PROCEEDINGS OF
THE ROYAL SOCIETY B
BIOLOGICAL SCIENCES

It's a small world for parasites; evidence supporting the North American invasion of European Echinococcus multilocularis.

Journal:	*Proceedings B*
Manuscript ID	RSPB-2023-0128.R1
Article Type:	Research
Date Submitted by the Author:	13-Feb-2023
Complete List of Authors:	Santa, Maria; University of Calgary, Biological Sciences Umhang, Gerald; National Agency for Food Environmental and Occupational Health and Safety Rabies and Wildlife Laboratory Nancy Klein, Claudia; University of Calgary; Friedrich-Loeffler-Institut Bundesforschungsinstitut fur Tiergesundheit Grant, Danielle; University of Bergen; University of Calgary, Department of Clinical and Veterinary Clinical Sciences Ruckstuhl, Kathreen; University of Calgary, Biological Sciences Musiani, Marco; University of Bologna, Dipartimento di Scienze Biologiche, Geologiche e Ambientali (BiGeA) Gilleard, John; University of Calgary, Department of Comparative Biology & Experimental Medicine, Host-Parasite Interactions Program Massolo, Alessandro; University of Pisa; University of Calgary, Department of Ecosystem and Public Health; Chrono-environnement, UMR CNRS 6249 Chrono-Environnement
Subject:	Ecology < BIOLOGY, Health and Disease and Epidemiology < BIOLOGY, Genetics < BIOLOGY
Keywords:	Parasite invasion, Echinococcus multilocularis, North America, Genetic characterization, Allochthonous strains
Proceedings B category:	Ecology

SCHOLARONE™
Manuscripts

Author-supplied statements

Relevant information will appear here if provided.

Ethics

Does your article include research that required ethical approval or permits?:
This article does not present research with ethical considerations

Statement (if applicable):
CUST_IF_YES_ETHICS :No data available.

Data

It is a condition of publication that data, code and materials supporting your paper are made publicly available. Does your paper present new data?:
Yes

Statement (if applicable):
All data generated or analyzed during this study are included in this article (and its supplementary information files). The dataset of EmsB profiles is available in the Figshare repository, (doi: 10.6084/m9.figshare.16818607).

Conflict of interest

I/We declare we have no competing interests

Statement (if applicable):
CUST_STATE_CONFLICT :No data available.

1 **It's a small world for parasites; evidence supporting the North American invasion of**
2 **European *Echinococcus multilocularis*.**

3 Maria A. Santa[1], Gerald Umhang[2], Claudia Klein[3,4], Danielle M. Grant[3,5], Kathreen E. Ruckstuhl[1], Marco
4 Musiani[6], John S. Gilleard[7], Alessandro Massolo[8,9,10](*).

5

6 [1]Department of Biological Sciences, Faculty of Science, University of Calgary, Calgary, Canada

7 [2]Nancy Laboratory for Rabies and Wildlife, National Reference Laboratory for Echinococcus spp., Wildlife
8 Surveillance and Eco-epidemiology Unit, ANSES, Technopôle Agricole et Vétérinaire, Malzéville, France.

9 [3]Department of Clinical and Veterinary Clinical Sciences, Faculty of Veterinary Medicine, University of
10 Calgary, Calgary, Canada

11 [4]Friedrich-Loeffler-Institut, Bundesforschungsinstitut für Tiergesundheit, Federal Research Institute for
12 Animal Health, Neustadt, Germany

13 [5]NORCE Climate, NORCE Norwegian Research Centre, Bjerknes Centre for Climate Research, Bergen,
14 Norway

15 [6]Dipartimento di Scienze Biologiche, Geologiche e Ambientali (BiGeA), University of Bologna, Italy

16 [7]Department of Comparative Biology & Experimental Medicine, Host-Parasite Interactions Program,
17 Faculty of Veterinary Medicine, Host-Parasite Interactions (HPI) Research Training Network, University of
18 Calgary, Canada

19 [8]Ethology Unit, Department of Biology, University of Pisa, Pisa, Italy

20 [9]Department of Ecosystem and Public Health, Faculty of Veterinary Medicine, University of Calgary,
21 Calgary, Canada

22 [10]UMR CNRS 6249 Chrono-environnement, Université Bourgogne Franche-Comté, Besançon, France

23 *Corresponding author

24 **Email:**alessandro.massolo@unipi.it

25

26

27

28

29

30

Abstract

Echinococcus multilocularis (*Em*), the causative agent of human alveolar echinococcosis (AE), is present in the Holarctic region, and several genetic variants deem to have differential infectivity and pathogenicity. An unprecedented outbreak of human AE cases in Western Canada infected with a European-like strain circulating in wild hosts warranted assessment of whether this strain was derived from a recent invasion or was endemic but undetected. Using nuclear and mitochondrial markers, we investigated the genetic diversity of *Em* in wild coyotes and red foxes from Western Canada, compared the genetic variants identified to global isolates, and assessed their spatial distribution to infer possible invasion dynamics. Genetic variants from Western Canada were closely related to the original European clade, with lesser genetic diversity than that expected for a long-established strain and spatial genetic discontinuities within the study area, supporting the hypothesis of a relatively recent invasion with various founder events.

Keywords: Parasite invasion, *Echinococcus multilocularis*, North America, genetic characterization, allochthonous strains

1. Introduction

In the current Anthropocene epoch, biological invasions, due to trans- and inter-continental movements, are generating a global biotic homogenization, influencing at the same time global patterns of disease [1, 2]. Such invasions can lead to the introduction of novel hosts and their parasites, opening up opportunities for the emergence of diseases that can lead to outbreaks or the establishment of new endemism [3]. Indeed, many aetiological agents of emerging infectious diseases can be considered biological invaders [4]. As an example, the cestode *Echinococcus multilocularis* (*Em*), the causative agent of Alveolar Echinococcosis (AE), is considered an emerging or re-emerging disease in some regions, becoming an important public health concern in many countries worldwide [5, 6]. This parasite is distributed across the northern hemisphere and is transmitted in sylvatic and semi-synanthropic cycles, involving small mammals as intermediate hosts (primarily arvicoline rodents), and canids as definitive hosts [e.g., red foxes (*Vulpes vulpes*), coyotes (*Canis latrans*) and dogs]. In small mammals, the parasite invades and develops in internal organs (primarily the liver), being highly pathogenic. Therefore, the host either dies from the infection or is easily preyed on by the definitive hosts [6]. Conversely, wild canids only develop an intestinal infection, causing no substantial pathology [7]. Humans can act as dead-end hosts and develop AE by accidental ingestion of eggs via contaminated food, water, or through contact with feces of infected definitive hosts [8].

Although the geographic distribution of *E. multilocularis* is limited to the Northern hemisphere, its distribution and prevalence in wild and domestic hosts are increasing across its range, likely due to human activities [9-11]. The parasite has been reported in new areas previously considered non-endemic [12-15]. Furthermore, synanthropic hosts and domestic animals now have important roles in its transmission, contributing to the establishment of urban cycles [16, 17]. Additionally, increased translocation of domestic and wild animals has generated new host species assemblages and introduced *E. multilocularis* (or particular genetic variants) into new areas [9, 18, 19].

Using concatenated sequences of three mitochondrial genes (*cox1*, *nad2* and *cob*), *Em* genetic variants have been grouped into four clades, mostly linked with their geographical origins: Asian, European, North American, and one restricted to Mongolia [20]. However, Asian haplotypes have also

74 been reported in Western Russia [21], Saint Lawrence Island (Alaska) [20] and, more recently, in central-
75 eastern Europe (Poland) [22, 23]. In addition, haplotypes clustering most closely with the North American
76 clade have also been detected in Siberia [21]. In North America, the situation is similar, with genetic
77 variants linked to geographically distant regions. In recent years, European-like haplotypes have been
78 detected in Western Canada (Alberta, British Columbia, and Saskatchewan) in wild and domestic hosts [12,
79 18, 24-26]. Additionally, despite a historically low incidence of human AE in continental North America,
80 with only two locally acquired cases ever reported (in 1923 and 1977, respectively), at least 17 cases have
81 been described in the province of Alberta since 2013 [25, 27], along with a case recently confirmed in the
82 adjacent province of Saskatchewan [28], and more recently, two cases in the US [29, 30]. Molecular
83 characterization of six of the Albertan human cases implicated a European-like haplotype (ECA) that was
84 also detected in local wild hosts [25, 26]. Similarly, European-like haplotypes (different from the ones
85 found in Canada) were identified in the two human cases in the US [29, 30]. However, the distribution of
86 European-like strains, their origin in North America, and potential interactions (e.g., competition or
87 hybridization) with native North American genetic variants have not been elucidated. It has been suggested
88 that the European strain could be more pathogenic and/or virulent than the North American strain, based on
89 the historically low incidence of human cases of AE reported for the region. Consequently, it is of great
90 importance to genetically characterize *E. multilocularis* in Western Canada, to assess the extent of the
91 spread of European strains, their distribution in the main definitive hosts (coyotes and red foxes), and to
92 elucidate potential sources of invasion of these strains.

93 In Canada, genetic characterization of *E. multilocularis* has been conducted primarily using
94 mitochondrial (mtDNA) markers, yet these markers have relatively low variability. Conversely, the
95 microsatellite EmsB, a multilocus nuclear DNA marker, has greater discriminatory power than classic
96 mtDNA markers, and has been used to identify spatio-temporal characteristics of *E. multilocularis*
97 transmission in several European countries, at continental, national, and local geographical scales [31-34].
98 Additionally, the combination of mitochondrial and nuclear markers had greater discriminatory power to
99 identify genetic profiles and detect potential introgression events among genetic variants [22].

