Joseon-era mummies (Cheongdo, Dangjin, Hwasung, YG2-4, YG2-6, and Yongin).
The specimens were ancient faeces from the intestines of mummies (Cheongdo and Hwasung) or the organic materials compiled upon the hipbones of half-mummified bodies (Dangjin, YG2-4, YG2-6 and Yongin).

To authenticate our work on ancient DNA (aDNA), we followed the Criteria of Authentication established by Hofreiter et al. (20). During aDNA analysis, we wore head caps, masks, protection gloves, and gowns. All the tools used in this study were sterilized before use. We also used specialised facilities that were exclusively dedicated to aDNA analysis. (21) We performed experiments in this facility that were developed in accordance with the Criteria. (20) The Institutional Review Board (IRB) of Seoul National University Hospital confirmed that our aDNA analyses could be exempt from board review (IRB No. 2017-001). We also followed the Vermillion Accord on Human Remains adopted by the World Archaeological Congress. (22)

For DNA extraction, we followed the method described in our previous report. (23) In brief, the samples (0.3 g each) were incubated at 56°C in 1 mL of lysis buffer (pH 8.0; including 1 % SDS; 0.1M DTT; 50 mM of EDTA; 1 mg/mL of proteinase K) for 24 h. After DNA was extracted with an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1), it was then treated with chloroform/isoamyl alcohol (24:1). DNA isolation and purification were performed using a QIAmp PCR purification kit (Qiagen, Hilden, Germany). The purified DNA was eluted in 40 μl of EB elution buffer (Qiagen, Hilden, Germany). Primers for the *P. westermani* COI and ITS2 regions were generated by Integrated DNA Technologies, Inc. (Iowa City, IA, USA). The information of the primers used in our study is summarised in Table I. DNA quantification was performed using a NanoDrop™ ND-1000 Spectrophotometer (Thermo Fisher Scientific, MA, USA). Extracted DNA (10 μL) was treated at 37°C with 1 unit of uracil-DNA-glycosylase (New England Biolabs, MA, USA) for 30 min. It (40 ng) was then mixed with the reagent premix containing 10 pmol of each primer and 1X AmpliTaq Gold® 360 Master Mix (Life Technologies, CA, USA). PCR conditions were as follows: pre-denaturation at 95°C for 10 min, 45 cycles of denaturation at 95°C for 30 s, annealing at 54-60°C for 30 s, extension at 72°C for 30 s, and final extension at 72°C for 10 min. The polymerase chain reaction (PCR) products were separated on 2.5% agarose gel (Invitrogen, CA, USA), and then stained with ethidium bromide. Negative controls (extraction controls) were also applied to the electrophoresis at the same time. Electrophoresis results were photographed using a Vilber Lourmat ETX-20.M equipped with BioCapt software (Vilber Lourmat, Collégien, France).

### Table I

List of primers used for amplification of *Paragonimus westermani* DNA in this study

| Region | Set | Primer | 5' to 3' | Annealing temp. (ºC) | Length (bp) |
|--------|-----|--------|----------|----------------------|-------------|
| COI    | COI1 | COI-1F | GGG CAT CCG GAG GTG TAT GT | 54 | 106 |
|        |      | COI-1R | TTC GGG TAC TAC GGG CTG G  | | |
|        | COI2 | COI-2F | CTG ACC AAC AAC GAT TCC T  | 54 | 150 |
|        |      | COI-2R | TCC CGT GAC AGA ACT AAG AA  | | |
|        | COI3 | COI-3F | GTC TGG GTA GTG TGT TGG  | 54 | 136 |
|        |      | COI-3R | AGC ATG AAC AAC CAA GAG AA  | | |
|        | COI4 | COI-4F | TTA GTT CTG TCA CGG GAG TG  | 57 | 114 |
|        |      | COI-4R | GAA TTC ACC ACA AAA CAG GA  | | |
|        | COI5 | COI-5F | TTC TCT TGG TTT TTC ATG CT  | 54 | 183 |
|        |      | COI-5R | GAC GTA ATG AAA ATG AGC C  | | |
|        | ITS2-1 | ITS2-1F | GCG CAG CCA ACT GTG TGA A  | 57 | 133 |
|        |      | ITS2-1R | GGC GTC GCG ATA GTT TAT  | | |
|        | ITS2-2 | ITS2-2F | TTA ATG CGA ACT GCA TAC TG  | 54 | 169 |
|        |      | ITS2-2R | AAG ACC AGA TGG GGG AGA T  | | |
|        | ITS2-3 | ITS2-3F | GGT CGG CTT ATA AAC TAT CG  | 54 | 129 |
|        |      | ITS2-3R | CCC GAG TAT GGT AGG GAA A  | | |
|        | ITS2-4 | ITS2-4F | AAT CTT GTA TGG TGc CTG T  | 60 | 165 |
|        |      | ITS2-4R | AAA CCA CAG ATG AAG ACA G  | | |
|        | ITS2-5 | ITS2-5F | GTG GCT CAG TGA ATG ATT TAT  | 54 | 170 |
|        |      | ITS2-5R | CCG CTT AGT GAT ATG CCT A  | | |
The PCR amplicons were isolated using a QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany). The bacteria were transformed using the pGEM-T Easy Vector system (Promega Corporation, Madison, USA). Transformed bacteria were grown on an agar plate containing X-GAL (40 μg/μL), ampicillin (50 μg/mL), and 0.5 mM IPTG for the 14 h. After selected colonies were grown in LB media for 12 h, the purity of cultured bacteria was performed using a QIAprep® Spin Miniprep kit (Qiagen, Hilden, Germany).

