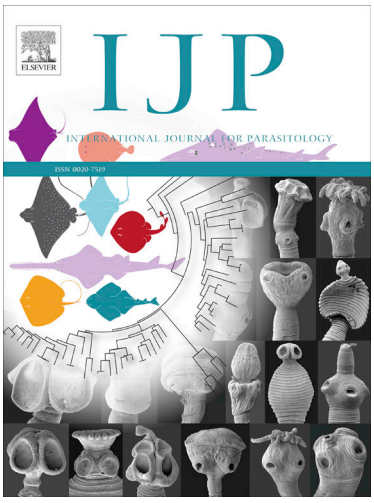
Journal Pre-proofs




# Challenging the phylogenetic relationships among *Echinococcus multilocularis* isolates from main endemic areas

Séverine Lallemand, Jorge Oyhenart, Benoit Valot, Romain Borne, Louis Bohard, Gérald Umhang, Jacek Karamon, Sergey Konyaev, Caroline Rönnberg, Bruno Gottstein, Delphine Weil-Verhoeven, Carine Richou, Solange Bresson-Hadni, Laurence Millon, Anne-Pauline Bellanger, Jenny Knapp

PII: S0020-7519(24)00110-3
DOI: https://doi.org/10.1016/j.ijpara.2024.05.004
Reference: PARA 4625

To appear in: *International Journal for Parasitology*

Received Date: 16 January 2024
Revised Date: 5 April 2024
Accepted Date: 23 May 2024

Please cite this article as: Lallemand, S., Oyhenart, J., Valot, B., Borne, R., Bohard, L., Umhang, G., Karamon, J., Konyaev, S., Rönnberg, C., Gottstein, B., Weil-Verhoeven, D., Richou, C., Bresson-Hadni, S., Millon, L., Bellanger, A-P., Knapp, J., Challenging the phylogenetic relationships among *Echinococcus multilocularis* isolates from main endemic areas, *International Journal for Parasitology* (2024), doi: https://doi.org/10.1016/j.ijpara.2024.05.004

# Challenging the phylogenetic relationships among *Echinococcus multilocularis* isolates from main endemic areas ★

Séverine Lallemand[a], Jorge Oyhenart[b], Benoit Valot[a], Romain Borne[a], Louis Bohard[c], Gérald Umhang[d], Jacek Karamon[e], Sergey Konyaev[f], Caroline Rönnberg[g], Bruno Gottstein[h], Delphine Weil-Verhoeven[i], Carine Richou[i], Solange Bresson-Hadni[a,j], Laurence Millon[a,j], Anne-Pauline Bellanger[a,j], Jenny Knapp[a,j,*]

[a] *UMR CNRS 6249 Chrono-Environnement – University of Franche-Comté – 16 Route de Gray, 25030 Besançon cedex, France*

[b] *Instituto de Ciencias de la Tierra y Ambientales de La Pampa, Consejo Nacional de Investigaciones Científicas y Técnicas – Facultad de Ciencias Naturales Universidad Nacional de La Pampa, La Pampa, Argentina*

[c] *Department of Infectiology, University Hospital of Besançon, 25030 Besançon, France*

[d] *INTERFAS Unit, National Reference Laboratory for Echinococcus spp., Rabies and Wildlife Laboratory, ANSES, Malzéville, France*

[e] *Department of Parasitology and Invasive Diseases, National Veterinary Research Institute, Puławy, Poland*

[f] *Institute of Systematics and Ecology of Animals, SB RAS, Novosibirsk, Russia*

[g] *Public Health Agency of Sweden, Department of Microbiology, Unit for Parasitology, Solna, Sweden*

[h] *Institute of Infectious Diseases, Faculty of Medicine, University of Bern, CH-3012, Bern, Switzerland*

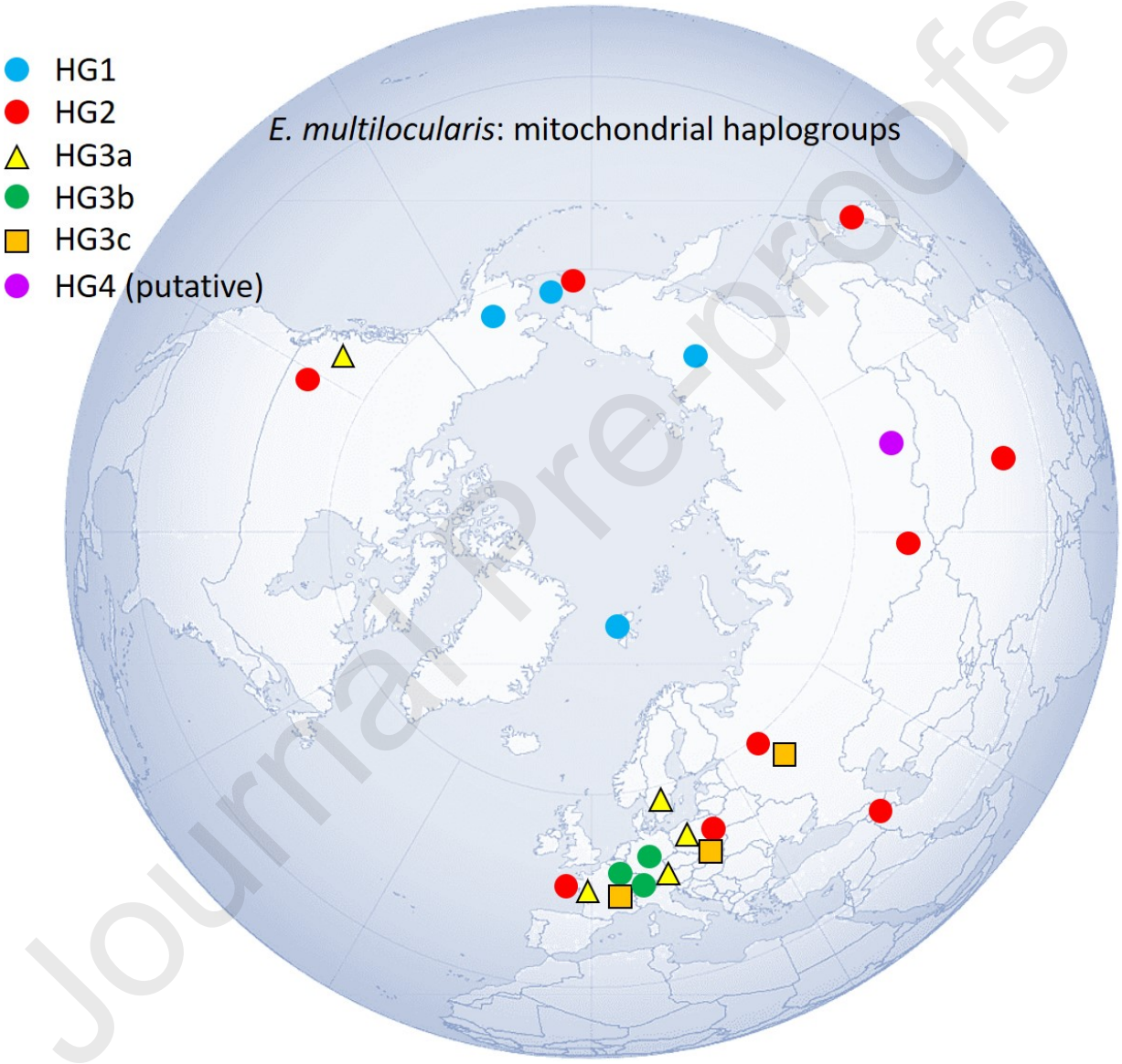
[i] *Department of Hepatology and Intensive Cares, University Hospital of Besançon, 25030 Besançon, France*

[j] *Department of Parasitology-Mycology, Reference National Centre for Echinococcoses, University Hospital of Besançon, 25030 Besançon, France*

* Corresponding author. Department of Parasitology-Mycology, University Hospital of Besançon, 3 boulevard Alexander Fleming, F-25003 Besançon, France.

*E-mail address*: jknapp@univ-fcomte.fr (J. Knapp)

★Note: Nucleotide sequence data reported in this paper are available in the GenBank™ database under the accession numbers **OR911371-OR911453**.






Highlights

- A phylogeny of *Echinococcus multilocularis* was done based on complete mitochondrial sequencing
- Samples (113) from intermediate and definitive hosts were collected worldwide
- Three main (and one putative) haplogroups were emphasized from the sample collection

- One haplogroup including French patients' isolates was divided into three micro-haplogroups
- Finer genetic diversity was described from full mitochondrial sequencing

**ABSTRACT**

Alveolar echinococcosis (AE) is a rare but severe disease that affects more than 18,000 people worldwide per year. The complete sequencing of the mitochondrial genome of *Echinococcus multilocularis* has made it possible to study the genetic diversity of the parasite and its spatial and temporal evolution. We amplified the whole mitochondrial genome by PCR, using one uniplex and two multiplex reactions to cover the 13,738 bp of the mitogenome, and then sequenced the amplicons with Illumina technology. In total, 113 samples from Europe, Asia, , the Arctic and North America were analyzed. Three major haplogroups were found: HG1, which clustered samples from Alaska (including Saint-Lawrence Island), Yakutia (Russia) and Svalbard; HG2, with samples from Asia, North America and Europe; and HG3, subdivided into three micro-haplogroups. HG3a included samples from North America and Europe, whereas HG3b and HG3c only include samples from Europe. In France, HG3a included samples from patients more recently diagnosed in a region outside the historical endemic area. A fourth putative haplogroup, HG4, was represented by only one isolate from Olkhon Island (Russia). The increased discriminatory power of the complete sequencing of the *E. multilocularis* mitogenome has made it possible to highlight four distinct geographical clusters, one being divided into three micro-haplogroups in France.

*Keywords*: *Echinococcus multilocularis*, Whole mitogenome sequencing, Haplogroup, Phylogenetic inference

## 1. Introduction

*Echinococcus multilocularis* is a flatworm of the Class Cestoda, Family Taeniidae. Its life cycle involves members of the Order Carnivora (mainly red foxes in Europe) as the definitive hosts (DH) and small rodents as the intermediate hosts (IH). Humans act as aberrant intermediate and dead-end hosts. The parasite is the causative agent of alveolar echinococcosis (AE), a rare but serious zoonosis, primarily affecting the liver, more rarely other organs and potentially secondarily metastasizing. In IH, the metacestode stage of the parasite leads to the formation of an infiltrating and destructive parasitic tumour of varying size, sometimes with central necrosis. The resulting AE is a disease that develops slowly, clinically manifesting between 5 - 15 years after infection, mainly affecting the liver (96%) but also other organs in approximately 4% of cases, e.g., the lungs, brain, or bones (Thompson and

McManus, 2002; Eckert and Deplazes, 2004; Piarroux et al., 2011). When the first clinical signs appear (pain, general alteration of the physical state, jaundice, hepatomegaly, cholangitis, pruritus), the stage of parasite development is often far advanced and thus also the metacestode mass. The main first-line diagnostic tool is ultrasound which, coupled with specific serology, confirms 95% of cases (Bresson-Hadni et al., 2021).