100 With this study, we aimed to: *i.* assess the genetic diversity of *E. multilocularis* in Western Canada, based
101 on EmsB profiles and mtDNA haplotypes; *ii*. investigate genetic relationships with *E. multilocularis*
102 isolates from other geographic regions across the globe, to identify possible sources of invasion; and *iii.*
103 evaluate the spatial distribution of the detected genetic variants, to elucidate the spread dynamics of the
104 European-like strains, possible spatio-temporal scenarios of its invasion process, and the likely origins of
105 these European strains in Western Canada.

106

107 **2. Materials and Methods**

108

109 **(a) Parasite collection and DNA extraction**

110 *E. multilocularis* specimens were collected from gastrointestinal (GI) tracts of red foxes and coyotes of
111 either road-killed or trap-harvested animals (trapped for purposes independent of this study), collected
112 between 2012 and 2017 in Western Canada. Trapped animals were obtained from licensed trappers with the
113 collaboration of the Alberta Trappers Association. GI tracts were screened using a modification of the
114 scraping, filtration, and counting technique (SFCT), to identify and collect *Echinococcus* spp. worms [35,
115 36]. We analyzed *Em* worms from 70 coyotes and 13 foxes from northern, central and southern Alberta
116 (AB); four coyotes from northwest British Columbia (BC); and 10 coyotes from southeast Saskatchewan
117 (SK). Extraction of DNA was performed on up to five individual worms per host using the Nucleospin 96

118 Tissue Kit (Macherey-Nagel, Germany) for samples processed in France (Anses Nancy Laboratory for
119 Rabies and Wildlife), and the E.Z.N.A. MicroElute Genomic DNA Kit (Omega Bio-tek, US), for samples
120 processed in Canada (University of Calgary, Faculty of Veterinary Medicine). Extraction was performed
121 following the manufacturer's instructions, and DNA was stored at − 20 °C until processed.

122 **(b) Genetic characterization using mtDNA**

123 Genetic diversity of *E. multilocularis* was characterized by sequencing the genes *nad2* (1068 bp), *cob* (882
124 bp) and *cox1* (1608 bp). We used one to five worms per host, depending on worm availability and quality
125 (intactness), using primers and PCR parameters as previously described (Nakao *et al*., 2009). Sequences
126 obtained from the same *Em* worms from our previous studies were also included [25, 26]. The sequences
127 obtained per genetic locus were concatenated and aligned in Geneious 10.0.9 (Biomatters Ltd, New
128 Zealand) and compared to nucleotide databases using the NCBI Nucleotide BLAST tool to identify the
129 strain/haplotype (https://blast.ncbi.nlm.nih.gov). To analyze genetic relationships and possible origins of
130 haplotypes identified in Western Canada, we constructed a haplotype network, based on hamming distance,
131 including reported European, Asian, Mongolian and North American haplotypes (based on full sequences
132 of *cob*, *nad2* and *cox1*) [20, 24, 25, 29]. R statistical software (R Development Core Team, 2022) and the
133 package *pegas 0.14* [37] were used to construct the haplotype network. A phylogenetic tree was built via
134 partitioned Bayesian analysis using MrBayes 3.2.7 [38], considering three partitions, corresponding to *cob*,
135 *nad2* and *cox1*, and using *E. granulosus sensu stricto* (G1) as the outgroup. Reversible-jump Markov chain
136 Monte Carlo (MCMC) analyses were run for 1 million generations, producing 10,000 trees, with the first
137 2,500 regarded as burn-in. The tree was plotted using the R-package *phangorn 2.3.1* [39].
138

139 **(c) Characterization of EmsB profiles**

140 A fluorescent PCR assay was used to amplify the EmsB microsatellite, following a previous protocol [34].
141 Fragment analysis of PCR products was performed by capillary electrophoresis on an automatic sequencer
142 (ABI Prism 310 and ABI 3500/3730; Life Technologies, CA). The resulting EmsB electropherograms were
143 composed of several peaks between 209 and 241 bp. The size (base-pair length) and height (fluorescent
144 signal intensity) of peaks present in each EmsB electropherogram were determined using GeneMapper 5.0
145 (Life Technologies, CA). Characterization of each EmsB profile was performed as described in the EmsB
146 guidelines from the EWET database (https://ewet-db.univ-fcomte.fr/) [40]. Briefly, peaks below 10% of the
147 sample's maximum peak height were considered artifacts and discarded. To normalize raw data, the height
148 of each peak was divided by the sum of the height of all peaks of a given sample.

149 *EmsB genotyping, clustering, and ordination analysis*

150 Genotyping and hierarchical cluster analysis were performed by calculating the Euclidean distance between
151 profiles and using the average link clustering method UPGMA [41]. The uncertainty of clusters was tested
152 using multiscale bootstrap resampling (1000 bootstrap replicates), obtaining approximately unbiased *p*-
153 values (au) [42]. A dendrogram was built using *E. granulosus sensu stricto* (G1) as an outgroup. A distance
154 threshold of 0.08 (average genetic distance observed after three generations in the rodent *Meriones*
155 *unguiculatus*) was used to identify unique EmsB profiles; any branching below this threshold was
156 considered genetically identical [43]. Unique EmsB profiles, represented by only one sample were
157 excluded from the analysis as they could not be technically validated [44]. To understand the structure of
158 the clusters detected and unravel relational patterns among the genetic profiles, we performed a non-metric
159 multidimensional scaling (NMDS) analysis based on the Euclidean distance matrix, using two and three
160 dimensions to visualize the genetic distance between samples and their arrangement in reduced dimensions
161 [45].

162 To evaluate genetic relationships with genotyped *E. multilocularis* isolates from Europe (historical
163 and peripheral endemic areas), Asia (Japan, China and Kyrgyzstan) and North America (Canada, Alaska),
164 each unique EmsB profile identified was compared to a reference world collection of profiles from the
165 EWET database (https://ewet-db.univ-fcomte.fr; updated until 2017) [40] and from profiles reported after
166 2017 [33, 44]. Profiles of 1275 samples from 17 countries were first compared to profiles obtained from
167 our study; thereafter, the five profiles with the lowest genetic distance to each of our profiles (P1-P16) were
168 used to build a dendrogram. Hierarchical cluster analysis was performed using R package *pvclust* [46] and
169 NMDS was done using Primer & Permanova + add-on, Version 6 (PRIMER-E Ltd.).

170 *Diversity analysis of EmsB profiles from Alberta*
171 We evaluated the diversity of EmsB profiles in the province of Alberta (AB), the most extensively sampled
172 area in our study. We divided this area into six geographical subregions: North-West (AB-NW), North-East
173 (AB-NE), Central-West (AB-CW), Central-East (AB-CE), South-West (AB-SW, and South-East (AB-SE).

174 We used an integrated approach based on the framework of Hill numbers to assess: (a) sample
175 completeness, (b) asymptotic diversity estimates to infer true diversities, (c) non-asymptotic standardization
176 via rarefaction and extrapolation, and (d) evenness of each subregion (except for AB-NW due to low
177 sample size), following the methodology described by Chao, *et al*., 2020 [47]. First, to evaluate if sample
178 effort was sufficient and assess the extent of undetected diversity, sample completeness was estimated per
179 each subregion based on orders $0 \geq q \leq 2$, where q is a number that determines the measures' sensitivity to
180 "species" (i.e., EmsB profiles) abundances. Second, asymptotic diversity estimates were calculated using a
181 sample-size-based rarefaction and extrapolation sampling method [48]. The sample was extrapolated to
182 double the size of the observed sample, if the curve stabilized and levelled off, then the asymptotic
183 estimates were used to infer true diversity. Here, Hill numbers for order $q \geq 0$ included the three most
184 widely used diversity measures (i.e., Chao1 richness estimator, exponential Shannon diversity and inverse
185 Simpson diversity, as special cases of orders $q = 0$, 1 and 2, respectively). Third, if data did not contain
186 sufficient information to accurately infer the true diversity, this was inferred for a standardized sample
187 coverage [49]. Here, we calculated non-asymptotic coverage-based rarefaction and extrapolation estimates
188 for diversity orders $q = 0$, 1 and 2. The standardized sample coverage (i.e., equal fraction of an
189 assemblage´s individuals; $C_{\max}$) was selected as the minimum among the coverage values for samples
190 extrapolated to double the size of the reference sample. Lastly, the evenness (i.e., distribution of EmsB
191 profiles abundances) was calculated for each subregion. A bootstrap method ($n = 100$) was applied to
192 obtain the associated 95% confidence intervals for all estimates.