Each amplified strand was sequenced using an ABI Prism BigDye Terminator v3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Waltham, USA) and the 3730xl Automatic Sequencer (Applied Biosystems, Waltham, USA). The obtained DNA sequences were aligned by MEGA7 program. The consensus sequences were compared to those available in GenBank using NCBI/BLAST tools. The web browser module and Alignment Explorer of MEGA7 were used to retrieve sequences homologous to those of interest from NCBI GenBank database.

The evolutionary relationship of *P. westermani* and other taxa from NCBI GenBank was inferred from the Phylogeny Reconstruction analysis implemented in MEGA7. We used the Maximum Likelihood (ML) method. Selected parameters of our method followed the Tamura-Nei model for COI or the Kimura 2-parameter model for ITS2; we selected Complete Deletion for *P. siamensis* and the 3730xl Automatic Sequencer (Applied Biosystems, Waltham, USA). The obtained DNA sequences were aligned by MEGA7 program. The consensus sequences were compared to those available in GenBank using NCBI/BLAST tools. The web browser module and Alignment Explorer of MEGA7 were used to retrieve sequences homologous to those of interest from NCBI GenBank database.

To obtain the consensus sequences of COI and ITS2 regions, we repeated cloning and sequencing for each specific amplicon. From these trials, 9-12 clone sequences were successfully acquired for the COI (390 bp) and ITS2 (417 bp) regions. Next, multiple sequence alignment was performed using Clustal W implemented in MEGA7. We then obtained consensus sequences after alignment of the cloned sequences [Supplementary data (Fig. 1)]. The consensus sequences of the *P. westermani* COI and ITS2 regions from the Cheongdo and YG2-4 mummies were identical [Supplementary data (Fig. 2)].

Using NCBI/BLAST tools, we compared the consensus sequences with those available in GenBank [Supplementary data (Fig. 2)]. The BLAST results are summarised in Table II. Briefly, the current *P. westermani* COI sequences obtained from Joseon mummies exhibited significant similarities to the *P. westermani* COI sequences reported in Korea (AF333280.1; AF540958.1), Japan (U97205.1; U97208.1), China (AY140680.1; U97209.1), Vietnam (FJ434988.1), India (JN656169.1; KM280646.1), the Philippines (U97213.1), Thailand (AB354224.1), and Sri Lanka (AY240940.1). We also found moderate similarities to the GenBank sequences of *P. siamensis* (AB354231.1), *P. paishuihoensis* (AB679289.1), and *P. mexicanus* (KC562301.1) [Supplementary data (Fig. 2A); Table II].

In the case of the ITS2 region, the current Joseon *P. westermani* sequences exhibit similarities to *P. westermani* ITS2 sequences found in Korea (AF333277.1), Japan (U96907.1), China (KC471492.1; AB713404.1), India (AB938198.1; JN656208.1; DQ836243.1), Taiwan (U96908.1), Vietnam (LC144492.1; FJ434982.1), Malaysia (U96909.1), Thailand (AB354217.1; AF159604.1), and Sri Lanka (AY240942.1). Similar sequences were also found in the sequences of *P. siamensis* (AB354222.1), *Euparagonimus cenocopiosus* (AF159601.1), and *P. skrjabini* (AB703444.1). The *P. westermani* ITS2 sequence obtained from a Korean mummy reported in our previous study (Shin et al. 2010) also demonstrated 100% identity (coverage 87%) [Supplementary data (Fig. 2B); Table II].