More than 18,000 new human cases occur annually worldwide, mostly in China (Torgerson et al., 2010). In Europe, the historical endemic zone covers the countries near the Alps (southern Germany, Switzerland, Austria, and eastern France) and was described as early as 1900 by Posselt (Eckert and Thompson, 2017). In France, the incidence of AE has risen in recent years, with an increasing number of diagnoses outside the historical endemic area (data from the National Reference Centre of Echinococcoses (NRC-E) University Hospital, Besançon, France). This area was described by Piarroux et al. in 2013 (Piarroux et al., 2013) and consists of two regions in the east and centre of France, including a total of 22 *départements* (second largest administrative areas in France, corresponding to the European NUTS3 level), known as *départements* at risk (DAR). The annual number of declared French AE cases has risen from seven to approximately 50 (data from the NRC-E), in the last 40 years. Simultaneously, the proportion of immunocompromised AE patients increased from 1% to 18% between 1982 and 2012. Chauchet and co-workers (2014) suggest a link between the increase in the number of immunocompromised AE patients due to immunosuppressive treatment and a change in therapeutic choices for cancer and chronic inflammatory diseases (Chauchet et al., 2014).

In North America, few human cases had been described outside St. Lawrence Island. However, human cases have been increasingly described in recent years. Houston and co-workers (2021) reported a cluster of 17 human cases in Alberta, Canada, between 2013 and 2020, partly associated with the emergence of a European-type strain in local wildlife. However, the strains from these patients have not been sequenced (Houston et al., 2021). The infectivity of strains according to origin has been little studied. Bartel and co-workers (1992) compared *E. multilocularis* specimens from Montana and Alaska in a rodent model. The specimen from Alaska (St. Lawrence Island) was found to be more pathogenic, with faster development, earlier formation of lung metastases, and a high number of deaths versus none for those infected with the Montana isolate (Bartel et al., 1992).

Studying the genetic diversity of *E. multilocularis* mitochondrial DNA (mtDNA) allows determination of haplotypes, as defined by unique sets of variants, followed by haplogroups, defined as collections of similar haplotypes, identified in phylogenetic trees as major branch points (Mitchell et al., 2014). Studies on genetic diversity have advanced considerably over the last 10 years. The genetic diversity of the parasite includes the study of single nucleotide polymorphisms (SNPs), insertion-deletions (indels), and heteroplasmy. Although the non-Mendelian inheritance pattern of mtDNA reduces its informative value in population genetics, it is widely used due to its haploid nature, high copy number, and faster mutation rate than nuclear markers (Ballard and Whitlock, 2004). In addition to the genes usually processed, determination of the complete mitochondrial genome provides a larger genetic diversity by including areas more susceptible to variations than genes, as the non-coding regions I and II and ribosomal DNA. (Nakao et al., 2002).

Mitochondrial sequence analyses have enabled the identification of distinct groups of *E. multilocularis* (Bowles et al., 1992; Bowles and McManus, 1993). Three mitochondrial genes have been particularly used: cytochrome b (*cob*), NADH dehydrogenase subunit 2 (*nad2*), and cytochrome c oxidase subunit 1 (*cox1*). Based on the study of these three genes, four genetically well-differentiated clades have been described at the continental level (Asian, European, North American, and Inner-Mongolian clades) (Nakao et al., 2009). Other mitochondrial genes have also been studied, such as NADH dehydrogenase subunit 1 (*nad1*) and the 12S gene (Gesy et al., 2014; Wang et al., 2018; Heidari et al., 2019). Santoro and co-authors (2024) described 43 haplotypes, including four main ones, from a collection of European *E. multilocularis* isolates, one haplotype being found in all the countries investigated but Svalbard. Additionally, two main haplotypes seemed to dominate, the first one mostly in western, central and eastern European countries, and the other in Baltic areas and north-eastern Poland. An Asian-type haplotype was also found in Latvia and north-eastern Poland (Santoro et al., 2024). In North America, the recent description of human cases (17 in the province of Alberta and one in Saskatchewan (Canada), and three in Vermont and Minnesota states, USA) (Yamasaki et al., 2008; Polish et al., 2022, 2021), has led Santa and co-authors (2023) to study the genetic diversity of *E. multilocularis* in wild coyotes and red foxes from western Canada. The genetic variants were closely related to the European clade and supported the hypothesis of a relatively recent invasion with various founder events (Santa et al., 2023).

To date, nuclear markers (Haag et al., 1997; Knapp et al., 2012) and microsatellite markers (Bretagne et al., 1996; Nakao et al., 2003; Bart et al., 2006; Knapp et al., 2020; Umhang et al., 2021) have also been used. Nuclear markers are generally highly conserved and therefore less suitable for establishing phylogeny. Microsatellite markers such as the tandem repeated multi-locus microsatellite EmsB, are highly discriminating from the local to the continental scale, but can be complex to interpret and do not permit phylogenetic considerations (Valot et al., 2015; Knapp et al., 2020).

The whole mitochondrial genome of *E. multilocularis* was sequenced in 2002. It is a closed circular molecule of 13,737 to 13,738 bp. The mitochondrial genome of *E. multilocularis* contains 12 protein-coding genes, two rRNA genes (l-rRNA, s-rRNA), and 22 tRNA genes. The ATP8 gene is missing from the mtDNA of cestodes (Nakao et al., 2002).

In a previous study (Bohard et al., 2023), a new functional protocol for the complete sequencing of the *E. multilocularis* mitochondrial genome was developed, allowing an improvement in our knowledge of genetic diversity among AE lesions from French patients.

The main objective of the present study was to refine the phylogenetic relationships, using complete mtDNA sequencing, among *E. multilocularis* isolates from different endemic areas and different hosts and, in particular, in French patients with AE. The secondary objective was to (i) investigate the existence of sub-haplogroups among those French strains and (ii) to examine whether clinical features varied according to those sub-haplogroups.

## 2. Materials and methods

*2.1. Sample collection*

In total, 113 samples of various origin were selected, of which 57 were from a collection of biological French human samples submitted to the NRC-E for molecular diagnosis and 56 selected from a worldwide collection or from previous studies (Table 1). Of the 113 samples, 74 were of human origin and 39 of animal origin (13 foxes (individual worms), 21 rodents, three monkeys, and two dogs (metacestode lesions). Thirty patients' mitogenomes had already been sequenced as part of a previous study (Bohard et al., 2023) (GenBank accession numbers **OQ599939-OQ599968**). The French patients were diagnosed with AE between 1987 and 2022. For each, retrospective information was collected, including epidemiological data (gender, age at diagnosis, residence at diagnosis, year of diagnosis, residence in a DAR, zone of residence), immunocompetence status, clinical data (organ involved, with or without extra-hepatic localization, presence of metastases, invasion of neighbouring organs). For geographical considerations, Europe has been considered in the present study as bounded to the east by the Ural Mountains and to the southeast by the Carpathian Mountains.

*2.2. PCR and sequencing*

A previously described protocol was applied to amplify the entire mitochondrial genome of *E. multilocularis* (13,738 bp) (Bohard et al., 2023). Briefly, PCR was carried out in a final volume of 20 µl containing 30 to 150 ng of purified DNA, 0.5 µM of each primer, and 0.5 U AmpliTaq Gold 360 DNA Polymerase master mix (Thermo Fisher Scientifics, Vilnius, Lithuania), and nuclease-free water. Seven pairs of previously described primers (Bohard et al., 2023), were used to amplify the whole mitochondrial genome of *E. multilocularis* (Table 2). The PCR conditions were an initial denaturation at 95°C for 5 min, followed by 40 cycles consisting of a denaturation step at 98°C for 20 s, an annealing step at 55°C for 15 s, and an elongation step at 65°C for 3 min, and a final elongation at 65°C for 5 min. All PCRs were performed with a Biometra T3 thermocycler (Whatman Biometra, Göttingen, Germany). PCR products were revealed by electrophoresis on a 1% agarose gel.

PCR products were purified separately and then mixed together in equimolar amounts after fluorometric measurement, (Qubit, Invitrogen, Thermo Fisher Scientific, Wilmington, DE, USA), to obtain a minimum concentration of mixed PCR products of 20 ng/µL. Sequencing was performed using Illumina technology (NextSeq 500 System, Illumina, San Diego, USA) after preparation of the library with a Nextera XT DNA Library Prep Kit (Illumina) by the microbiology core laboratory of the Pasteur Institute (P2M, Paris, France).

*2.3. Data analyses*

*2.3.1. Determination of the consensus sequences*

Raw paired-end reads of 150 bp were obtained by Illumina sequencing and trimmed using Sickle v1.33.2 (Joshi and Fass, 2011) at a quality threshold of 20 and a length threshold of 75 bp. Quality control of the reads was performed using FastQC v0.11.9

(https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Trimmed reads were mapped against GenBank accession number **AB018440.2** using the Burrows-Wheeler Aligner (BWA)-MEM algorithm (v0.7.17.2) (Li and Durbin, 2010). Variant calling was performed using FreeBayes (v1.3.6) (Garrison and Marth, 2012). A minimum mapping quality score of 20 and a minimal alternate fraction of observations supporting an alternate allele of 0.2 were retained. This means that mutations occurring in at least 20% of reads were analysed. This threshold allowed the examination of heteroplasmy (Pakendorf and Stoneking, 2005). Mutations occurring in 20 to 80% of reads were considered to be ambiguous sites and those occurring in more than 80% of reads were considered to be real mutations. Ambiguous sites appearing in nucleotide sequences were coded using the standard International Union of Pure and Applied Chemistry (IUPAC) for combinations of two or more bases. To obtain consensus sequences, a variant call format (VCF) filter was then applied to conserve only the mutations with a quality score Q > 500. The VCF files resulting from this mapping have been published in the European Variation Archive (EVA) (https://www.ebi.ac.uk/eva/?eva-study=PRJEB52628). Mutations related to primer sequences were found using IGV viewer (Thorvaldsdóttir et al., 2013) and were not considered for the remaining analysis. A consensus sequence in FASTA formatted files was obtained for each sample using BCFtools consensus (v1.9) (Li et al., 2009).

*2.3.2. Determination of haplotypes and population diversity*

All *E. multilocularis* mitogenome sequences and the *E. multilocularis* mitogenome reference sequence (GenBank accession number **AB018440**), were aligned using MAFFT v.7.508 (Katoh and Standley, 2013). All loci given in this article are numbered relative to this reference. Haplotypes were determined using DNA sequence polymorphism v.6.12.03 (Rozas et al., 2017), indels not being considered. Arlequin software v3.5.2.2 (Excoffier and Lischer, 2010) was used to calculate population diversity indices, such as the number of haplotypes (Hn), the haplotype diversity (Hd), corresponding to the probability that two randomly sampled genome specimens are different, nucleotide diversity (defined as the average number of nucleotide differences per site in pairwise comparisons among DNA sequences) and neutrality indices based on Tajima's D (Tajima, 1989) and Fu's fs (Fu and Li, 1993) tests to assess the random evolution of a DNA sequence.