193 Analysis of β-diversity was performed based on the Bray-Curtis dissimilarity index and complete
194 linkage agglomerative clustering for hierarchical agglomeration [41]. Diversity of EmsB profiles in Alberta
195 was compared to reported diversity in historical European endemic areas (i.e., Switzerland, Germany,
196 Czech Republic, Austria, and France) [31, 50]. Diversity estimates were recalculated for these areas based
197 on the standardized sample coverage ($C_{\max}$). Analyses were performed using the R package *iNEXT- 4steps*
198 [47] and *vegan* [51].
199

200 *Analysis of the spatial distribution of EmsB profiles*
201 Alberta is divided into 174 wildlife management units (WMUs). Samples were collected from 38 WMUs
202 and the location of each host was assigned as a random point within a radius of 15 km (minimum radius of
203 a WMU) from the centroid point of the WMU where it was collected. This methodology was used, as the
204 exact geo-reference (latitude, longitude) for all the hosts was not available. A Mantel test was used to
205 assess the hypothesis of genetic isolation by geographical distance, comparing a matrix of the genetic
206 distance between EmsB profiles, and a matrix of the geographical distances between samples based on

207 Euclidean pairwise distances [52]. A Mantel correlogram was used to investigate the underlying structure
208 of the relationship, and to measure the correlation between each class of distances [41]. A distance-based
209 redundancy analysis (dbRDA), based on Bray-Curtis dissimilarity matrix of EmsB profiles, was performed
210 using the two spatial variables (latitude and longitude) to estimate the proportion of the genetic variation
211 explained by spatial structures [53]. Statistical significance of the model and variables was evaluated with a
212 permutation test. Analyses were done using the R package *vegan* [51] and Primer & Permanova + add-on,
213 Version 6 (PRIMER-E Ltd.).
214
215
216 **3. Results**
217
218 **(a) Genetic relationships of mtDNA haplotypes from North America supported the hypothesis of**
219 **multiple invasion events of European origin**
220 To characterize the genetic diversity of *E. multilocularis* based on mtDNA, we analyzed 96 concatenated
221 sequences of *cox1*, *nad2* and *cob* genes in adult worms recovered from 13 foxes and 77 coyotes. Virtually
222 all (93/96) were identified as the previously described European-like haplotypes ECA, EAB, ESK, ESK2
223 and BC1 [25, 26], and only three as the North American haplotype N2 (in two coyotes and one fox).
224 Furthermore, the percent identity to previously described haplotypes was 100%, with no new SNPs
225 identified. The haplotypes ECA, EAB and N2, were present only in Alberta; ESK2 was in Saskatchewan
226 and Alberta; ESK only in Saskatchewan; and BC1 only in British Columbia. The ECA haplotype was the
227 most prevalent, with a frequency of 78.1% (75/96), followed by ESK2 with a frequency of 11.5% (11/96).
228 Haplotypes BC1, EAB and N2 had a frequency of 3.1% (3/96) each, and only one sample was identified as
229 ESK. The haplotype network showed that the European-like haplotypes found in Western Canada were
230 closely related to the European clade, with few mutational steps (range, 1 to 6) between them, and that
231 these were more similar to the European haplotype E4 (AB461395.1, AB461404.1, AB461414.1) than any
232 other, suggesting a recent origin derived from the European clade. The most genetically distant haplotype
233 was one from British Columbia (BC1), with six mutational steps to E4, followed by one from South-East
234 Alberta and Saskatchewan (ESK2). Conversely, the haplotypes ECA, EAB, and ESK were closely related,
235 differing by only one or two mutations, having the lowest number of mutational steps to E4 (figure 1*a*),
236 indicating a possible stem from a single invasion event. The phylogenetic tree confirmed the close
237 relationship between the European-like haplotypes in Western Canada and the original European clade.
238 Within this clade, the three similar haplotypes (EAB, ECA and ESK) were in the same branch with the E4,
239 whereas BC1 was in another (along with the haplotype SK1), and more deeply branched, indicating a
240 distinct most recent common ancestor. These branching nodes were supported by the Bayesian posterior
241 probabilities (figure 1*b*).

242 **(b) Cluster and ordination analysis of EmsB profiles from Western Canada confirms the invasion**
243 **hypothesis**
244 Through fragment analysis of the multi-locus EmsB microsatellite, 204 individual EmsB profiles were
245 computed and technically validated [40] (doi: 10.6084/m9.figshare.16818607). Comparing the similarity of
246 these 204 profiles, we identified 16 unique EmsB profiles based on the 0.08 threshold average genetic
247 distance. These profiles formed three significant clusters: the first with samples from AB only, the second
248 with samples from AB and SK, and the third with samples from AB and BC (figure 2*a*).

249 Two profiles (P13 and P15; exclusively from AB) were the most frequent, each representing
250 32.8% of the worms. When comparing the two definitive hosts, profile P13 was present in 69% of foxes
251 and 64% of coyotes, and was the most prevalent in both hosts, representing 56% and 45% of worms in
252 foxes and coyotes, respectively. Profile P4, the third most common (11.8%), was present in AB and SK.

253 Profile P11 was only in BC and AB, corresponding to 5.9% of the worms. The other profiles represented 1
254 to 3.4% of worms and were primarily from AB (electronic supplementary material, table S1).

255 When comparing our profiles with profiles from the EWET world database (1275 samples, from
256 17 countries) to evaluate genetic relationships with other *E. multilocularis* isolates, 13 of our 16 unique
257 profiles (81.2%) clustered with isolates from Europe, with eight being genetically indistinguishable (based
258 on the 0.08 threshold) to profiles present in several European countries (i.e., Germany, France, Poland,
259 Austria, Switzerland, Denmark, Slovakia, and Czech Republic), and only three profiles clustering with
260 samples from the Arctic [i.e., Saint Lawrence Island (Alaska) and Svalbard archipelago (Norway)] (figure
261 2*b*; electronic supplementary material, table S2).

262 When the NMDS ordination analysis was performed, the representation of the genetic distances
263 between profiles in two dimensions had the lowest stress (0.05; indicating the best representation), with the
264 three clusters identified being linked with the location (i.e., provinces) (figure 3*a*) and the mtDNA
265 haplotypes (figure 3*b*).

266 **(c) Genetic diversity of EmsB profiles in Alberta was lower than would be expected in endemic areas**
267 The sample completeness profiles increased with diversity order ($q \geq 0$), implying that there was undetected
268 diversity (mostly less abundant EmsB profiles) in all subregions (except for AB-CW), and nearly all
269 abundant EmsB profiles had been detected. The areas with the lowest sample coverage ($q = 1$) were AB-NE
270 and AB-CE (89%) and the area with the highest coverage was AB-CW (100%). When Alberta was
271 analyzed as a single area, the sample completeness profile was 100% for all diversity orders (electronic
272 supplementary material, table S3).

273 The size-based rarefaction and extrapolation analysis revealed that for each subarea the sampling
274 curve stabilized for orders $q = 1$ and 2, but not for $q = 0$ (electronic supplementary material, figure S1),
275 implying that the asymptotic diversity estimates could be used to infer true diversities (Shannon and
276 Simpson), but not to estimate true genetic richness. When comparing the asymptotic estimates for $q = 1$ and
277 2 (electronic supplementary material, table S3), the highest diversity indices were found in AB-SW (5.9,
278 4.5), and the lowest in AB-CW (2.5, 2.4) and AB-SE (3.1, 2.2), with similar values for the other
279 subregions.

280 Since data were insufficient to infer the true richness, non-asymptotic coverage-based rarefaction
281 and extrapolation curves were calculated (figure 4) and measures were computed up to a standardized
282 coverage value of C_{max}= 94.5% (electronic supplementary material, table S3). For this C_{max}, the
283 corresponding highest richness estimate ($q = 0$) was found in AB-CE (8.2) and the lowest in AB-CW (2.7).
284 For evenness profiles, all values of q showed the highest evenness of EmsB profiles in AB-CW and the
285 lowest in AB-SE.

286 Analysis of β-diversity was performed by constructing a dendrogram based on the Bray-Curtis
287 distance to compare genetic diversity between areas. There was a higher similarity between the central and
288 northern regions, and a second branch grouping for the southern regions (electronic supplementary
289 material, figure S2), with low dissimilarity among all regions (0.27 to 0.53). When comparing non-
290 asymptotic diversity indices from subregions in Alberta with diversity reported in historical endemic areas
291 from Europe (recalculated based on C_{max} = 94.5%), the range of values (min. and max.) for all q orders was
292 higher in European countries than in most subregions in Alberta. Moreover, when Alberta was analyzed as
293 a single area, estimates for orders $q = 1$ and $q = 2$ were lower than in all areas in Europe (except for South
294 Germany) (electronic supplementary material, table S3). The same pattern was observed for the evenness

295 profiles, with Alberta being the most uneven area (0.2 - 0.3), which indicates the presence of a dominant
296 EmsB profile and low diversity.

297 **(d) Spatial distribution of EmsB profiles in Western Canada**

298 Profiles were spatially clustered, with profiles P13 and P15 found exclusively in Alberta (figure 5), and
299 those from BC (P6 and P11) mostly localized only there, except for P11 which was found also in northern
300 AB, close to the BC border. Similarly, in southern SK, only P4 and P5 were detected, with both profiles
301 also present in southern AB. No profiles from BC were detected in SK, or vice versa. To assess the
302 hypothesis of genetic isolation by geographical distance, we used a Mantel test comparing a genetic
303 distance matrix between EmsB profiles and a matrix of the geographical distances between samples, based
304 on Euclidean pairwise distances [52]. There was no significant correlation between genetic and
305 geographical distances when testing samples only from AB ($r = 0.02$, $p = 0.30$). However, the Mantel test
306 yielded a significant correlation ($r = 0.4$, $p < 0.001$) when including samples from BC and SK in the
307 analysis. The Mantel correlogram indicated a significant positive correlation (more similar samples) within
308 50 to 550 km, but a negative correlation (more dissimilar samples) when > 600 km apart (electronic
309 supplementary material, figure S3*a*).

310 In a distance-based redundancy analysis (dbRDA) using the two spatial variables, latitude and
311 longitude, 21.1% of the variance of the genetic matrix was explained by geographic distance (for samples
312 from all three provinces). Still, only the longitude, the axis dividing provinces, was significant ($F = 30.36$, p
313 $< .001$) (electronic supplementary material, figure S3*b*).
314
315
316 **4. Discussion**
317
318 In this study, we investigated the origin of the European-like strains of *E. multilocularis* in Western
319 Canada, its spread dynamics, and potential spatio-temporal scenarios of their process of invasion in North
320 America. We assessed the genetic diversity of the parasite in this region, comparing EmsB profiles and
321 mtDNA sequences; most genetic variants from Western Canada were closely related to the original
322 European clade. Furthermore, the diversity of the European genetic variants found in Western Canada was
323 low compared to that expected for a long-established endemic strain, supporting the hypothesis of relatively
324 recent introductions, and discounting the hypothesis of North America as an ancestral endemic focus for
325 this strain. Moreover, the spatial analysis indicated genetic discontinuity, only evident over large
326 geographic scales, suggesting various introduction events in Western Canada.