Next, we performed phylogenetic analysis of *P. westermani* COI and ITS2 regions (Fig. 2). Every *P. westermani* taxon in the COI and ITS2 regions was clearly distinct from other *Paragonimus* species (*P. siamensis, P. skrjabini, P. paishuihoensis and P. mexicanus*). *P. westermani* sequences were clustered into several groups. In the case of the COI region, Group I included *P. westermani* sequences from East Asia (Korea, Japan, and China) whereas Group II included those from South and Southeast Asia (India, Philippines, and Thailand) and formed another cluster. We note that some sequences reported from India and Sri Lanka formed a separate cluster of the third separate group (Group III) (Fig. 2A). *P. westermani*
COI sequences obtained from Korean mummies in this study evidently belong to the East Asia group. Our result was similar to those of the phylogenetic analyses conducted by Iwagami et al.\(^{(13,17)}\) and Devi et al.,\(^{(8,16)}\) which yielded that the *P. westermani* COI region was clustered into East, Southeast, and South Asia groups (Fig. 2A). Like the COI region, *P. westermani* ITS2 sequences reported from East Asia (Korea, Japan, and China; Group I), Southeast Asia (Thailand; Group II), and South Asia (India; Group III) were also separately clustered in the phylogenetic analysis (Fig. 2B). The *P. westermani* ITS2 sequences from Cheongdo and YG2-4 Korean mum-
Fig. 2: maximum likelihood tree of *Paragonimus* (A) COI and (B) ITS2 DNA region. The percentage of replicate trees in which the associated taxa clustered in the bootstrap test are marked next to the branches. The ancient *P. westermani* sequences revealed in this study are represented by red dots.

Our findings in our study belonged to the East Asia group as well. Taken together, our phylogenetic tree of the ITS2 region was generally similar to those reported in previous studies conducted by Iwagami et al.\(^{(13,17)}\) and Devi et al.\(^{(8,16)}\) However, different from the COI region, there was also a unique finding of our examination of the ITS2 region of *P. westermani*. In brief, some *P. westermani* ITS2 sequences from South Asia (India: AB938198.1, JN656208.1, JN656205.1) and Southeast Asia (Vietnam: LC144902.1, FJ434982.1) were clustered with those of the East Asia group in this study (Fig. 2B). Similar findings have rarely been reported except for the research conducted by Doanh et al.\(^{(12)}\)

According to the paleoparasitological estimation, the prevalence of *P. westermani* infection in the Joseon society might reach as high as 33.3%\(^{(18,19)}\). Since Joseon people commonly enjoyed raw crayfish or raw crab dishes,\(^{(7,29,30)}\) it is understandable that they were heavily
infected by *P. westermani*. In this study, using coprolite samples obtained from Joseon dynasty mummies, we successfully analysed the COI and ITS2 regions of aDNA from *P. westermani*. As very few studies have examined ancient *P. westermani* DNA so far, the current study significantly expands the existing gene pool of *P. westermani* paleogenetics. Nevertheless, we admit that aDNA reports of much wider geo-historical scope are still required to improve our knowledge about the exact evolutionary history of *P. westermani*.

AUTHORS’ CONTRIBUTION

DHS and MS designed this study as leading authors; MS performed microscopic examination; JHH and OCS performed the molecular experiments; JHH, MS, and DHS analysed the data and wrote the manuscript.

REFERENCES

1. Blair D, Nawa Y, Mitreva M, Doanh PN. Genetic diversity and genetic variation in lung flukes (genus Paragonimus). Trans R Soc Trop Med Hyg. 2016; 110(1): 6-12.

2. Bogitsh B, Carter CE, Oeltmann TN. Human parasitology. 3rd ed. Cambridge: Academic Press; 2005. p. 218-20.

3. Blair D, Xu ZB, Agatsuma T. Paragonimiasis and the genus *Paragonimus*. Adv Parasitol. 1999; 42: 113-222.

4. Choi DW. *Paragonimus* and paragonimiasis in Korea. Korean J Parasitol. 1990; 28(Suppl.): 79-102.

5. Kim EA, Juhe SK, Kim HW, Kim GD, Lee YW, Cho HJ, et al. Imaging findings of hepatic paragonimiasis: a case report. J Korean Med Sci. 2004; 19(5): 759-62.

6. Koh EI, Kim S-K, Wang K-C, Chai J-Y, Chong S, Park S-H, et al. The return of an old worm: cerebral paragonimiasis presenting with intracerebral hemorrhage. J Korean Med Sci. 2012; 27(11): 1428-32.

7. Shin DH, Oh CS, Lee SJ, Lee E-J, Yim SG, Kim MJ, et al. Ectopic paragonimiasis from 400-year-old female mummy of Korea. J Archaeol Sci. 2012; 39(4): 1103-10.