*2.3.3. Phylogeny and determination of haplogroups*

A phylogenetic tree was constructed using all the complete mtDNA sequences to determine the genetic relationships between samples and compare them with data available in the literature. The tree was generated with MrBayes (Ronquist et al., 2012) using the previously complete sequence alignment. A generalized time-reversible (GTR) model with variable sites and a gamma distribution of evolution rates were used. The *Echinococcus shiquicus* sequence (**NC_009460.1**) (Nakao et al., 2007) was used as an outgroup. One million iterations were computed with a sampling every 500, on two independent Markov chains. Convergence was verified using an effective sample size (ESS) above 100 and a potential scale reduction factor (PSRF) close to 1 (Gelman and Rubin, 1992). A consensus tree was determined after removing the first 25% of samples as burn-in. Visualization and annotation of the tree

were performed using Itol v.6 (Letunic and Bork, 2021). A median-joining network was then created to visualize the relationship between the complete mtDNA sequences of samples from the French AE patients in greater detail. Previously defined French haplotypes were used to construct a haplotype network with PopART software (http://popart.otago.ac.nz) using the median-Joining algorithm as the inference method (Bandelt et al., 1999).

### *2.3.4. Statistical analyses*

Statistical analyses were performed using Jamovi (v.2.2.3) (https://www.jamovi.org). A potential association between the clinical forms of AE and the haplogroup of the *E. multilocularis* strains in French AE patients was investigated. The variables analysed were for (i) a quantitative variable:age at diagnosis, for (ii) qualitative variables:diagnosis after 2013, gender, immunosuppression status prior to diagnosis, and for (iii) clinical forms: hepatic forms alone, extra-hepatic forms alone, intra and extra-hepatic forms, invasion of neighbouring tissues and/or organs and presence of metastases. For quantitative variables, normality (Shapiro-Wilk) and homogeneity of variance (Levene) tests were assessed. If normality and/or homogeneity were not verified, a non-parametric test was applied (Kruskall-Wallis). Otherwise, Fisher's test was used. For qualitative variables, verification of the expected number of observations led to the use of the chi-square test or Fisher's exact test when the number of observations was less than five. Values of $P < 0.05$ were considered statistically significant.

## 3. Results

### *3.1. Description of polymorphism*

In total, 113 DNA extracts were sequenced and analysed by alignment and variant calling. The sizes of the mitogenomes ranged from 13,735 to 13,740 bp. Among the 13,747 sites (corresponding to the size of the multiple alignment including indels), 13,311 were invariable (sequence conservation C = 0.969). Of the 436 polymorphic sites, 17 were indels and 419 were substitutions. Among the substitutions, 207 were parsimony informative sites (205 with two variants and two with three variants at loci 3798 and 6811 according to the reference sequence **AB018440**) and 212 were singleton variable sites (210 with two variants and two with three variants at loci 11,948 and 13,526). The nucleotide content was: T (48.42%), A (20.59%), G (23.36%) C (7.63%). The molecular diversity index values are summarized in Table 3. Selective neutrality tests were not significant (Tajima's D = -1.27271, $P = 0.065$; Fu's Fs = 0.71923, $P = 0.664$).

SNPs were evenly distributed throughout the mitogenome. Parsimony informative sites concerned 0 (*nd*4L) to 2.9% (*cox*3) of the nucleotide sites in the gene sequences and up to 3.8% in intergenic region 1 (Fig. 1). The nucleotides at positions 112 and 3490 differed from those in the reference sequence, which suggested that it was rather the reference sequence that contained the SNPs. One major insertion of four bases was detected at locus 8376, at the end of the *nad*1 gene (position 886 in the gene), for all Yakutia (Russia) and Svalbard samples and for 6/7 samples from Alaska.

*3.2. Determination of haplotypes and phylogenetic inference*

From the 113 aligned mitochondrial sequences with the mitogenome reference sequence **AB018440**, 58 haplotypes were determined. Forty-three were represented by a single sample, whereas 15 haplotypes clustered 2 - 11 samples. Haplotype diversity was 0.971 ± 0.006 (Table 3).

Using the 113 *E. multilocularis* mitogenomes described in the present study, the reference sequence **AB018440**, and the reference sequence of *E. shiquicus* **NC_009460.1** as an outgroup, three haplogroups, HG1 to HG3, were clearly distinguishable from the phylogenetic analysis. One sample (S76) with a high number of mutations and undetermined nucleotides from position 10,843 to 12,850, due to the lack of PCR amplification for one primer pair (coded as N according to the IUPAC code) was the only representative of a putative fouth group, distant from the others.

There was a strong posterior probability of 100% for the branch separating HG1, HG2, and HG3 (Fig. 2A). In haplogroup HG3, three micro-haplogroups HG3a, HG3b, and HG3c were highighted and HG3a was clearly separated from the other two other with a posterior probability of 100%. HG3b and HG3c were less clearly separated in the tree, with posterior probabilities of 87% to 100% (Fig. 2B and 2C).

Among the 113 sequenced samples, 12 (11%) were clustered in haplogroup HG1, 25 (22%) in HG2, 75 (66%) in HG3 and one (1%) in HG4. In haplogroup HG3, three micro-haplogroups were determined, HG3a with 20 samples (27%), HG3B with 26 samples (35%) and HG3c with 29 samples (38%). In terms of geographical distribution, 82 (73%) showed a European origin, 74 (90%) belonging to haplogroup HG3. Five samples belonged to haplogroup HG2 and originated from the Ryazan Oblast on the Russian-European plain (*n*=2), Poland (*n*=2), or France (*n*=1). The French sample came from a patient who was hypothetically contaminated in Tibet (China). The last three samples belonged to haplogroup HG1 and originated from Svalbard Island (Norway).

Twenty samples were of Asian origin (18% of the sample collection) and 17 (85%) belonged to haplogroup HG2; two other samples belonged to haplogroup HG1 and originated from Yakutia and one sample belonged to a new putative haplogroup HG4, originating from Olkhon Island on Lake Baikal. The *cox1* sequence from the Lake Baikal sample (**OR911453**) was aligned with eight other *cox1* sequences from the Mongolian clade, available (**AB510023**, **AB510025**, **AB777920-21**, **AB461420**, **AB813186-188**), the reference sequences **AB018440** and all the present samples. Table 4 shows the SNP loci described in these Mongolian *cox1* sequences. Among the 28 SNPs, 17 (60%) seemed specific to HG4 and were never found in another samples.

Twelve samples (11% of the sample collection) were of North American origin and seven (58%) belonged to haplogroup HG1 and originated from St. Lawrence Island (Alaska). Four belonged to haplogroup HG2 and originated from Canada and St. Lawrence Island. Another sample belonged to haplogroup HG3 and originated from Canada. This last sample originated from a dog with hepatic AE, native to the British Columbia province (Quesnel area), Canada, with no trip abroad declared (Peregrine et al., 2012) (Fig. 3).

In France, among the 56 French patient samples belonging to haplogroup HG3, 10 clustered in micro-haplogroup HG3a (18%), 21 in HG3b (37%) and 25 in HG3c (45%). All samples from patients residing in north-western France shared the same haplotype HG3a (Fig. 4). Interestingly, only micro-haplogroup HG3a was also found outside Europe (in Alaska and Canada), instead of HG3b and HG3c (Fig. 3).

### *3.3. Comparison of epidemiological and clinical data according to haplogroups among French AE patients*

Among the 57 French AE patients, 56 (98%) presented an AE lesion belonging to the HG3 haplogroup, 10 (18%) in the micro-haplogroup HG3a, 21 (37%) in HG3b, and 25 (45%) in HG3c. One patient presented an AE lesion clustered in the HG2 haplogroup after reporting several visits to China years before. Comparisons of the results of epidemiological and clinical data for French patients for the HG3 micro-haplogroups are summarized in Table 5. Briefly, the proportion of diagnoses after 2013 differed between the HG3a and HG3c micro-haplogroups (from 2014 to 2022 for HG3a versus 1997 to 2022 for HG3c, Fisher's exact test, $P = 0.007$).

## 4. Discussion

Sequencing of the complete *E. multilocularis* mtDNA based on a worldwide collection of AE lesions and *E. multilocularis* adult worms has enabled us to establish a phylogeny of *E. multilocularis*. Until now, the common use of concatenated genes has not allowed a distinction between haplogroups in France. Bohard et al. (2023) showed that concatenation of the *cob*, *cox1*, and *nad2* genes made it possible to distinguish just six haplotypes, versus 13 based on the complete mitochondrial genome obtained from samples collected from 30 French AE patients. Here, the study of the complete mitochondrial genome of a panel including 113 samples revealed 419 mutations and enabled us to differentiate 58 haplotypes and three major haplogroups. The classic use of concatenated genes made it possible to solely determine continental clusters as European, Asian, central or Arctic North American, and Alaskan clades (Nakao et al., 2009; Wu et al., 2017; Santoro et al., 2024). Because the complete mitochondrial DNA of *E. multilocularis* was sequenced in the present study, the term "haplogroup" instead of the "clade" was preferred, which is generally applied in the literature.

Haplogroup HG1 included haplotypes from locations close to the Arctic Circle: Svalbard Island (Norway), St. Lawrence Island (Alaska), and Yakutia (Russia). Given the geographical origin of these samples and their genetic distance one to another, we hypothesized that the subspecies described by Rausch and Schiller could belong to haplogroup 1 (Rausch and Schiller, 1954). It was named *Echinococcus sibiricensis* and characterized by shorter rostellar hooks in arctic foxes, *Vulpes lagopus*, and sled dogs on St. Lawrence Island in the Bering Sea. Moreover, in 1957, Vogel (Vogel, 1957) considered *E. sibiricensis* to be a geographic subspecies of *E. multilocularis*. In 1985 and 1998-1999, parasites from the Hulunbeier Pasture in Inner Mongolia, northeastern China, were investigated and differences in the evolution between *E. multilocularis* and *E. sibiricensis* were reported (Tang et al., 2001). However, molecular studies confirmed that these two parasites belong to the same species (Nakao et

al., 2009). Moreover, the place of the recently described species *Echinococcus russicensis* in phylogeny also has to be studied with regard to the present findings (Romig et al., 2015).

Haplogroup HG2 included samples from Asia (Russia, Armenia, Japan, and China), as well as North America (Canada and Alaska) and Europe (Poland, France). From the present findings, it appears to be the most frequently distributed haplogroup worldwide.

Haplogroup HG3 combined mainly European samples (France, Switzerland, Germany, Luxembourg, Belgium, Sweden, and Poland), as well as Russian (Ryazan oblast) and Canadian samples. This haplogroup accounted for the majority of the present studied samples (65%) and analysis of the phylogenetic tree enabled us to determine the three micro-haplogroups HG3a, HG3b, and HG3c. In France, HG3b and HG3c constituted the two major micro-haplogroups found in the historical endemic zone of AE (Piarroux et al., 2013). Of note, mitochondrial intergenic zone 1 (positions 67 to 249 according to AB018440) features a variable zone specific to European micro-haplogroups HG3a and HG3c. Santoro and co-workers described a predominant haplotype in French red fox samples, which is characterized by the presence of an 'A' in position 222 in the *nad2* gene (position 6610 according to AB018440) (Santoro et al., 2024). This SNP was not found in any of the present French samples considered, although heteroplasmy occurrence was studied. This difference could be due to the fox origin in the former paper and the human origin of the present collection of samples. Haplogroup HG3a included all samples from more recently diagnosed patients included in the present study and all were from north-western France, an area with a recent emergence of AE cases. This region of France is not part of the historical endemic zone. It is, thus, difficult to know whether these cases i) are only better recorded today with the active registration of AE cases, as this is the case e.g. in France with the NRC-E, mandated by government authorities, ii) were detected due to improved diagnostic techniques for the parasite in humans and animals, or iii) indicate the presence of an emerging area contaminated by *E. multilocularis* with the spreading of foxes, as previously observed (Combes et al., 2012; Umhang et al., 2014). Of note, all patients from this western area presented specimens genetically clustering in the same haplotype, resulting in low genetic diversity, which could occur, for example, after a bottleneck event. Haplogroup HG3a was also described in two monkeys from a zoo in Zürich (Switzerland), in a Swedish patient, and in two Polish foxes from the region of Śląskie and Podlaskie. For the isolates of the six patients with HG3a for whom the information was available, all had an identical microsatellite EmsB profile (P04 in Knapp and co-workers (2020)). This EmsB profile was also found to be predominant in isolates from red foxes in France in a previous study (corresponds to P17 in Umhang (2014)).