327 When using mtDNA as a genetic marker, most haplotypes detected were closely related to the
328 European clade, with few point mutations from European isolates [18, 24-26]. However, there was a
329 limited number of European samples with sequences available for the three mtDNA genes. Therefore,
330 although not likely, it cannot be ruled out that haplotypes from Western Canada may also be present in
331 European countries but not yet detected. The Bayesian phylogenetic tree implied different evolutionary
332 origins and pathways of the haplotypes found in Western Canada. The ECA, EAB, and ESK were grouped
333 in the same branch, whereas BC1 was in a different branch and showed a higher divergence with the closest
334 haplotype SK1. This division was also evident in their geographical distribution, as BC1 was detected in
335 British Columbia, but not in Saskatchewan, whereas ECA and EAB were only in Alberta. These results
336 suggest different spatio-temporal scenarios with multiple introductions of the European strains to Western
337 Canada, possibly occurring at different times in the last centuries, causing worm populations to be more or
338 less isolated. Indeed, various sources of invasion have been proposed, including the translocation of
339 domestic dogs from European endemic areas [12], introduced European red foxes imported for sport

340 hunting [54], and/or introduction of the parasite via translocation of intermediate hosts with international
341 shipping [9].

342 During the process of invasion, the likelihood of successful establishment will strongly depend on
343 the propagule pressure (i.e., the number of introduction events and the number of infective stages released).
344 Therefore, the repeated release of a large number of individuals in multiple locations can facilitate the long-
345 term establishment and increase the rate of spread of non-native populations [55]. However, multiple
346 introductions and high genetic variation do not seem to be indispensable for a successful invasion [56]. For
347 a parasite like *E. multilocularis*, reproductive traits (i.e., hermaphroditism, asexual multiplication in
348 intermediate hosts, self- and cross-fertilization, and development of thousands of sexually mature adult
349 worms in each definitive host) could have also been determinant in the first stages of invasion, boosting
350 local spread, but producing in turn geographic isolation of the different genotypes introduced.

351 The EmsB microsatellite enabled detection of more profiles (16 in total) in comparison to only
352 seven mtDNA haplotypes, helping to identify local genetic variation and trace the invasion process of the
353 parasite. When compared to a world database, most of the profiles we found (13 of 16) clustered within the
354 European group, having a low genetic distance to European isolates (ranging from 0.05 to 0.1). Moreover,
355 half of the profiles were genetically indistinguishable from known European ones, supporting the
356 hypothesis of a recent introduction of these European genetic variants. In contrast, the remaining three
357 EmsB profiles (P1-P3), which were genotyped as North American haplotype N2 based on mtDNA,
358 clustered within the Arctic clade and were genetically more distant from the Alaskan and Svalbard isolates
359 (0.21 to 0.42), suggesting a more prolonged genetic isolation, consistent with the greater divergence
360 between the haplotypes N1 and N2 [20, 57].

361 Complete agreement between the two markers, in correlation with the geographic location, is not
362 guaranteed. Umhang, Knapp 2021 [22] described samples from Poland clustering within the European
363 clade, based on mtDNA, whereas they were genotyped as Asian-type profiles using EmsB, suggesting
364 some degree of introgression between the two strains. Regardless, these two markers facilitated
365 differentiation between North American and European genetic variants, with no evidence of introgression
366 between them. All EmsB profiles clustered within the European clade were identified as European-like
367 haplotypes based on mtDNA, whereas EmsB profiles identified as North American strains, were in a
368 different cluster. Moreover, most mtDNA haplotypes were represented by more than one EmsB profile
369 from the same cluster, with samples from Saskatchewan clustering apart from samples from British
370 Columbia, as observed with mtDNA haplotypes.

371 In Europe, the expansion history of the parasite from the historical endemic core area into
372 peripheral regions has been governed by a mainland-island system of transmission, in which the ancestral
373 focus in central Europe served as a ‘mainland’ supplying the peripheral areas (islands), perhaps due to
374 dispersals of fox populations. From a genetic perspective, this resulted in invasion events with introduction
375 of only a few genetic profiles, resulting in low genetic diversity in the colonized region compared to greater
376 diversity in the ancestral endemic foci [33, 50, 58]. In our study, the genetic diversity indices recalculated
377 for Switzerland (endemic area), based on C_{max} = 94.5%, were as high as 15 (q = 1) and 11.5 (q = 2), with an
378 estimated richness of 23 profiles. In comparison, in the province of Alberta (with 661,848 km² and almost
379 twice as big as Germany), only 15 EmsB profiles were found, with genetic diversity indices between 2.3 to
380 5.1 (q = 1) and 2.1 to 4.1 (q = 2) in the five subareas, indicating low genetic diversity, which is inconsistent
381 with the hypothesis of an undetected historical endemic area of the European-like clade. Moreover, the
382 genetic indices recalculated for Switzerland were more than three times higher than the ones calculated for
383 all of Alberta.

384 In this Canadian province, the highest genetic richness was found in the central east region, and
385 the highest genetic diversity based on the abundance of each EmsB profile was found in the southwestern
386 region. However, the degree of differentiation between *E. multilocularis* populations from all five sub-
387 regions was low, with widespread distribution of two profiles (P13 and P15) representing 77 % of the
388 samples from this province. Moreover, these two profiles might be considered as part of one population,
389 due to: i) the low genetic distance between them (0.14), ii) their overlap in geographical distribution, and
390 iii) the relatively high frequency of co-occurrence in the same host (19 hosts). These results were consistent
391 with previous studies using mtDNA showing a higher prevalence and wider geographic distribution of the
392 haplotype ECA, the causative agent of the most recent human cases of AE in Alberta [25, 26]. Therefore,
393 the presence of a single predominant genetic variant in Alberta supported the hypothesis of a single
394 invasion event responsible for the initial establishment of a small number of *E. multilocularis* individuals in
395 that province.

396 The subsequent spread across the province was likely aided by red foxes, but mostly by the most
397 abundant host, the coyote, which has a larger home range (> 100 km^2) and higher dispersal distances (up to
398 300 km) [59-62]. In a previous study comparing coyotes and red foxes from Alberta infected with the ECA
399 haplotype, the difference in intensity of infection (worm burden) between both hosts was significantly
400 higher in coyotes than foxes [26], which could be related to a lack of co-evolved resistance in coyotes to the
401 European strains. Therefore, the coyote, being a naïve host, might be the primary source of environmental
402 contamination with eggs of these strains. In this study, we did not perform an analysis of intra-host
403 diversity to compare red foxes and coyotes, due to the low sample size of red foxes. However, we observed
404 a higher prevalence of the profile P13 (identified as the ECA haplotype based on mtDNA) in both coyotes
405 and red foxes, and a low prevalence of North American profiles, which is consistent with previous studies
406 using deep amplicon sequencing [26]. Further analyses of differences in the prevalence and intensity of
407 infection of European and North American genetic variants in coyotes compared to red foxes, will be
408 important to understand the role of the two hosts in the transmission and spread of the European strains and
409 the intra-host competition between North American and European strains.

410 To estimate the expansion rate of the parasite, Takumi and colleagues [63] used a mathematical
411 model predicting a spreading rate of 2.7 km per year, from an endemic to a non-endemic area in the
412 Netherlands. In contrast, the presence of highly vagile hosts, like coyotes in Western Canada, could have
413 facilitated and accelerated the spread of the European strain. Thus, it is likely that the process of colonizing
414 the entire province of Alberta was completed in a shorter time frame, expanding from the south to the
415 north, as evidenced by the highest differentiation between the Southern and the Northern regions of
416 Alberta, and the distribution of the least common profiles in distant regions. This pattern, with one single
417 cluster of profiles being the most prevalent, and low genetic diversity, was similar to that in European
418 regions initially not considered endemic. For example, in a survey conducted in Poland, only one profile
419 was the most prevalent and was distributed across the country [33]. Similarly, in Denmark and Sweden, one
420 profile represented 68.4 and 55.5% of worms, respectively [44].

421 In a previous study using EmsB to assess the genetic diversity of the parasite at a continental scale
422 in Europe [50], only 5% of the genetic variability was explained by geographical distance. In our study,
423 there was no significant correlation when analyzing samples from Alberta alone. However, there was a
424 significant correlation when including samples from British Columbia and Saskatchewan, with more
425 dissimilar samples that were far apart (> 600 km). In the distance-based redundancy analysis, 21.1% of the
426 genetic variability was explained by geographic distance, with longitude being the only significant spatial
427 variable. Our results could indicate independent invasion events in each province and isolation by distance,
428 with the exchange of some genetic variants between neighboring regions, thanks to the dispersal movement

429 of definitive hosts. Furthermore, these invasion events were likely influenced by geographical barriers, e.g.,
430 the Rocky Mountains between Alberta and British Columbia.