8. Devi KR, Narain K, Mahanta J, Nirmolia T, Blair D, Saikia SP, et al. Presence of three distinct genotypes within the *Paragonimus westermani* complex in northeastern India. Parasitol. 2013; 140(1): 76-86.

9. Kerbert C. Zur Trematoden-Kenntnis. Zoolog Anz. 1878; 1: 271-3.

10. Zarrin-Khameh N, Citron DR, Stager CE, Laucirica R. Pulmonary paragonimiasis diagnosed by fine-needle aspiration biopsy. J Clin Microbiol. 2008; 46(6): 2137-40.

11. Bincahi S, Rangsiruji A, Ketudat P, Morishima Y, Sugiyama H. Molecular systematics of a new form of *Paragonimus westermani* discovered in Thailand. Southeast Asian J Trop Med Public Health. 2007; 38: 92-6.

12. Doanh PN, Shinohara A, Horii Y, Habe S, Nawa Y. Discovery of *Paragonimus westermani* in Vietnam and its molecular phylogenetic status in *P. westermani* complex. Parasitol Res. 2009; 104(5): 1149-55.

13. Iwagami M, Rajapakse RPVJ, Paranagama W, Agatsuma T. Identities of two *Paragonimus* species from Sri Lanka inferred from molecular sequences. J Helminthol. 2003; 77(3): 239-45.

14. Prasad PK, Tandon V, Biswal DK, Goswami LM, Chatterjee A. Phylogenetic reconstruction using secondary structures and sequence motifs of ITS2 rDNA of *Paragonimus westermani* (Kerbert, 1878) Braun, 1899 (Digenea: Paragonimidae) and related species. BMC Genomics. 2009; 10(suppl. 3): S25.

15. Sugiyama H, Morishima Y, Bincahi S, Rangsiruji A, Ketudat P. New form of *Paragonimus westermani* discovered in Thailand: morphological characteristics and host susceptibility. Southeast Asian J Trop Med Public Health. 2007; 38(Suppl. 1): 87-91.

16. Devi KR, Narain K, Agatsuma T, Blair D, Nagataki M, Wickramasinghe S, et al. Morphological and molecular characterization of *Paragonimus westermani* in northeastern India. Acta Trop. 2010; 116(1): 31-8.

17. Iwagami M, Rajapakse RP, Paranagama W, Okada T, Kano S, Agatsuma T. Ancient divergence of *Paragonimus westermani* in Sri Lanka. Parasitol Res. 2008; 102(5): 845-52.

18. Seo M, Oh CS, Chai JY, Jeong MS, Hong SW, Seo YM, et al. The changing pattern of parasitic infection among Korean population by paleoparasitological study of Joseon dynasty mummies. J Parasitol. 2014; 100(1): 147-50.

19. Seo M, Oh CS, Hong JH, Chai J-Y, Cha SC, Bang Y, et al. Estimation of parasite infection prevalence of Joseon people by paleoparasitological data updates from the ancient feces of pre-modern Korean mummies. Anthropol Sci. 2017; 125(1): 9-14.

20. Hofreiter M, Serre D, Poinar HN, Kuch M, Pääbo S. Ancient DNA. Nat Rev Genet. 2001; 2(5): 353-9.

21. Ho SY, Gilbert MT. Ancient mitogenomes. Mitochondrion. 2010; 10(1): 1-11.

22. Forde C. Vermilion accord on human remains (1989) (indigenous archaeology). In: Smith C, editor. Encyclopedia of global archaeology. New York: Springer; 2014.

23. Kim Y-S, Oh CS, Lee SJ, Park JB, Kim MJ, Shin DH. Sex determination of Joseon skeletons based on anatomical, cultural and molecular biological clues. Ann Anat. 2011; 193(6): 539-43.

24. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Mol Biol Evol. 2016; 33(7): 1870-4.

25. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997; 25(17): 3389-402.

26. Tamura K. Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C-content biases. Mol Biol Evol. 1992; 9(4): 678-87.

27. Kimura M. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 1980; 16(2): 111-20.

28. Hall BG. Building phylogenetic trees from molecular data with MEGA. Mol Biol Evol. 2013; 30(5): 1229-35.

29. Sohn BS, Bae YJ, Cho YS, Moon HB, Kim TB. Three cases of paragonimiasis in a family. Korean J Parasitol. 2009; 47(3): 281-5.

30. Yun DJ. Paragonimiasis in children in Korea: related to the custom of ingesting raw crayfish for the treatment of measles. J Pediatr. 1960; 56(6): 736-51.