A fourth putative haplogroup was described in a single Asian sample obtained from a rodent native of the Olkhon Island in Russia. *Alticola olchonensis* is an endemic species of the Lake Baikal area with an extremely restricted geographical distribution. The red fox is known as the DH of *E. multilocularis* in the area and the winter ice bridge allows free island-continent migration (Konyaev et al., 2013a). However, no other representatives of this haplogroup have been found in any of the seven Altai patients in the present collection of samples. In previous studies, this specimen was identified as part of the Mongolian clade, based on partial mitochondrial and microsatellite EmsB studies (Konyaev et al., 2013; Umhang et al., 2021). Moreover, genetic relationships with Mongolian samples previously isolated from humans has to be investigated (Ito et al., 2010).

Four distinct haplogroups have been identified, each with an apparently distinct geographic origin, Arctic, Asian, European and "Mongolian". However, their distribution in the endemic area of the Northern hemisphere shows a more or less marked expansion of natural or anthropic origin. The natural migration of red foxes over long periods would explain the east-west gradient of strains of Asian origin, as far as the Ural Mountains (Ryazan oblast near Moscow), Poland, Latvia (Karamon et al., 2017; Santoro et al., 2024). In Turkey, strains of Asian origins were also described (Avcioglu et al., 2021; Santoro et al., 2024). In the opposite west-east direction, the expansion of Asian strains could be partly explained by the fact that Arvicoline rodents most likely crossed the Beringia land bridge during the Pleistocene, carrying with them Asian strains of *E. multilocularis* to North America (Rausch, 1994). Hayashi and co-workers (2023) described transfers of DH and DI between St. Lawrence, Bering, Kurile, Rebun and Hokkaido Islands, leading to the colonisation by arctic-type strains (Hayashi et al., 2023). In Alaska, in addition to Arctic and Asian strains, two European-type strains were present in this study. In recent decades, several authors have demonstrated the presence of European-type strains in Canada (British Columbia (Jenkins et al., 2012), Alberta and Saskatchewan (Gesy and Jenkins, 2015) and in the United States, as in New York State (Conlon et al., 2023), Vermont (Polish et al., 2021), Virginia (Zajac et al., 2020), Missouri, (Kuroki et al., 2022)). Various anthropogenic factors may explain this colonization. The importation of red foxes of European origin for sport hunting (in the mid-1700s in the eastern USA and over the following 150 years), followed by other intentional imports over the last century (fur trade, for example) importing dogs, may have led to the establishment of these strains (Kamler and Ballard, 2002; Davidson et al., 2012).

The Arctic fox harboured a proper strain and the Svalbard example is an obvious simple system with one DH (*Vulpes lagopus*), one IH (*Microtus levis*), and a well-defined and restricted area. Henttonen et al. (2001) concluded that the migrating arctic fox was the source of spread of *E. multilocularis* to Svalbard and that the introduction of an IH through human activities was the reason for the establishment of the parasite. A microsatellite EmsB study found a profile similar to some described in St. Lawrence Island isolates. A circumpolar origin of the parasite, harboured by arctic foxes and transmitted to imported rodents, was hypothesized (Knapp et al., 2012). Our study confirms this result with the close relative relationship between samples from Svalbard and those from St. Lawrence Island clustering in HG1.

Distinguishing haplogroups could help in the investigation of genetic diversity related to infectivity and nuclear gene expression. Among the 56 French patients belonging to the HG3 haplogroup, there were no differences in the type of organ affected (liver versus other organs) according to the HG3a, 3b, or 3c micro-haplogroup ($P$ = 0.459). In addition, the immunocompetent status of the patients did not differ between micro-haplogroups (all with approximately 40% immunosuppressed patients). These micro-haplogroups may be too close genetically to show a true difference in infectivity. It could be informative to compare the infectivity of *E. multilocularis* between the main haplogroups (HG1, HG2, and HG3). To date and to our knowledges, no study has demonstrated a correlation between *E. multilocularis* genetic characteristics and clinical data in patients with alveolar echinococcosis. It could be informative to compare the pathogenicity of the haplogroups determined by mtDNA studies through the analysis of functional markers of parasite infectivity, as expression, or regulation studies of nuclear genes involved in the variation of the pathogenicity of the parasite.

In conclusion, from a worldwide collection of *E. multilocularis* samples, we describe three haplogroups that are distributed differently around the world and a putative fourth haplogroup represented by only one sample originating from the Inner Mongolian region. Haplogroup HG1 appears to be restricted to the polar region, whereas haplogroup HG2 is distributed worldwide. The third haplogroup, HG3, is composed of three micro-haplogroups. The micro -haplogroups HG3b and HG3c were only found in Europe in the present collection of samples, whereas micro-haplogroup HG3a was also present in North America. HG3a also appears to predominate in north-western France, outside the historical endemic area.

## Acknowledgements

This work was funded by the National Reference Centre for Echinococcoses, University Hospital, Besançon, France.



## References

Avcioglu, H., Guven, E., Balkaya, I., Kirman, R., Akyuz, M., Mebarek Bia, M., Gulbeyen, H., Yaya, S., 2021. The situation of echinococcosis in stray dogs in Turkey: the first finding of *Echinococcus multilocularis* and *Echinococcus ortleppi*. Parasitology 148, 1092–1098.

Ballard, J.W.O., Whitlock, M.C., 2004. The incomplete natural history of mitochondria. Mol. Ecol. 13, 729–744.

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Bart, J.M., Knapp, J., Gottstein, B., El-Garch, F., Giraudoux, P., Glowatzki, M.L., Berthoud, H., Maillard, S., Piarroux, R., 2006. EmsB, a tandem repeated multi-loci microsatellite, new tool to investigate the genetic diversity of *Echinococcus multilocularis*. Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis. 6, 390–400.

Bartel, M.H., Seesee, F.M., Worley, D.E., 1992. Comparison of Montana and Alaska isolates of *Echinococcus multilocularis* in gerbils with observations on the cyst growth, hook characteristics, and host response. J. Parasitol. 78, 529–532.

Bohard, L., Lallemand, S., Borne, R., Courquet, S., Bresson-Hadni, S., Richou, C., Millon, L., Bellanger, A.-P., Knapp, J., 2023. Complete mitochondrial exploration of *Echinococcus multilocularis* from French alveolar echinococcosis patients. Int. J. Parasitol. S0020-7519(23)00076–0.

Bowles, J., Blair, D., McManus, D.P., 1992. Genetic variants within the genus *Echinococcus* identified by mitochondrial DNA sequencing. Mol. Biochem. Parasitol. 54, 165–173.

Bowles, J., McManus, D.P., 1993. NADH dehydrogenase 1 gene sequences compared for species and strains of the genus *Echinococcus*. Int. J. Parasitol. 23, 969–972.

Bresson-Hadni, S., Spahr, L., Chappuis, F., 2021. Hepatic Alveolar Echinococcosis. Semin. Liver Dis. 41, 393–408.

Bretagne, S., Assouline, B., Vidaud, D., Houin, R., Vidaud, M., 1996. *Echinococcus multilocularis*: microsatellite polymorphism in U1 snRNA genes. Exp. Parasitol. 82, 324–328.

Chauchet, A., Grenouillet, F., Knapp, J., Richou, C., Delabrousse, E., Dentan, C., Millon, L., Di Martino, V., Contreras, R., Deconinck, E., Blagosklonov, O., Vuitton, D.A., Bresson-Hadni, S., Virginie, V., Karine, B., Brigitte, B., Isabelle, B.-C., Oleg, B., Solange, B.-H., Pascale, B.M., Sylvie, C., Remy, C., Eric, D., Vincent, D.M., Philippe, E., Sophie, F., Patrick, G., Frédéric, G., Bruno, H., Séverine, V.-D., Jenny, K., Stéphane, K., Georges, M., Laurence, M., Francis, R., Carine, R., Claire, V., Angèle, V.D., Lucine, V., Patricia, P., Claire, G., Armand, A., Jean, B., Monique, C., François, B.J., Bernadette, C., Patrick, H., Anne, M., Odile, F., Christian, L., Jérôme, D., Olivier, G., Christian, P., Meja, R., Martine, W., Martine, P., Eric, C., Marc, J., Jean-Jacques, R., Daniel, S., Alain, G., Lorraine, L., Marie, M., Jérôme, W., Olivier, F., Didier, S., Cathy, C., Francois, D.J., Bruno, G., Monique, D., Ahmed, A.-B., Maxime, A., Yves, H., Nicolas, L., Véronique, L., Blandine, A.-L., for the FrancEchino Network, 2014. Increased Incidence and Characteristics of Alveolar Echinococcosis in Patients With Immunosuppression-Associated Conditions. Clin. Infect. Dis. 59,

Combes, B., Comte, S., Raton, V., Raoul, F., Boué, F., Umhang, G., Favier, S., Dunoyer, C., Woronoff, N., Giraudoux, P., 2012. Westward spread of *Echinococcus multilocularis* in foxes, France, 2005-2010. Emerg. Infect. Dis. 18, 2059–2062.

Conlon, C.L., Schuler, K.L., Lejeune, M., Whipps, C.M., 2023. Novel report of the European variant of *Echinococcus multiloculairs* coyotes (*Canis latrans*) in the New York State. J. Parasitol. 109, 357–361.

Davidson, R.K., Romig, T., Jenkins, E., Tryland, M., Robertson, L.J., 2012. The impact of globalisation on the distribution of *Echinococcus multilocularis*. Trends Parasitol. 28, 239–247.

Eckert, J., Deplazes, P., 2004. Biological, epidemiological, and clinical aspects of echinococcosis, a zoonosis of increasing concern. Clin. Microbiol. Rev. 17, 107–135.

Eckert, J., Thompson, R.C.A., 2017. Historical Aspects of Echinococcosis. Adv. Parasitol. 95, 1–64.

Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10, 564–567.

Fu, Y.X., Li, W.H., 1993. Statistical Tests of Neutrality of Mutations. Genetics 133, 693–709.

Gelman, A., Rubin, D.B., 1992. Inference from Iterative Simulation Using Multiple Sequences. Stat. Sci. 7, 457–472.