431 Despite the smaller sample size in Saskatchewan and British Columbia, our results were consistent
432 with previous studies using mtDNA characterization in these areas. In British Columbia, Gesy, *et al*., 2013
433 [18] reported only one European-like haplotype (BC1) in the immediate area of Quesnel, BC. Likewise, in
434 a study in Saskatchewan, only one European-like haplotype (SK1) was detected in the central area of the
435 province. Yet, seven haplotypes that belonged to the North American clade were detected in the southern
436 region [24]. However, in our study in Alberta, the abundance of the North American strain was surprisingly
437 low, compared to the European strains. Nonetheless, the genetic variability of the North American strain
438 (i.e., number of disparate profiles in a few hosts) was high (three EmsB profiles in eight worms from two
439 hosts), which is consistent with what would be expected from an ancestral endemic strain. The difference in
440 the prevalence, genetic diversity, and distribution between the North American and the European strains
441 observed in Alberta, using nuclear and mitochondrial markers was relevant, suggesting some degree of
442 competitive interactions, with a population expansion of the European over the North American strain [26].

443 This was the first large-scale study using both nuclear and mitochondrial markers to assess the
444 genetic diversity of *E. multilocularis* in North America. The use of EmsB improved characterization of
445 parasite diversity, even at a fine geographical scale, whereas the use of mtDNA aided in unravelling the
446 evolutionary process of the invasion of the European strain. As evidenced in this and previous studies, the
447 implementation of genetic population structure analysis is a powerful tool to trace the origins and history
448 expansion of parasites and their hosts in the areas invaded and for the assessment of public health risks. For
449 example, a genetic study of raccoons (*Procyon lotor*) in the Netherlands (where they were previously
450 absent), and their gastrointestinal nematode parasite (*Baylisascaris procyonis*), showed that most of the
451 Dutch raccoons and their roundworms were introduced through ex-captive individuals, which ultimately
452 aided in the development of control measures of these invasive populations [64]. Recently, a genomic
453 analysis allowed the evaluation of the process of colonization of the Americas by *Schistosoma mansoni*, a
454 blood fluke that infects humans, and that was introduced into the Americas from Africa during the Trans-
455 Atlantic slave trade [65]. In this study, no evidence of population bottlenecks was observed, suggesting that
456 *S. mansoni* parasites were pre-adapted to the Americas and able to establish with relative ease.

457 Further work on the study of genetic diversity of *E. multilocularis* including wild and domestic
458 hosts from British Columbia and Saskatchewan, as well as from eastern Canadian provinces, where the
459 parasite has been recently detected, and in central US states, will be pivotal to understanding the current
460 distribution and expansion trend of the European strains in continental North America, including the role of
461 the main definitive hosts (coyotes and foxes) in harboring and spreading these strains.
462

463 **Funding**
464
465 MITACS (Mathematics of Information Technology and Complex Systems), Accelerate Program, Grant no.
466 10018836 (to A.M.); ELANCO Canada Limited, Grant no. 10017067 (to A.M.); ACA (Alberta
467 Conservation Association), Grant no. 10010809 (to A.M.); Natural Sciences & Engineering Research
468 Council of Canada (NSERC), Discovery Grant, Grant no.10013904, URGC SEED Grant; Grant
469 no.10007230 (to A.M.); NSERC, Discovery Grant, Grant no. RGPIN-2018-03913 (to K.E.R.); City of
470 Calgary, Grant no.10007719 (to A.M.).
471
472 **Acknowledgments**
473

474 We gratefully acknowledge the ATA (Alberta Trappers' Association), the trappers who provided animal
475 carcasses, Vanessa Bastid (Nancy Laboratory for Rabies and Wildlife) who helped with the molecular
476 analysis, Dimitri Giunchi (University of Pisa) who helped with the Bayesian analysis, Maria Cavedon and
477 Jessica Theoret (University of Calgary) who helped with spatial analysis, John Patrick Kastelic (University
478 of Calgary) for revising the final draft of the paper, and all undergraduate students that contributed to
479 laboratory work.
480
481
482 **References**

483 1. Sax D.F., Gaines S.D. 2003 Species diversity: from global decreases to local increases. *Trends*
484 *Ecol. Evol.* **18**(11), 561-566.
485 2. Young H.S., Parker I.M., Gilbert G.S., Sofia Guerra A., Nunn C.L. 2017 Introduced Species,
486 Disease Ecology, and Biodiversity–Disease Relationships. *Trends Ecol. Evol.* **32**(1), 41-54.
487 (doi:10.1016/j.tree.2016.09.008).
488 3. Hatcher M.J., Dick J.T.A., Dunn A.M. 2012 Disease emergence and invasions. *Funct. Ecol.* **26**(6),
489 1275-1287. (doi:doi:10.1111/j.1365-2435.2012.02031.x).
490 4. Ogden N.H., Wilson J.R.U., Richardson D.M., Hui C., Davies S.J., Kumschick S., Le Roux J.J.,
491 Measey J., Saul W.-C., Pulliam J.R.C. 2019 Emerging infectious diseases and biological invasions: a call
492 for a One Health collaboration in science and management. *Royal Soc. Open Sci.* **6**(3), 181577-181577.
493 (doi:10.1098/rsos.181577).
494 5. Thompson R.C.A., Deplazes P., Lymbery A.J. 2017 *Advances in Parasitology*: *Echinococcus and*
495 *Echinococcosis, Part A*. 1st ed., Academic Press; 525 p.
496 6. Eckert J., Gemmell M.A., Meslin F.X., Pawlowski Z.S., WHO. 2001 WHO/OIE manual on
497 echinococcosis in humans and animals: a public health problem of global concern. Paris, World
498 Organization for Animal Health; 265 p.
499 7. Thompson R.C.A. 2017 Biology and Systematics of Echinococcus. *Adv. Parasitol.* **95**, 65-109.
500 (doi:10.1016/bs.apar.2016.07.001).
501 8. Torgerson P.R., Robertson L.J., Enemark H.L., Foehr J., Van Der Giessen J.W.B., Kapel C.M.O.,
502 Klun I., Trevisan C. 2020 Source attribution of human echinococcosis: A systematic review and meta-
503 analysis. *PLoS Negl. Trop. Dis* **14**(6), e0008382. (doi:10.1371/journal.pntd.0008382).
504 9. Davidson R.K., Romig T., Jenkins E., Tryland M., Robertson L.J. 2012 The impact of
505 globalisation on the distribution of *Echinococcus multilocularis*. *Trends Parasitol.* **28**(6), 239-247.
506 (doi:10.1016/j.pt.2012.03.004).
507 10. Eckert J., Conraths F.J., Tackmann K. 2000 Echinococcosis: an emerging or re-emerging
508 zoonosis? *Int. J. Parasitol.* **30**(12–13), 1283-1294. (doi:10.1016/S0020-7519(00)00130-2).
509 11. Giraudoux P., Craig P.S., Delattre P., Bao G., Bartholomot B., Harraga S., Quere J.P., Raoul F.,
510 Wang Y., Shi D., et al. 2003 Interactions between landscape changes and host communities can regulate
511 *Echinococcus multilocularis* transmission. *Parasitology* **127 Suppl**, S121-131.
512 12. Jenkins E.J., Peregrine A.S., Hill J.E., Somers C., Gesy K., Barnes B., Gottstein B., Polley L. 2012
513 Detection of European Strain of *Echinococcus multilocularis* in North America. *Emerg. Infect. Dis.* **18**(6),
514 1010-1012. (doi:10.3201/eid1806.111420).
515 13. Gottstein B., Stojkovic M., Vuitton D.A., Millon L., Marcinkute A., Deplazes P. 2015 Threat of
516 alveolar echinococcosis to public health: A challenge for Europe. *Trends Parasitol.* **31**(9), 407-412.
517 (doi:10.1016/j.pt.2015.06.001).
518 14. Massolo A., Valli D., Wassermann M., Cavallero S., D'Amelio S., Alberto M., Torretta E.,
519 Serafini M., Casulli A., Zambon L., et al. 2018 Unexpected *Echinococcus multilocularis* infections in
520 shepherd dogs and wolves in south-western Italian Alps: A new endemic area? *Int. J. Parasitol. Parasites*
521 *Wildl.* **7**(3), 309-316. (doi:10.1016/j.ijppaw.2018.08.001).
522 15. Kotwa J., Isaksson M., Jardine C., Douglas Campbell G., Berke O., L. Pearl D., J. Mercer N.,
523 Osterman-Lind E., Peregrine A. 2019 *Echinococcus multilocularis* Infection, Southern Ontario, Canada.
524 *Emerg. Infect. Dis.* **25**(2), 265-272. (doi:10.3201/eid2502.180299).