Gesy, K.M., Jenkins, E.J., 2015. Introduced and Native Haplotypes of *Echinococcus multilocularis* in Wildlife in Saskatchewan, Canada. J. Wildl. Dis. 51, 743–748.

Gesy, K.M., Schurer, J.M., Massolo, A., Liccioli, S., Elkin, B.T., Alisauskas, R., Jenkins, E.J., 2014. Unexpected diversity of the cestode *Echinococcus multilocularis* in wildlife in Canada. Int. J. Parasitol. Parasites Wildl. 3, 81–87.

Haag, K., Zaha, A., Araújo, A., Gottstein, B., 1997. Reduced genetic variability within coding and non-coding regions of the *Echinococcus multilocularis* genome. Parasitology 115 ( Pt 5), 521–9.

Hayashi, N., Nakao, R., Ohari, Y., Irie, T., Kouguchi, H., Chatanga, E., Mohamed, W.M.A., Moustafa, M.A.M., Kinoshita, G., Okamoto, M., Yagi, K., Nonaka, N., 2023. Mitogenomic exploration supports the historical hypothesis of anthropogenic diffusion of a zoonotic parasite *Echinococcus multilocularis*. iScience 26, 107741.

Heidari, Z., Sharbatkhori, M., Mobedi, I., Mirhendi, S.H., Nikmanesh, B., Sharifdini, M., Mohebali, M., Zarei, Z., Arzamani, K., Kia, E.B., 2019. *Echinococcus multilocularis* and *Echinococcus granulosus* in canines in North-Khorasan Province, northeastern Iran, identified using morphology and genetic characterization of mitochondrial DNA. Parasit. Vectors 12, 606.

Henttonen, H., Fuglei, E., Gower, C.N., Haukisalmi, V., Ims, R.A., Niemimaa, J., Yoccoz, N.G., 2001. *Echinococcus multilocularis* on Svalbard: introduction of an intermediate host has enabled the local life-cycle. Parasitology 123, 547–552.

Houston, S., Belga, S., Buttenschoen, K., Cooper, R., Girgis, S., Gottstein, B., Low, G., Massolo, A., MacDonald, C., Müller, N., Preiksaitis, J., Sarlieve, P., Vaughan, S., Kowalewska-Grochowska, K., 2021. Epidemiological and Clinical Characteristics of Alveolar Echinococcosis: An Emerging Infectious Disease in Alberta, Canada. Am. J. Trop. Med. Hyg.

Ito, A., Agvaandaram, G., Bat-Ochir, O.-E., Chuluunbaatar, B., Gonchigsenghe, N., Yanagida, T., Sako, Y., Myadagsuren, N., Dorjsuren, T., Nakaya, K., Nakao, M., Ishikawa, Y., Davaajav, A., Dulmaa, N., 2010. Histopathological, serological, and molecular confirmation of indigenous alveolar echinococcosis cases in Mongolia. Am. J. Trop. Med. Hyg. 82, 266–269.

Jenkins, E.J., Peregrine, A.S., Hill, J.E., Somers, C., Gesy, K., Barnes, B., Gottstein, B., Polley, L., 2012. Detection of European Strain of *Echinococcus multilocularis* in North America. Emerg. Infect. Dis. 18, 1010–1012.

Kamler, J.F., Ballard, W.B., 2002. A Review of Native and Nonnative Red Foxes in North America. Wildl. Soc. Bull. 1973-2006 30, 370–379.

Karamon, J., Stojecki, K., Samorek-Pieróg, M., Bilska-Zajac, E., Rózycki, M., Chmurzyńska, E., Sroka, J., Zdybel, J., Cencek, T., 2017. Genetic diversity of *Echinococcus multilocularis* in red foxes in Poland: the first report of a haplotype of probable Asian origin. Folia Parasitol. 64, 2017.007.

Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. 30, 772–780.

Knapp, J., Gottstein, B., Bretagne, S., Bart, J.-M., Umhang, G., Richou, C., Bresson-Hadni, S., Millon, L., 2020. Genotyping *Echinococcus multilocularis* in Human Alveolar Echinococcosis Patients: An EmsB Microsatellite Analysis. Pathog. Basel Switz. 9, 282.

Knapp, J., Staebler, S., Bart, J.M., Stien, A., Yoccoz, N.G., Drögemüller, C., Gottstein, B., Deplazes, P., 2012. *Echinococcus multilocularis* in Svalbard, Norway: microsatellite genotyping to investigate the origin of a highly focal contamination. Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis. 12, 1270–1274.

Konyaev, S.V., Yanagida, T., Nakao, M., Ingovatova, G.M., Shoykhet, Y.N., Bondarev, A.Y., Odnokurtsev, V.A., Loskutova, K.S., Lukmanova, G.I., Dokuchaev, N.E., Spiridonov, S., Alshinecky, M.V., Sivkova, T.N., Andreyanov, O.N., Abramov, S.A., Krivopalov, A.V.,

Karpenko, S.V., Lopatina, N.V., Dupal, T.A., Sako, Y., Ito, A., 2013. Genetic diversity of *Echinococcus* spp. in Russia. Parasitology 140, 1637–1647.

Kuroki, K., Morishima, Y., Dorr, L., Cook, C.R., 2022. Alveolar echinococcosis in a dog in Missouri, USA. J Vet Diagn Invest 34, 746–751.

Letunic, I., Bork, P., 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49, W293–W296.

Li, H., Durbin, R., 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinforma. Oxf. Engl. 26, 589–595.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Mitchell, S.L., Goodloe, R., Brown-Gentry, K., Pendergrass, S.A., Murdock, D.G., Crawford, D.C., 2014. Characterization of mitochondrial haplogroups in a large population-based sample from the United States. Hum. Genet. 133, 861–868.

Nakao, M., McManus, D.P., Schantz, P.M., Craig, P.S., Ito, A., 2007. A molecular phylogeny of the genus *Echinococcus* inferred from complete mitochondrial genomes. Parasitology 134, 713–722.

Nakao, M., Sako, Y., Ito, A., 2003. Isolation of polymorphic microsatellite loci from the tapeworm *Echinococcus multilocularis*. Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis. 3, 159–163.

Nakao, M., Xiao, N., Okamoto, M., Yanagida, T., Sako, Y., Ito, A., 2009. Geographic pattern of genetic variation in the fox tapeworm *Echinococcus multilocularis*. Parasitol. Int. 58, 384–389.

Nakao, M., Yokoyama, N., Sako, Y., Fukunaga, M., Ito, A., 2002. The complete mitochondrial DNA sequence of the cestode *Echinococcus multilocularis* (Cyclophyllidea: Taeniidae). Mitochondrion 1, 497–509.

Pakendorf, B., Stoneking, M., 2005. Mitochondrial DNA and human evolution. Annu. Rev. Genomics Hum. Genet. 6, 165–183.

Peregrine, A.S., Jenkins, E.J., Barnes, B., Johnson, S., Polley, L., Barker, I.K., De Wolf, B., Gottstein, B., 2012. Alveolar hydatid disease (*Echinococcus multilocularis*) in the liver of a Canadian dog in British Columbia, a newly endemic region. Can. Vet. J. 53, 870–874.

Piarroux, M., Piarroux, R., Giorgi, R., Knapp, J., Bardonnet, K., Sudre, B., Watelet, J., Dumortier, J., Gérard, A., Beytout, J., Abergel, A., Mantion, G., Vuitton, D.A., Bresson-Hadni, S., 2011. Clinical features and evolution of alveolar echinococcosis in France from 1982 to 2007: results of a survey in 387 patients. J. Hepatol. 55, 1025–1033.

Piarroux, M., Piarroux, R., Knapp, J., Bardonnet, K., Dumortier, J., Watelet, J., Gerard, A., Beytout, J., Abergel, A., Bresson-Hadni, S., Gaudart, J., FrancEchino Surveillance Network, 2013. Populations at risk for alveolar echinococcosis, France. Emerg. Infect. Dis. 19, 721–728.

Polish, L.B., O'Connell, E.M., Barth, T.F.E., Gottstein, B., Zajac, A., Gibson, P.C., Bah, A., Kirchgessner, M., Estrada, M., Seguin, M.A., Ramirez-Barrios, R., 2022. European Haplotype of *Echinococcus multilocularis* in the United States. N. Engl. J. Med. 387, 1902–1904.

Polish, L.B., Pritt, B., Barth, T.F.E., Gottstein, B., O'Connell, E.M., Gibson, P.C., 2021. First European Haplotype of *Echinococcus multilocularis* Identified in the United States: An Emerging Disease? Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am. 72, 1117–1123.

Rausch, R., Schiller, E.L., 1954. Studies on the helminth fauna of Alaska. XXIV. *Echinococcus sibiricensis* n. sp., from St. Lawrence Island. J. Parasitol. 40, 659–662.

Rausch, R.L., 1994. Transberingian dispersal of cestodes in mammals. Int. J. Parasitol. 24, 1203–1212.

Romig, T., Ebi, D., Wassermann, M., 2015. Taxonomy and molecular epidemiology of *Echinococcus granulosus* sensu lato. Vet. Parasitol. 213, 76–84.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Syst. Biol. 61, 539–542.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sánchez-Gracia, A., 2017. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Mol. Biol. Evol. 34, 3299–3302.

Santa, M.A., Umhang, G., Klein, C., Grant, D.M., Ruckstuhl, K.E., Musiani, M., Gilleard, J.S., Massolo, A., 2023. It's a small world for parasites: evidence supporting the North American invasion of European *Echinococcus multilocularis*. Proc. Biol. Sci. 290, 20230128.

Santoro, A., Santolamazza, F., Cacciò, S.M., La Rosa, G., Antolová, D., Auer, H., Bagrade, G., Bandelj, P., Basso, W., Beck, R., Citterio, C.V., Davidson, R.K., Deksne, G., Frey, C.F., Fuglei, E., Glawischnig, W., Gottstein, B., Harna, J., Huus Petersen, H., Karamon, J., Jansen, F., Jarošová, J., Jokelainen, P., Lundström-Stadelmann, B., Maksimov, P., Miljević, M., Miterpáková, M., Moks, E., Origgi, F., Ozolina, Z., Ryser, M.-P., Romig, T., Šarkūnas, M., Scorrano, N., Saarma, U., Šnábel, V., Sréter, T., Umhang, G., Vengušt, G., Žele Vengušt, D., Casulli, A., 2024. Mitochondrial genetic diversity and phylogenetic relationships of *Echinococcus multilocularis* in Europe. Int. J. Parasitol. 54, 233-245.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123, 585–595.

Tang, C.T., Chen, J.A., Tang, L., Cui, G.W., Qian, Y.C., Kang, Y.M., Lu, H.C., 2001. [Comparison observation on the mature alveolar of *Echinococcus sibiricensis* and *Echinococcus multilocularis* in the experimentally infected white mice]. Shi Yan Sheng Wu Xue Bao 34, 261–268.

Thompson, R.C.A., McManus, D.P., 2002. Towards a taxonomic revision of the genus *Echinococcus*. Trends Parasitol. 18, 452–457.

Thorvaldsdóttir, H., Robinson, J.T., Mesirov, J.P., 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. 14, 178–192.

Torgerson, P.R., Keller, K., Magnotta, M., Ragland, N., 2010. The Global Burden of Alveolar Echinococcosis. PLoS Negl. Trop. Dis. 4, e722.

Umhang, G., Bastid, V., Avcioglu, H., Bagrade, G., Bujanić, M., Bjelić Čabrilo, O., Casulli, A., Dorny, P., van der Giessen, J., Guven, E., Harna, J., Karamon, J., Kharchenko, V., Knapp, J., Kolarova, L., Konyaev, S., Laurimaa, L., Losch, S., Miljević, M., Miterpakova, M., Moks, E., Romig, T., Saarma, U., Snabel, V., Sreter, T., Valdmann, H., Boué, F., 2021. Unravelling the genetic

diversity and relatedness of *Echinococcus multilocularis* isolates in Eurasia using the EmsB microsatellite nuclear marker. Infect. Genet. Evol. 92, 104863.

Umhang, G, Knapp, J., Hormaz, V., Raoul, F., Boué, F., 2014. Using the genetics of *Echinococcus multilocularis* to trace the history of expansion from an endemic area. Infect. Genet. Evol. 22, 142–149.

Valot, B., Knapp, J., Umhang, G., Grenouillet, F., Millon, L., 2015. Genomic characterization of EmsB microsatellite loci in *Echinococcus multilocularis*. Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis. 32, 338–341.

Vogel, H., 1957. [*Echinococcus multiocularis* in South Germany. I. The tapeworm stage of strains from humans and animals]. Z. Tropenmed. Parasitol. 8, 404–454.

Wang, X., Liu, J., Zuo, Q., Mu, Z., Weng, X., Sun, X., Wang, J., Boufana, B., Craig, P.S., Giraudoux, P., Raoul, F., Wang, Z., 2018. *Echinococcus multilocularis* and *Echinococcus shiquicus* in a small mammal community on the eastern Tibetan Plateau: host species composition, molecular prevalence, and epidemiological implications. Parasit. Vectors 11, 302.

Wu, C., Zhang, Wenbao, Ran, B., Fan, H., Wang, H., Guo, B., Zhou, C., Shao, Y., Zhang, Wei, Giraudoux, P., Knapp, J., Wen, H., Kuang, L., Li, J., 2017. Genetic variation of mitochondrial genes among *Echinococcus multilocularis* isolates collected in western China. Parasit. Vectors 10, 265.

Yamasaki, H., Nakao, M., Nakaya, K., Schantz, P.M., Ito, A., 2008. Genetic analysis of *Echinococcus multilocularis* originating from a patient with alveolar echinococcosis occurring in Minnesota in 1977. Am. J. Trop. Med. Hyg. 79, 245–247.

Zajac, A., Fairman, D., McGee, E., Wells, B., Peregrine, A., Jenkins, E., LeRoith, T., St John, B., 2020. Alveolar echinococcosis in a dog in the eastern United States. J. Vet. Diagn. Investig. 32, 742–746.

**Figure legends**

Fig. 1. Distribution of single nucleotide polymorphisms (SNPs) in the *Echinococcus multilocularis* mitogenome. Only the distribution of mutations (total number of sites, number of parsimony informative sites (PIS) and percentage of PIS) for the 12 genes, rRNAs, and intergenic zone 1 appear in this graph to facilitate readability.

Fig. 2. Phylogenetic tree constructed using the Bayesian method on whole mitogenome sequences from *Echinococcus multilocularis* isolates and the reference sequences of *Echinococcus shiquicus* **NC_009460.1** and *E. multilocularis* **AB018440**. (A) The branch representing *E. shiquicus* has been deliberately shortened by a factor of 10 to facilitate readability. (B) Subtree representing the haplogroup HG3. The bold lines represent posterior probability values ≥ 90%. (C) Entire mitochondrial haplotype network of *Echinococcus multilocularis* based on the median-joining method. Dashes on the lines represent mutations. Each circle represents one of the 38 haplotypes. The sizes of the circles depend on the numbers of samples representing the haplotypes and the colours the geographical origins. BC, British Columbia.

Fig. 3. Schematic representation of *Echinococcus multilocularis* haplogroup distribution. Each dot indicates the presence of a haplogroup in a given country.

Fig. 4. French *Echinococcus multilocularis* haplogroup distribution map based on the patients' "*département*" of residence. The areas delimit the repartition zones of haplogroups in France in this study. The sizes of the circles are proportional to the numbers of patients in a haplogroup in a given French "*département*" (administrative division). The red (black) dot indicates a patient known to have probably been contaminated in the Tibetan region.

Table 1. Characteristics of *Echinococcus multilocularis* samples and results of haplotypic analysis.