525 16. Deplazes P., Hagglin D., Gloor S., Romig T. 2004 Wilderness in the city: the urbanization of
526 *Echinococcus multilocularis*. *Trends Parasitol.* **20**(2), 77-84.
527 17. Liccioli S., Giraudoux P., Deplazes P., Massolo A. 2015 Wilderness in the 'city' revisited:
528 different urbes shape transmission of *Echinococcus multilocularis* by altering predator and prey
529 communities. *Trends Parasitol.* **31**(7), 297-305. (doi:10.1016/j.pt.2015.04.007).
530 18. Gesy K., Hill J.E., Schwantje H., Liccioli S. 2013 Establishment of a European-type strain of
531 *Echinococcus multilocularis* in Canadian wildlife. *Parasitology* **140**(9), 1133-1137.
532 (doi:10.1017/S0031182013000607).
533 19. Henttonen H., Fuglei E., Gower C.N., Haukisalmi V., Ims R.A., Niemimaa J., Yoccoz N.G. 2001
534 *Echinococcus multilocularis* on Svalbard: introduction of an intermediate host has enabled the local life-
535 cycle. *Parasitology* **123**. (doi:10.1017/s0031182001008800).
536 20. Nakao M., Xiao N., Okamoto M., Yanagida T., Sako Y., Ito A. 2009 Geographic pattern of
537 genetic variation in the fox tapeworm *Echinococcus multilocularis*. *Parasitol. Int.* **58**(4), 384-389.
538 21. Konyaev S.V., Yanagida T., Nakao M., Ingovatova G.M., Shoykhet Y.N., Bondarev A.Y.,
539 Odnokurtsev V.A., Loskutova K.S., Lukmanova G.I., Dokuchaev N.E., et al. 2013 Genetic diversity of
540 Echinococcus spp. in Russia. *Parasitology* **140**(13), 1637-1647. (doi:10.1017/S0031182013001340).
541 22. Umhang G., Knapp J., Wassermann M., Bastid V., Peytavin de Garam C., Boué F., Cencek T.,
542 Romig T., Karamon J. 2021 Asian Admixture in European Echinococcus multilocularis Populations: New
543 Data From Poland Comparing EmsB Microsatellite Analyses and Mitochondrial Sequencing. *Front. Vet.*
544 *Sci.* **7**(1202). (doi:10.3389/fvets.2020.620722).
545 23. Karamon J., Stojecki K., Samorek-Pierog M., Bilska-Zajac E., Rozycki M., Chmurzynska E.,
546 Sroka J., Zdybel J., Cencek T. 2017 Genetic diversity of Echinococcus multilocularis in red foxes in
547 Poland: the first report of a haplotype of probable Asian origin. *Folia Parasitol. (Praha)* **64**.
548 (doi:10.14411/fp.2017.007).
549 24. Gesy K.M., Jenkins E.J. 2015 Introduced and Native Haplotypes of *Echinococcus multilocularis*
550 in Wildlife in Saskatchewan, Canada. *J. Wildl. Dis.* **51**(3), 743-748. (doi:10.7589/2014-08-214 [doi]).
551 25. Massolo A., Klein C., Kowalewska-Grochowska K., Belga S., MacDonald C., Vaughan S., Girgis
552 S., Giunchi D., Bramer S.A., Santa M.A., et al. 2019 European *Echinococcus multilocularis* Identified in
553 Patients in Canada. *N. Engl. J. Med.* **381**(4), 384-385. (doi:10.1056/NEJMc1814975).
554 26. Santa M.A., Rezansoff A.M., Chen R., Gilleard J.S., Musiani M., Ruckstuhl K.E., Massolo A.
555 2021 Deep amplicon sequencing highlights low intra-host genetic variability of Echinococcus
556 multilocularis and high prevalence of the European-type haplotypes in coyotes and red foxes in Alberta,
557 Canada. *PLoS Negl. Trop. Dis.* **15**(5), e0009428. (doi:10.1371/journal.pntd.0009428).
558 27. Houston S., Belga S., Buttenschoen K., Cooper R., Girgis S., Gottstein B., Low G., Massolo A.,
559 MacDonald C., Müller N., et al. 2021 Epidemiological and Clinical Characteristics of Alveolar
560 Echinococcosis: An Emerging Infectious Disease in Alberta, Canada. *Am. J. Trop. Med. Hyg.*
561 (doi:10.4269/ajtmh.20-1577).
562 28. Schurer J.M., Tsybina P., Gesy K.M., Kolapo T.U., Skinner S., Hill J.E., Jenkins E.J. 2020
563 Molecular Evidence for Local Acquisition of Human Alveolar Echinococcosis in Saskatchewan, Canada.
564 *J.Infect.Dis*. (doi:10.1093/infdis/jiaa473).
565 29. Polish L.B., O'Connell E.M., Barth T.F.E., Gottstein B., Zajac A., Gibson P.C., Bah A.,
566 Kirchgessner M., Estrada M., Seguin M.A., et al. 2022 European Haplotype of *Echinococcus multilocularis*
567 in the United States. *N. Engl. J. Med.* **387**(20), 1902-1904. (doi:10.1056/nejmc2210000).
568 30. Polish L.B., Pritt B., Barth T.F.E., Gottstein B., O'Connell E.M., Gibson P.C. 2020 First European
569 Haplotype of Echinococcus multilocularis Identified in the United States: An Emerging Disease? *Clin.*
570 *Infect. Dis.* (doi:10.1093/cid/ciaa245).
571 31. Umhang G., Knapp J., Hormaz V., Raoul F., Boué F. 2014 Using the genetics of *Echinococcus*
572 *multilocularis* to trace the history of expansion from an endemic area. *Infect. Genet. Evol.* **22**, 142-149.
573 (doi:10.1016/j.meegid.2014.01.018).
574 32. Knapp J., Bart J.M., Maillard S., Gottstein B., Piarroux R. 2010 The genomic *Echinococcus*
575 *microsatellite* EmsB sequences: from a molecular marker to the epidemiological tool. *Parasitology*
576 **137**(Special Issue 03), 439-449. (doi:10.1017/S0031182009991612).
577 33. Umhang G., Karamon J., Hormaz V., Knapp J., Cencek T., Boué F. 2017 A step forward in the
578 understanding of the presence and expansion of *Echinococcus multilocularis* in Eastern Europe using

579 microsatellite EmsB genotyping in Poland. *Infect. Genet. Evol.* **54**, 176-182.
580 (doi:10.1016/j.meegid.2017.07.004).
581 34. Knapp J., Guislain M.H., Bart J.M., Raoul F., Gottstein B., Giraudoux P., Piarroux R. 2008
582 Genetic diversity of *Echinococcus multilocularis* on a local scale. *Infect. Genet. Evol.* **8**(3), 367-373.
583 (doi:10.1016/j.meegid.2008.02.010).
584 35. Santa M.A., Pastran S.A., Klein C., Duignan P., Ruckstuhl K., Romig T., Massolo A. 2018
585 Detecting co-infections of Echinococcus multilocularis and Echinococcus canadensis in coyotes and red
586 foxes in Alberta, Canada using real-time PCR. *Int. J. Parasitol. Parasites Wildl.* **7**(2), 111-115.
587 (doi:https://doi.org/10.1016/j.ijppaw.2018.03.001).
588 36. Gesy K., Pawlik M., Kapronczai L., Wagner B., Elkin B., Schwantje H., Jenkins E. 2013 An
589 improved method for the extraction and quantification of adult Echinococcus from wildlife definitive hosts.
590 *Parasitol. Res.* **112**(5), 2075-2078. (doi:10.1007/s00436-013-3371-x).
591 37. Paradis E. 2010 pegas: an R package for population genetics with an integrated-modular approach.
592 *Bioinformatics* **26**(3), 419-420. (doi:10.1093/bioinformatics/btp696).
593 38. Ronquist F., Teslenko M., van der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L.,
594 Suchard M.A., Huelsenbeck J.P. 2012 MrBayes 3.2: efficient Bayesian phylogenetic inference and model
595 choice across a large model space. *Syst. Biol.* **61**(3), 539-542. (doi:10.1093/sysbio/sys029).
596 39. Schliep K.P. 2011 phangorn: phylogenetic analysis in R. *Bioinformatics* **27**(4), 592-593.
597 40. Knapp J., Damy S., Brillaud J., Tissot J.-D., Navion J., Mélior R., Afonso E., Hormaz V.,
598 Gottstein B., Umhang G., et al. 2017 EWET: Data collection and interface for the genetic analysis of
599 Echinococcus multilocularis based on EmsB microsatellite. *PLoS One* **12**(10), e0183849.
600 (doi:10.1371/journal.pone.0183849).
601 41. Legendre P., Legendre L. 2012 *Numerical Ecology*. Oxford, Elsevier; 990 p.
602 42. Shimodaira H. 2004 Approximately Unbiased Tests of Regions Using Multistep-Multiscale
603 Bootstrap Resampling. *Ann. Stat.* **32**(6), 2616-2641.
604 43. Knapp J., Bart J.M., Glowatzki M.L., Ito A., Gerard S., Maillard S., Piarroux R., Gottstein B. 2007
605 Assessment of Use of Microsatellite Polymorphism Analysis for Improving Spatial Distribution Tracking
606 of *Echinococcus multilocularis*. *J. Clin. Microbiol.* **45**(9), 2943-2950. (doi:10.1128/JCM.02107-06).
607 44. Knapp J., Umhang G., Wahlström H., Al-Sabi M.N.S., Ågren E.O., Enemark H.L. 2019 Genetic
608 diversity of Echinococcus multilocularis in red foxes from two Scandinavian countries: Denmark and
609 Sweden. *Food Waterborne Parasitol.* **14**, e00045. (doi:https://doi.org/10.1016/j.fawpar.2019.e00045).
610 45. Kruskal J.B. 1964 Nonmetric multidimensional scaling: A numerical method. *Psychometrika*
611 **29**(2), 115-129. (doi:10.1007/bf02289694).
612 46. Suzuki R., Shimodaira H. 2006 Pvclust: an R package for assessing the uncertainty in hierarchical
613 clustering. *Bioinformatics* **22**(12), 1540-1542. (doi:10.1093/bioinformatics/btl117).
614 47. Chao A., Kubota Y., Zelený D., Chiu C.-H., Li C.-F., Kusumoto B., Yasuhara M., Thorn S., Wei
615 C.-L., Costello M.J., et al. 2020 Quantifying sample completeness and comparing diversities among
616 assemblages. *Ecol. Res.* **35**(2), 292-314. (doi:https://doi.org/10.1111/1440-1703.12102).
617 48. Chao A., Gotelli N.J., Hsieh T.C., Sander E.L., Ma K.H., Colwell R.K., Ellison A.M. 2014
618 Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species
619 diversity studies. *Ecol. Monogr.* **84**(1), 45-67. (doi:https://doi.org/10.1890/13-0133.1).
620 49. Chao A., Jost L. 2012 Coverage-based rarefaction and extrapolation: standardizing samples by
621 completeness rather than size. *Ecology* **93**(12), 2533-2547. (doi:https://doi.org/10.1890/11-1952.1).
622 50. Knapp J., Bart J.-M., Giraudoux P., Glowatzki M.-L., Breyer I., Raoul F., Deplazes P., Duscher
623 G., Martinek K., Dubinsky P., et al. 2009 Genetic Diversity of the Cestode Echinococcus multilocularis in
624 Red Foxes at a Continental Scale in Europe. *PLoS Negl. Trop. Dis.* **3**(6), e452.
625 (doi:10.1371/journal.pntd.0000452).
626 51. Oksanen J., Blanchet F.G., Kindt R., Legendre P., Minchin P., O'hara R., Simpson G., Solymos
627 P., Stevens M., Wagner H. 2013 Community ecology package. *R package version*, 2.0-2.
628 52. Mantel N. 1967 The Detection of Disease Clustering and a Generalized Regression Approach.
629 *Cancer Res.* **27**(2), 209-220.
630 53. Legendre P., Anderson M.J. 1999 Distance-based Redundancy Analysis: testing multispecies
631 responses in multifactorial ecological experiments. *Ecol. Monogr.* **69**(1), 1-24. (doi:10.1890/0012-
632 9615(1999)069[0001:DBRATM]2.0.CO;2).