| ID | Sample name | Geographical origin | Original host | DH /IH | Sampling year | Reference | Haplotype | Haplogroup | VCF Acc. No. | GenBank Acc. No. |
|---|---|---|---|---|---|---|---|---|---|---|
| FE-007 | NRC-E-007 | France | *Homo sapiens* | IH | 2000 | this study | MtG_h1 | HG3c | 9710-007 | OR911371 |
| FE-191 | NRC-E-191 | France | *Homo sapiens* | IH | 2011 | this study | MtG_h2 | HG3b | 11652-191 | OR911372 |
| FE-238 | NRC-E-238 | France | *Homo sapiens* | IH | 1999 | this study | MtG_h1 | HG3c | 9908-238 | OR911373 |
| FE-270 | NRC-E-270 | France | *Homo sapiens* | IH | 1985 | this study | MtG_h2 | HG3b | 0011-270 | OR911374 |
| FE-276 | NRC-E-276 | France | *Homo sapiens* | IH | 2001 | this study | MtG_h3 | HG3c | 0105-276 | OR911375 |
| FE-292 | NRC-E-292 | France | *Homo sapiens* | IH | 2004 | this study | MtG_h4 | HG3b | 0403-292 | OR911376 |
| FE-333 | NRC-E-333 | France | *Homo sapiens* | IH | 2005 | this study | MtG_h5 | HG3c | 0412-333 | OR911377 |
| FE-351 | NRC-E-351 | France | *Homo sapiens* | IH | 2005 | this study | MtG_h4 | HG3b | 0502-351 | OR911378 |
| FE-355 | NRC-E-355 | France | *Homo sapiens* | IH | 2005 | this study | MtG_h1 | HG3c | 0503-355 | OR911379 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| FE-363 | NRC-E-363 | France | *Homo sapiens* | IH | 2006 | this study | MtG_h3 | HG3c | 0506-363 | OR911380 |
| FE-366 | NRC-E-366 | France | *Homo sapiens* | IH | 2005 | this study | MtG_h6 | HG3c | 0512-366 | OR911381 |
| FE-371 | NRC-E-371 | France | *Homo sapiens* | IH | 2007 | this study | MtG_h7 | HG3b | 0512-371 | OR911382 |
| FE-404 | NRC-E-404 | France | *Homo sapiens* | IH | 2007 | this study | MtG_h3 | HG3c | 0707-404 | OR911383 |
| FE-423 | NRC-E-423 | France | *Homo sapiens* | IH | 2008 | this study | MtG_h3 | HG3c | 0707-423 | OR911384 |
| FE-427 | NRC-E-427 | France | *Homo sapiens* | IH | 2010 | this study | MtG_h4 | HG3b | 0804-427 | OR911385 |
| FE-478 | NRC-E-478 | France | *Homo sapiens* | IH | 2010 | this study | MtG_h8 | HG3c | 10529-478 | OR911386 |
| FE-491 | NRC-E-491 | France | *Homo sapiens* | IH | 2019 | Bohard et al., 2023 | MtG_h3 | HG3c | 19746-491 | OQ599939 |
| FE-652 | NRC-E-652 | France | *Homo sapiens* | IH | 2018 | Bohard et al., 2023 | MtG_h3 | HG3c | 18685-652 | OQ599947 |
| FE-653 | NRC-E-653 | France | *Homo sapiens* | IH | 2018 | Bohard et al., 2023 | MtG_h9 | HG3a | 18818-653 | OQ599962 |
| FE-655 | NRC-E-655 | France | *Homo sapiens* | IH | 2017 | this study | MtG_h10 | HG3b | 17166-655 | OR911387 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| FE-735 | NRC-E-735 | France | *Homo sapiens* | IH | 2017 | this study | MtG_h11 | HG3b | 17792-735 | OR911388 |
| FE-749 | NRC-E-749 | France | *Homo sapiens* | IH | 2017 | Bohard et al., 2023 | MtG_h12 | HG3c | 17872-749 | OQ599942 |
| FE-755 | NRC-E-755 | France | *Homo sapiens* | IH | 2020 | Bohard et al., 2023 | MtG_h13 | HG3c | 20084-755 | OQ599950 |
| FE-763 | NRC-E-763 | France | *Homo sapiens* | IH | 2017 | this study | MtG_h14 | HG3b | 17124-763 | OR911389 |
| FE-766 | NRC-E-766 | France | *Homo sapiens* | IH | 2018 | Bohard et al., 2023 | MtG_h1 | HG3c | 18083-766 | OQ599940 |
| FE-779 | NRC-E-779 | France | *Homo sapiens* | IH | 2017 | this study | MtG_h15 | HG3b | 17857-779 | OR911390 |
| FE-780 | NRC-E-780 | France | *Homo sapiens* | IH | 2017 | Bohard et al., 2023 | MtG_h15 | HG3b | 17068-780 | OQ599953 |
| FE-781 | NRC-E-781 | France | *Homo sapiens* | IH | 2017 | this study | MtG_h16 | HG3c | 17062-781 | OR911391 |
| FE-782 | NRC-E-782 | France | *Homo sapiens* | IH | 2017 | this study | MtG_h17 | HG3a | 17853-782 | OR911392 |
| FE-790 | NRC-E-790 | France | *Homo sapiens* | IH | 2017 | this study | MtG_h18 | HG3a | 17077-790 | OR911393 |
| FE-791 | NRC-E-791 | France | *Homo sapiens* | IH | 2018 | Bohard et al., 2023 | MtG_h3 | HG3c | 18317-791 | OQ599948 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| FE-792 | NRC-E-792 | France | *Homo sapiens* | IH | 2018 | Bohard et al., 2023 | MtG_h12 | HG3c | 18488-792 | OQ599943 |
| FE-796 | NRC-E-796 | France | *Homo sapiens* | IH | 2018 | Bohard et al., 2023 | MtG_h17 | HG3a | 18328-796 | OQ599965 |
| FE-800 | NRC-E-800 | France | *Homo sapiens* | IH | 2018 | Bohard et al., 2023 | MtG_h15 | HG3b | 18250-800 | OQ599954 |
| FE-805 | NRC-E-805 | France | *Homo sapiens* | IH | 2018 | Bohard et al., 2023 | MtG_h3 | HG3c | 18718-805 | OQ599949 |
| FE-826 | NRC-E-826 | France | *Homo sapiens* | IH | 2018 | Bohard et al., 2023 | MtG_h1 | HG3c | 18544-826 | OQ599944 |
| FE-829 | NRC-E-829 | France | *Homo sapiens* | IH | 2018 | Bohard et al., 2023 | MtG_h19 | HG3b | 18022-829 | OQ599951 |
| FE-836 | NRC-E-836 | France | *Homo sapiens* | IH | 2019 | Bohard et al., 2023 | MtG_h20 | HG3a | 19104-836 | OQ599963 |
| FE-840 | NRC-E-840 | France | *Homo sapiens* | IH | 2022 | this study | MtG_h21 | HG3b | 22694-840 | OR911394 |
| FE-841 | NRC-E-841 | France | *Homo sapiens* | IH | 2019 | Bohard et al., 2023 | MtG_h4 | HG3b | 19674-841 | OQ599955 |
| FE-842 | NRC-E-842 | France | *Homo sapiens* | IH | 2021 | Bohard et al., 2023 | MtG_h22 | HG3c | 21798-842 | OQ599941 |
| FE-860 | NRC-E-860 | France | *Homo sapiens* | IH | 2019 | Bohard et al., 2023 | MtG_h17 | HG3a | 19518-860 | OQ599964 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| FE-878 | NRC-E-878 | France | *Homo sapiens* | IH | 2021 | Bohard et al., 2023 | MtG_h9 | HG3a | 21229-878 | OQ599961 |
| FE-879 | NRC-E-879 | France | *Homo sapiens* | IH | 2021 | Bohard et al., 2023 | MtG_h23 | HG3b | 21939-879 | OQ599956 |
| FE-884 | NRC-E-884 | France | *Homo sapiens* | IH | 2020 | Bohard et al., 2023 | MtG_h24 | HG3b | 20298-884 | OQ599960 |
| FE-888 | NRC-E-888 | France | *Homo sapiens* | IH | 2020 | Bohard et al., 2023 | MtG_h25 | HG3b | 20659-888 | OQ599959 |
| FE-889 | NRC-E-889 | France | *Homo sapiens* | IH | 2022 | this study | MtG_h15 | HG3b | 22986-889 | OR911395 |
| FE-890 | NRC-E-890 | France | *Homo sapiens* | IH | 2020 | Bohard et al., 2023 | MtG_h39 | HG2 | 20928-890 | OQ599968 |
| FE-892 | NRC-E-892 | France | *Homo sapiens* | IH | 2020 | Bohard et al., 2023 | MtG_h4 | HG3b | 20381-892 | OQ599957 |
| FE-900 | NRC-E-900 | France | *Homo sapiens* | IH | 2020 | Bohard et al., 2023 | MtG_h1 | HG3c | 20549-900 | OQ599945 |
| FE-907 | NRC-E-907 | France | *Homo sapiens* | IH | 2021 | Bohard et al., 2023 | MtG_h17 | HG3a | 21012-907 | OQ599966 |
| FE-908 | NRC-E-908 | France | *Homo sapiens* | IH | 2020 | Bohard et al., 2023 | MtG_h1 | HG3c | 20742-908 | OQ599946 |
| FE-921 | NRC-E-921 | France | *Homo sapiens* | IH | 2022 | this study | MtG_h26 | HG3c | 22963-921 | OR911397 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| FE-922 | NRC-E-922 | France | *Homo sapiens* | IH | 2022 | this study | MtG_h17 | HG3a | 22958-922 | OR911398 |
| FE-935 | NRC-E-935 | France | *Homo sapiens* | IH | 2021 | Bohard et al., 2023 | MtG_h27 | HG3b | 21006-935 | OQ599952 |
| FE-938 | NRC-E-938 | France | *Homo sapiens* | IH | 2021 | Bohard et al., 2023 | MtG_h17 | HG3a | 21516-938 | OQ599967 |
| FE-988 | NRC-E-988 | France | *Homo sapiens* | IH | 2022 | this study | MtG_h1 | HG3c | 22993-988 | OR911396 |
| S1 | St Lawrence Isl _ 3 | Alaska (St Lawrence island) | *Microtus oeconomus* | IH | 1995 | Knapp et al., 2007 | MtG_h40 | HG1 | SLI-S1 | OR911399 |
| S2 | St Lawrence Isl _ 5 | Alaska (St Lawrence island) | *Microtus oeconomus* | IH | 1995 | Knapp et al., 2007 | MtG_h40 | HG1 | SLI-S2 | OR911400 |
| S3 | St Lawrence Isl _ 6 | Alaska (St Lawrence island) | *Microtus oeconomus* | IH | 1995 | Knapp et al., 2007 | MtG_h40 | HG1 | SLI-S3 | OR911401 |
| S4 | StLawrence Isl _ 7 | Alaska (St Lawrence island) | *Microtus oeconomus* | IH | 1995 | Knapp et al., 2007 | MtG_h40 | HG1 | SLI-S4 | OR911402 |
| S5 | StLawrence Isl _ 9 | Alaska (St Lawrence island) | *Microtus oeconomus* | IH | 1995 | Knapp et al., 2007 | MtG_h40 | HG1 | SLI-S5 | OR911403 |
| S33 | SLI60 | Alaska (St Lawrence island) | *Homo sapiens* [a] | IH | 1990 | Gottstein et al., 1991 | MtG_h40 | HG1 | SLI-S33 | OR911417 |
| S31 | 27_Alaska.r | Alaska | *Microtus oeconomus* [a] | IH | 1989 | Bretagne et al., 1996 | MtG_h40 | HG1 | ALK-S31 | OR911415 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| S34 | SLI144 | Alaska (St Lawrence island) | *Homo sapiens* [a] | IH | 1990 | Gottstein et al., 1991 | MtG_h43 | HG2 | SLI-S34 | OR911418 |
| S64 | 13264 | Armenia | *Homo sapiens* | IH | ND | Umhang et al., 2021 | MtG_h51 | HG2 | ARM-S64 | OR911444 |
| S20 | 1-Human-HP-102039 | Belgium | *Homo sapiens* | IH | 2010 | Knapp et al., 2020 | MtG_h29 | HG3c | BLG-S20 | OR911408 |
| S16 | CND85_b43 | Canada | rodent | IH | 1985 | Knapp et al., 2007 | MtG_h41 | HG2 | CND-S16 | OR911406 |
| S35 | CND-84_b36 | Canada | rodent[1] | IH | <1980 | Knapp et al., 2007 | MtG_h41 | HG2 | CND-S35 | OR911419 |
| S36 | CND-85_b36 | Canada | rodent[1] | IH | <1980 | Knapp et al., 2007 | MtG_h41 | HG2 | CND-S36 | OR911420 |
| S17 | 52-Dog-09D2679 | Canada (British Columbia) | *Canis lupus familiaris* [b] | IH | 2009 | Peregrine et al., 2012 | MtG_h28 | HG3a | CND-S17 | OR911407 |
| S44 | Em Chine 1.2 | China | *Canis lupus familiaris* | IH | ND | this study | MtG_h45 | HG2 | CHN-S44 | OR911426 |
| S28 | 3_Germany.r | Germany (Stuttgart 50 km west) | *Ondatra zibethicus* | IH | 1984 | Bretagne et al., 1996 | MtG_h32 | HG3b | GER-S28 | OR911414 |
| S26 | Germany monkey 1 – 2022 | Germany | monkey | IH | 2022 | this study | MtG_h4 | HG3b | GER-S26 | OR911412 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| S25 | Em genomic DNA 01 | Japan (Hokkaido) | rodent | IH | <2009 | Knapp et al., 2011 | MtG_h42 | HG2 | JPN-S25 | OR911411 |
| S27 | 15_Japan.r | Japan (Hokkaido) | *Clethrionomys rufocanus* | IH | 1979 | Bretagne et al., 1996 | MtG_h42 | HG2 | JPN-S27 | OR911413 |
| S47 | A50 | Japan (Hokkaido) | *Vulpes vulpes* | DH | ND | this study | MtG_h42 | HG2 | JPN-S47 | OR911427 |
| S48 | E50 | Japan (Hokkaido) | *Vulpes vulpes* | DH | ND | this study | MtG_h46 | HG2 | JPN-S48 | OR911428 |
| S49 | I50 | Japan (Hokkaido) | *Vulpes vulpes* | DH | ND | this study | MtG_h42 | HG2 | JPN-S49 | OR911429 |
| S50 | M50 | Japan (Hokkaido) | *Vulpes vulpes* | DH | ND | this study | MtG_h42 | HG2 | JPN-S50 | OR911430 |
| S51 | N50 | Japan (Hokkaido) | *Vulpes vulpes* | DH | ND | this study | MtG_h42 | HG2 | JPN-S51 | OR911431 |
| S40 | LUX | Luxembourg | *Homo sapiens* | IH | 2021 | Bohard et al., 2023 | MtG_h4 | HG3b | 21169-LX1 | OQ599958 |
| S37 | Em1 Norv Spitzberg | Norway (Svalbard) | *Microtus levis* | IH | 2004-2006 | Knapp et al., 2012 | MtG_h44 | HG1 | SVB-S37 | OR911421 |
| S42 | Em7 Norv Spitzberg | Norway (Svalbard) | *Microtus levis* | IH | 2004-2006 | Knapp et al., 2012 | MtG_h44 | HG1 | SVB-S42 | OR911424 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| S43 | Em22 Norv Spitzberg | Norway (Svalbard) | *Microtus levis* | IH | 2004-2006 | Knapp et al., 2012 | MtG_h44 | HG1 | SVB-S43 | OR911425 |
| S52 | Fox3 | Poland (Warmińsko-Mazurskie) | *Vulpes vulpes* | DH | 2009-2013 | Karamon et al., 2017 | MtG_h47 | HG2 | POL-S52 | OR911432 |
| S53 | Fox7 | Poland (Zachnodniopomorskie) | *Vulpes vulpes* | DH | 2009-2013 | Karamon et al., 2017 | MtG_h34 | HG3a | POL-S53 | OR911433 |
| S54 | Fox12 | Poland (Lubuskie) | *Vulpes vulpes* | DH | 2009-2013 | Karamon et al., 2017 | MtG_h35 | HG3a | POL-S54 | OR911434 |
| S55 | Fox28 | Poland (Świętokrzyskie) | *Vulpes vulpes* | DH | 2009-2013 | Karamon et al., 2017 | MtG_h36 | HG3c | POL-S55 | OR911435 |
| S56 | Fox33 | Poland (Śląskie) | *Vulpes vulpes* | DH | 2009-2013 | Karamon et al., 2017 | MtG_h17 | HG3a | POL-S56 | OR911436 |
| S57 | Fox47 | Poland (Warmińsko-Mazurskie) | *Vulpes vulpes* | DH | 2009-2013 | Karamon et al., 2017 | MtG_h37 | HG3c | POL-S57 | OR911437 |
| S58 | Fox64 | Poland (Podlaskie) | *Vulpes vulpes* | DH | 2009-2013 | Karamon et al., 2017 | MtG_h17 | HG3a | POL-S58 | OR911438 |
| S59 | Fox78 | Poland (Mazowieckie) | *Vulpes vulpes* | DH | 2009-2013 | Karamon et al., 2017 | MtG_h47 | HG2 | POL-S59 | OR911439 |
| S60 | 13254 | Russia (Altaï) | *Homo sapiens* | IH | ND | Umhang et al., 2021 | MtG_h48 | HG2 | RUS-S60 | OR911440 |
| S61 | 13256 | Russia (Altaï) | *Homo sapiens* | IH | ND | Umhang et al., 2021 | MtG_h49 | HG2 | RUS-S61 | OR911441 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| S62 | 13259 | Russia (Altaï) | *Homo sapiens* | IH | ND | Umhang et al., 2021 | MtG_h50 | HG2 | RUS-S62 | OR911442 |
| S63 | 13260 | Russia (Altaï) | *Homo sapiens* | IH | ND | Umhang et al., 2021 | MtG_h50 | HG2 | RUS-S63 | OR911443 |
| S65 | 13269 | Russia (Altaï) | *Homo sapiens* | IH | ND | Umhang et al., 2021 | MtG_h52 | HG2 | RUS-S65 | OR911445 |
| S66 | 13270 | Russia (Altaï) | *Homo sapiens* | IH | ND | Umhang et al., 2021 | MtG_h53 | HG2 | RUS-S66 | OR911446 |
| S67 | 13271 | Russia (Altaï) | *Homo sapiens* | IH | ND | Umhang et al., 2021 | MtG_h54 | HG2 | RUS-S67 | OR911447 |
| S70 | 13276 | Russia (Ryazan Oblast) | *Meriones unguiculatus* | IH | ND | Umhang et al., 2021 | MtG_h55 | HG2 | RUS-S70 | OR911448 |
| S71 | 13277 | Russia (Ryazan Oblast) | *Meriones unguiculatus* | IH | ND | Umhang et al., 2021 | MtG_h38 | HG3c | RUS-S71 | OR911449 |
| S72 | 13278 | Russia (Ryazan Oblast) | *Meriones unguiculatus* | IH | ND | Umhang et al., 2021 | MtG_h56 | HG2 | RUS-S72 | OR911450 |
| S73 | 13279 | Russia (Anabar-Yakutia) | *Microtus gregalis* | IH | ND | Umhang et al., 2021 | MtG_h57 | HG1 | RUS-S73 | OR911451 |
| S74 | 13280 | Russia (Anabar-Yakutia) | *Microtus gregalis* | IH | ND | Umhang et al., 2021 | MtG_h57 | HG1 | RUS-S74 | OR911452 |
| S76 | 13286 | Russia (Olkhon Island-Irkutsk Oblast) | *Alticola olchonensis* | IH | 2010-2012 | Konyaev et al., 2013 | MtG_h58 | [c] | RUS-S76 | OR911453 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| S41 | SW1 | Sweden | *Homo sapiens* | IH | 2017 | this study | MtG_h17 | HG3a | SWD-S40 | OR911423 |
| S13 | ADN_42 | Switzerland (Zürich zoo) | *Chlorocebus pygerythrus* | IH | 1989 | Knapp et al., 2007 | MtG_h17 | HG3a | SWT-S13 | OR911404 |
| S21 | WR | Switzerland | *Homo sapiens* | IH | 2014 | this study | MtG_h30 | HG3b | SWT-S21 | OR911409 |
| S23 | SJ | Switzerland | *Homo sapiens* | IH | 2013 | this study | MtG_h31 | HG3b | SWT-S21 | OR911410 |
| S39 | 9_Züric h.h | Switzerland | *Homo sapiens* | IH | 1981 | Bretagne et al., 1996 | MtG_h9 | HG3a | SWT-S39 | OR911422 |
| S32 | CH5 | Switzerland | *Homo sapiens* [a] | IH | ND | Gottstein et al., 1991 | MtG_h33 | HG3a | SWT-S32 | OR911416 |
| S14 | ADN_43 | Switzerland (Zürich zoo) | *Chlorocebus pygerythrus* | IH | 1990 | Knapp et al., 2007 | MtG_h17 | HG3a | SWT-S14 | OR911405 |

[a] Maintained in rodents

[b] Dog with hepatic AE, from British Columbia (BC) province (Quesnel area), Canada, without stay abroad declared

[c] Unique representative of a new group: putative haplotype HG4

DH, definitive host; IH, intermediate host; VCF, variant call format; Acc. No, accession number

Table 2. Oligonucleotide sequences used to amplify the complete mitochondrial genome of *Echinococcus multilocularis*. Primer positions are given according to the *Echinococcus multilocularis* reference mitochondrial genome, GenBank accession number **AB018440.2**.

| Primer name | Primer sequence (5' - 3') | Primer position | PCR product length (bp) |
|---|---|---|---|
| E1f | TTCAAAGAGTGCTGGATTTC | 1081…1100 | 2153 |
| E1r | CAATACTCCAGTAAAAATTTGAAGAATC | 3206…3233 | |
| E2f | GAAGTAGTGGGTTTGTATTGTC | 3175…3196 | 1887 |
| E2r | CACCCAACAAACCTAACTTC | 5042…5061 | |
| E3f | CATGCTGAGGCTACTAGAG | 4992…5010 | 2072 |
| E3r | GTATCCCTTTAAATCTGTTGAATCTG | 7038…7063 | |
| E4f | GGGGGTTATGTAGAATTATTTATTTTG | 7002…7028 | 2098 |
| E4r | TACACCAACCAAAGAGAAAG | 9080…9099 | |
| E5f | ATAGGTGTGGTTAGAGAAAATTG | 9018…9040 | 1827 |
| E5r | CAGCTATACAACCAGGTTAATG | 10823…10844 | |
| E6f | ACCTGGTTGTATAGCTGTGAAG | 10828…10849 | 2099 |
| E6r | CACCAACAGGAAACGAATC | 12908…12926 | |
| E7f | GACTATAAAGATTGTTGGTCATCAG | 12847…12871 | 1997 |
| E7r | GAAAGGAAATCCAGCACTCTTTG | 1083…1105 | |

Table 3. Molecular and haplotypic diversity indices for 113 *Echinococcus multilocularis* mitogenome samples.

| Molecular parameters and index | Value |
|---|---|
| Sample size | 113 |
| Number of observed nucleotide sites | 13,747 |
| Number of polymorphic sites (%) | 436 (3.1%) |
| Number of parsimony informative sites | 207 |
| Number of singleton variable sites | 212 |
| Number of indels sites | 17 |
| Number of substitutions | 419 |
| Number of observed transitions | 348 |
| Number of observed transversions | 71 |
| Number of INDELs | 10 |
| GC content | 31% |
| π (mean number of pairwise differences) | 52.020494 ± 22.662985 |
| average over loci ± S.D. | 0.003785 ± 0.001826 |
| Hn (number of haplotypes) | 58 |
| Hd (haplotype diversity) ± S.D. | 0.971 ± 0.006 |

Table 4. Single nucleotide polymorphisms (SNPs) described for the *cox1* gene in samples belonging to the *Echinococcus multilocularis* Mongolian clade. Shaded columns report SNP loci present in all samples and which appear to be clade-specific. Positions are given according to (i) the *E. multilocularis* reference mitochondrial genome, GenBank accession number **AB018440.2**, and (ii) according to the *cox1* gene in **AB018440.2**.

| **ABO18440** | **A** | **A** | **T** | **A** | **G** | **T** | **A** | **A** | **T** | **G** | **C** | **A** | **T** | **C** | **T** | **G** | **T** | **C** | **A** | **G** | **A** | **T** | **G** | **G** | **A** | **A** | **T** | **T** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Position in AB018440** | **9221** | **9243** | **9329** | **9363** | **9440** | **9453** | **9497** | **9588** | **9650** | **9671** | **9852** | **9854** | **9899** | **9906** | **9959** | **10109** | **10179** | **10271** | **10316** | **10379** | **10413** | **10478** | **10493** | **10562** | **10598** | **10652** | **10673** | **10712** |
| **Position in *cox1* gene** | **57** | **79** | **165** | **199** | **276** | **289** | **333** | **424** | **486** | **507** | **688** | **690** | **735** | **742** | **795** | **945** | **1015** | **1107** | **1152** | **1215** | **1249** | **1314** | **1329** | **1398** | **1434** | **1488** | **1509** | **1548** |
| **AB510023 (9199-10746)** | . | G | C | . | A | C | T | . | G | . | T | G | G | . | . | . | . | T | G | A | G | G | A | A | T | G | C | C |
| **AB510025 (9165-10772)** | G | G | . | G | A | C | T | G | G | A | T | G | G | T | G | . | C | T | G | A | G | G | A | A | T | G | C | C |
| **AB777920 (9165-10772)** | G | G | . | G | A | C | T | . | G | A | T | G | G | T | G | . | C | T | G | A | G | G | A | A | T | G | C | C |
| **AB777921 (9165-10772)** | . | G | . | G | A | C | T | . | G | . | T | G | G | T | G | . | C | T | G | A | G | G | A | A | T | G | C | . |
| **AB461420 (9165-10772)** | . | G | C | G | A | C | T | . | G | A | T | G | G | T | G | . | C | T | G | A | G | G | A | A | T | G | C | C |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **AB813186 (9165-10772)** | **.** | G | . | G | A | C | T | . | G | . | T | G | G | T | G | . | C | T | G | A | G | G | A | A | T | G | C | . |
| **AB813187 (9165-10772)** | **.** | G | C | G | A | C | T | . | G | A | T | G | G | T | G | . | C | T | G | A | G | G | A | A | T | G | C | C |
| **AB813188 (9165-10772)** | **.** | G | . | G | A | C | T | . | G | . | T | G | G | T | G | T | C | T | G | A | G | G | A | A | T | G | C | . |
| **OR911453 (9165-10772) S76** | **.** | G | . | G | A | C | T | . | G | . | T | G | G | T | G | . | C | T | G | A | G | G | A | A | T | G | C | . |

Table 5. Comparison of epidemiological and clinical data between haplogroups in French patients.

| | | Sub-haplogroup | | | |
|---|---|---|---|---|---|
| | No of patients | HG3a | HG3b | HG3c | Statistical test |
| | | (*n*=10) | (*n*=21) | (*n*=25) | |
| Epidemiology | | | | | |
| Proportion of males | 56 | 0.60 (6/10) | 0.52 (11/21) | 0.48 (12/25) | SI [a] |
| Median age at diagnosis [min-max] | 56 | 66 [15-76] | 58 [11-81] | 57 [10-76] | K=2.96, SI [b] |
| AE diagnoses after 2013 | 56 | 1 (10/10) | 0.71 (15/21) | 0.52 (13/25) | *P*=0.016 [a] |
| Clinical status | | | | | |
| Immunosuppression | 43 | 0.40 (4/10) | 0.36 (5/14) | 0.37 (7/19) | SI [a] |
| Clinical form | | | | | |
| Hepatic form only | 52 | 0.80 (8/10) | 0.47 (9/19) | 0.52 (12/23) | SI [a] |
| Extra-hepatic form only | 55 | 0.00 (0/10 | 0.10 (2/21 | 0.17 (4/24) | SI [a] |
| Intra- and extra- hepatic forms | 52 | 0.20 (2/10) | 0.42 (8/19) | 0.35 (8/23) | SI [a] |

| | | | | | |
|---|---|---|---|---|---|
| Invasion neighbour organs/tissues | 47 | 0.20 (2/10) | 0.38 (6/16) | 0.33 (7/21) | SI [a] |
| Metastasis | 44 | 0.00 (0/9) | 0.27 (4/15) | 0.25 (5/20) | SI [a] |

[a] Fisher's exact test, [b] Kruskall-Wallis test.

AE, alveolar echinococcosis; SI, statistically insignificant.