633 54. Kamler J.F., Ballard W.B. 2002 A Review of Native and Nonnative Red Foxes in North America.
634 *Wildl. Soc. Bull. (1973-2006)* **30**(2), 370-379.
635 55. Lockwood J.L., Cassey P., Blackburn T. 2005 The role of propagule pressure in explaining species
636 invasions. *Trends Ecol. Evol.* **20**(5), 223-228. (doi:https://doi.org/10.1016/j.tree.2005.02.004).
637 56. Dlugosch K.M., Anderson S.R., Braasch J., Cang F.A., Gillette H.D. 2015 The devil is in the
638 details: genetic variation in introduced populations and its contributions to invasion. *Mol. Ecol.* **24**(9),
639 2095-2111. (doi:10.1111/mec.13183).
640 57. Davidson R.K., Lavikainen A., Konyaev S., Schurer J., Miller A.L., Oksanen A., Skírnisson K.,
641 Jenkins E. 2016 Echinococcus across the north: Current knowledge, future challenges. *Food Waterborne*
642 *Parasitol.* **4**, 39-53. (doi:10.1016/j.fawpar.2016.08.001).
643 58. Umhang G., Bastid V., Avcioglu H., Bagrade G., Bujanić M., Bjelić Čabrilo O., Casulli A., Dorny
644 P., van der Giessen J., Guven E., et al. 2021 Unravelling the genetic diversity and relatedness of
645 E*chinococcus multilocularis* isolates in Eurasia using the EmsB microsatellite nuclear marker. *Infect.,*
646 *Genet. Evol.* **92**, 104863. (doi:10.1016/j.meegid.2021.104863).
647 59. Gosselink T.E., Van Deelen T.R., Warner R.E., Joselyn M.G. 2003 Temporal Habitat Partitioning
648 and Spatial Use of Coyotes and Red Foxes in East-Central Illinois. *J. Wildl. Manage.* **67** (1), 90-103.
649 (doi:10.2307/3803065).
650 60. Dlugosch K.M., Parker I.M. 2008 Founding events in species invasions: genetic variation,
651 adaptive evolution, and the role of multiple introductions. *Mol. Ecol.* **17**(1), 431-449.
652 (doi:doi:10.1111/j.1365-294X.2007.03538.x).
653 61. Gese E.M., Rongstad O.J., Mytton W.R. 1988 Home Range and Habitat Use of Coyotes in
654 Southeastern Colorado. *J. Wildl. Manage.* **52**(4), 640-646. (doi:10.2307/3800923).
655 62. Harrison D.J. 1992 Dispersal Characteristics of Juvenile Coyotes in Maine. *J. Wildl. Manage.*
656 **56**(1), 128-138. (doi:10.2307/3808800).
657 63. Takumi K., de Vries A., Chu M.L., Mulder J., Tenuis P., van der Giessen J. 2008 Evidence for an
658 increasing presence of *Echinococcus multilocularis* in foxes in the Netherlands. *Int. J. Parasitol.* **38**.
659 (doi:10.1016/j.ijpara.2007.09.014).
660 64. Maas M., Tatem-Dokter R., Rijks J.M., Dam-Deisz C., Franssen F., van Bolhuis H., Heddergott
661 M., Schleimer A., Schockert V., Lambinet C., et al. 2022 Population genetics, invasion pathways and
662 public health risks of the raccoon and its roundworm Baylisascaris procyonis in northwestern Europe.
663 *Transbound. Emerg. Dis.* **69**(4), 2191-2200. (doi:https://doi.org/10.1111/tbed.14218).
664 65. Platt R.N., 2nd, Le Clec'h W., Chevalier F.D., McDew-White M., LoVerde P.T., Ramiro de Assis
665 R., Oliveira G., Kinung'hi S., Djirmay A.G., Steinauer M.L., et al. 2022 Genomic analysis of a parasite
666 invasion: Colonization of the Americas by the blood fluke Schistosoma mansoni. *Mol. Ecol.* **31**(8), 2242-
667 2263. (doi:10.1111/mec.16395).

668

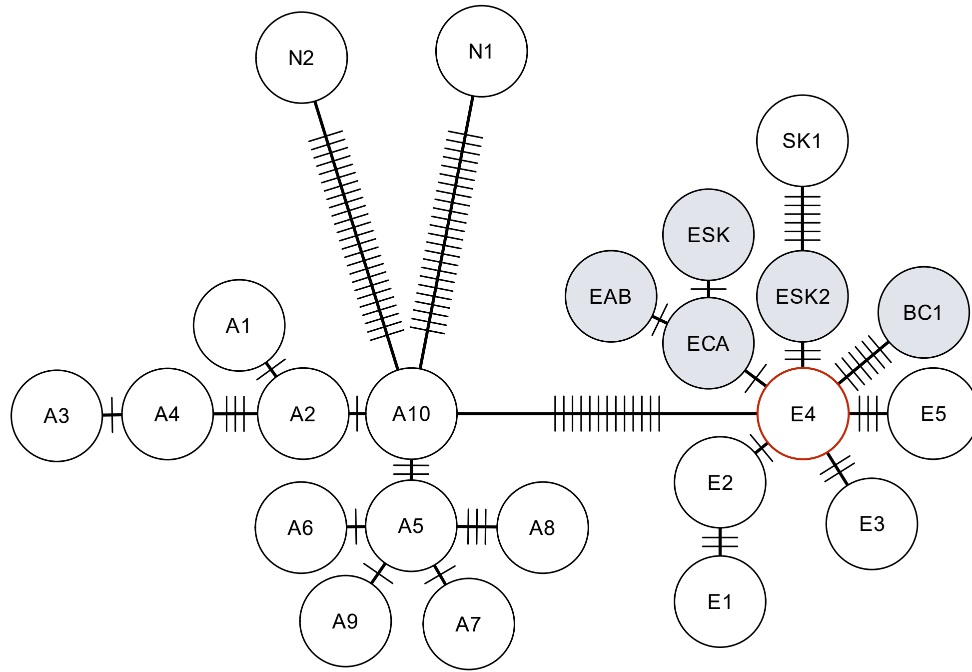
669

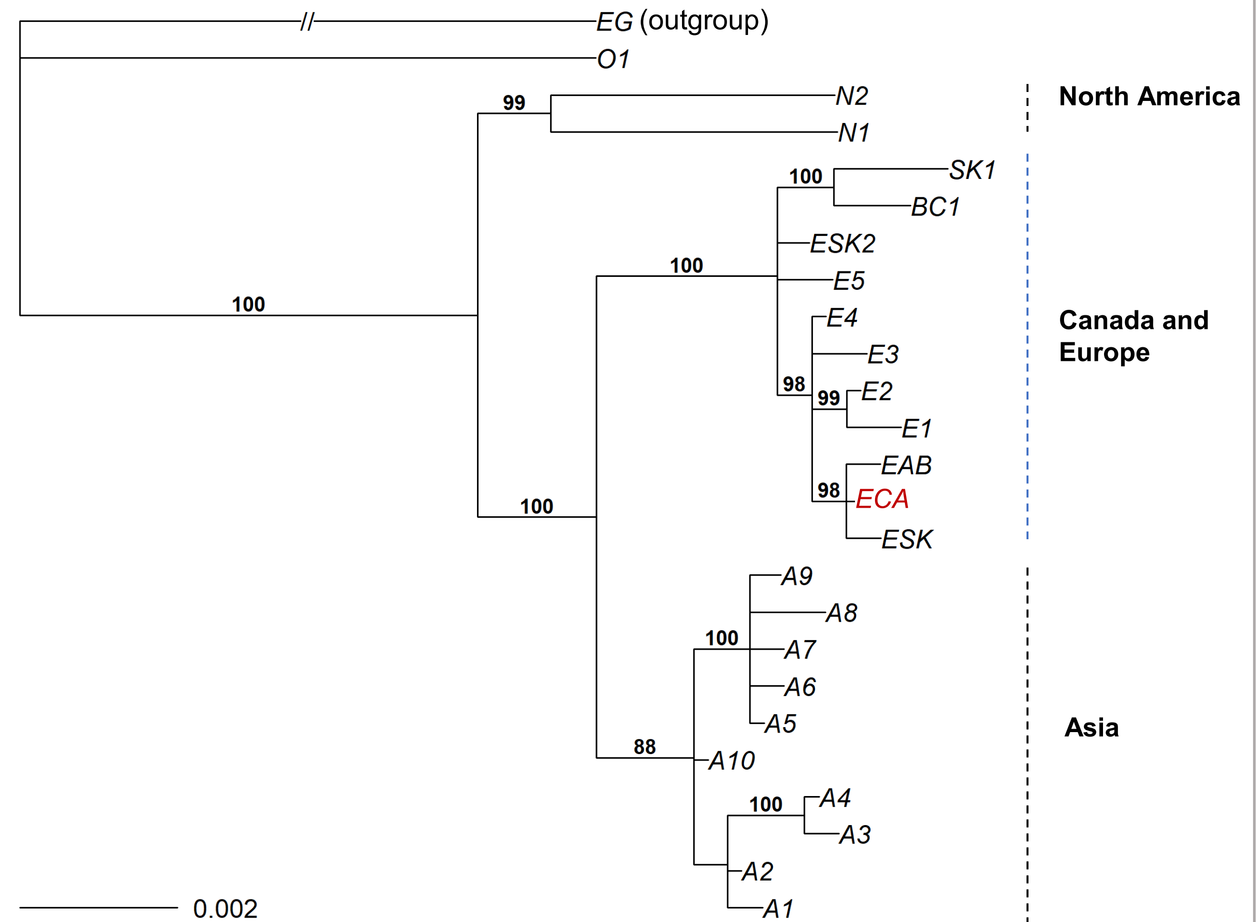
670

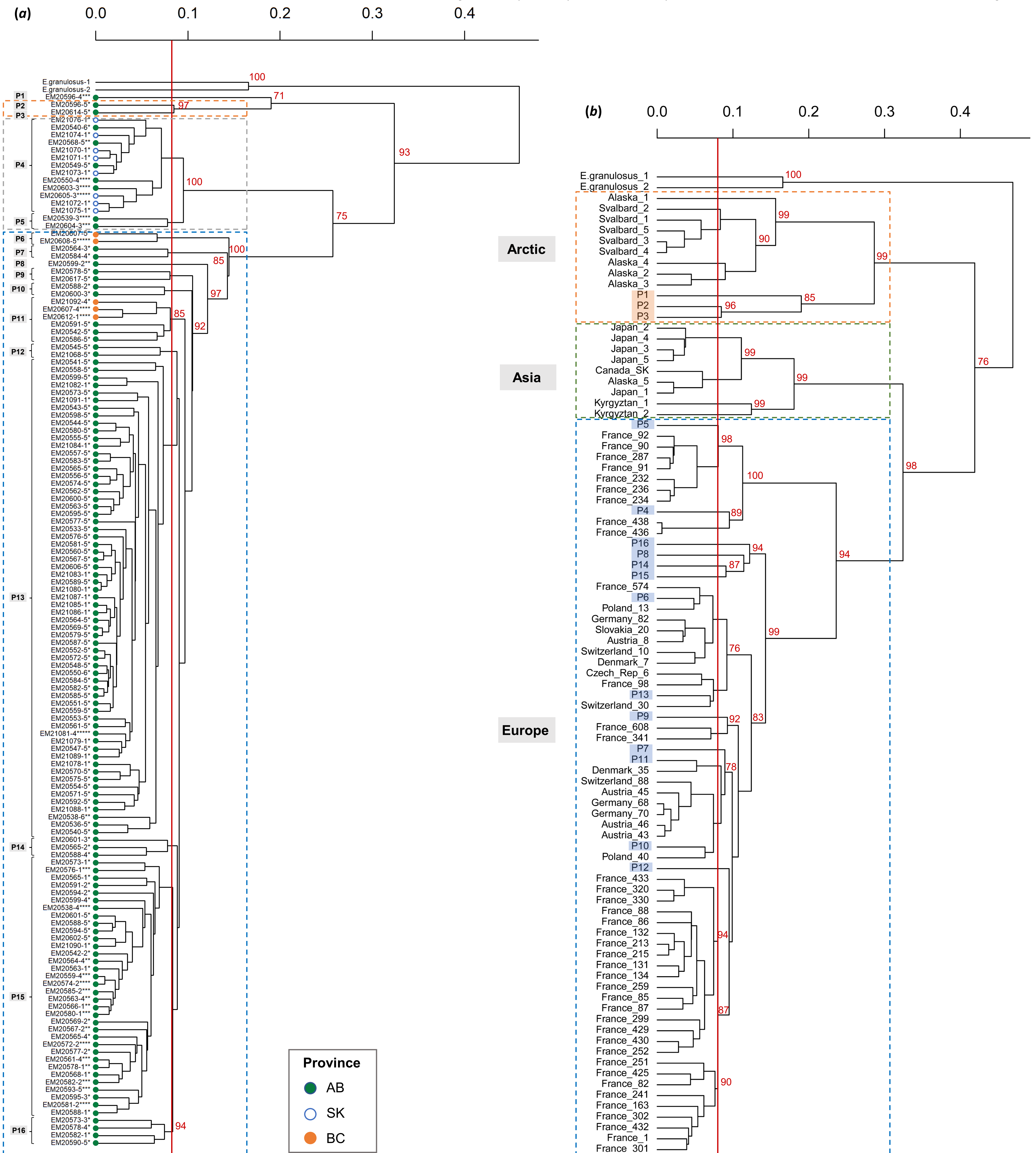
671 **Figures Captions**
672
673 **Figure 1. Genetic relationships among haplotypes of *Echinococcus multilocularis* from Western**
674 **Canada and other historically endemic regions. (*a*)** Mitochondrial haplotype network based on
675 Hamming distance, including European, Asian, and North American haplotypes previously reported (white
676 circles) [20, 24, 25], compared to haplotypes from Western Canada detected in our study (grey circles). The
677 network was constructed based on concatenated sequences of mitochondrial genes *nad2*, *cob* and *cox1*. The
678 ECA haplotype (in red) is the most prevalent in Alberta and has been associated with the most recent cases
679 of human AE in this province. **(*b*)** Phylogenetic tree inferred by partitioned Bayesian analysis performed on
680 concatenated mitochondrial DNA, using *E. granulosus* (G1) as an outgroup. Values on tree nodes are
681 Bayesian posterior probabilities. The 0.002 scale bar denotes genetic distance (nucleotide substitutions per
682 site).
683
684 **Figure 2. Cluster analysis of EmsB profiles of *Echinococcus multilocularis* from Western Canada and**
685 **correspondence with the most similar profiles globally. (*a*)** Dendrogram constructed by hierarchical
686 clustering analysis based on EmsB genotypic data from *E. multilocularis* worms collected in Alberta (AB),
687 Saskatchewan (SK), and British Columbia (BC) from 2012 to 2017. Two samples of *E. granulosus sensu*
688 *stricto* (G1) were used as an outgroup. Approximately unbiased *p*-values (red numbers on nodes, in
689 percent) were calculated with multiscale bootstrap resampling (1000 bootstrap replicates). A genetic
690 distance threshold of 0.08 (red line) was used to identify unique profiles. Worms for the same host with
691 indistinguishable profiles were pooled and the number of asterisks represents the number of worms for each
692 profile. In total, 16 unique profiles were identified and grouped in three clusters based on unbiased *p*-values
693 (sig. $\alpha = 0.05$). **(*b*)** Dendrogram constructed with EmsB profiles from Europe, Asia, and North America
694 (from the EWET database), and the 16 EmsB profiles identified in our study in Alberta (orange and blue
695 for North American and European clades, respectively). In total, 1275 samples from 17 countries were
696 compared with the profiles obtained from our study; thereafter, five profiles with the lowest genetic
697 distance to each of our profiles (P1-P16) were used to build the dendrogram.
698
699 **Figure 3. Non-metric multidimensional scaling (NMDS) ordination plot of EmsB profiles of**
700 ***Echinococcus multilocularis* collected in Western Canada from 2012 to 2017.** Genetic distance was
701 calculated based on the Euclidean distance. Coloured symbols represent the location **(*a*)** and mtDNA
702 haplotype **(*b*)**, 2D stress = 0.05.
703
704 **Figure 4. *Echinococcus multilocularis* samples collected in Alberta from 2012 to 2017, and coverage-**
705 **based rarefaction and extrapolation sampling curves per subregion in Alberta, Canada. (*a*)**
706 Subregions in Alberta: North-West (AB-NW), North-East (AB-NE), Central-West (AB-CW), Central-West
707 (AB-CE), South-West (AB-SW, and South-East (AB-SE). The dots represent the geographical locations of
708 the parasite samples analyzed. **(*b*)** Non-asymptotic coverage-based rarefaction (solid lines) and
709 extrapolation (dashed lines) curves for orders $q = 0$, 1 and 2, (i.e., Chao1 richness, Shannon and inverse
710 Simpson diversity) for each subregion in Alberta. Confidence intervals (95%) are represented by shading.
711 Solid dots denote the observed data points. (AB-NW was excluded from the analysis due to the small
712 sample size).
713
714 **Figure 5. Spatial distribution of the EmsB profiles of *Echinococcus multilocularis* collected in**
715 **Western Canada from 2012 to 2017. (*a*) and (*b*)** are the distributions of P13 and P15, the two main EmsB
716 profiles in Alberta. **(*c*)** illustrates the distribution of profiles P4 and P5, present only in Alberta (AB) and
717 Saskatchewan (SK). **(*d*)** is the distribution of the profiles P6 and P11, present exclusively in BC and AB. **(*e*)**
718 shows the distribution of EmsB profiles per wildlife management unit (WMU) in Alberta. The size of the
719 pie chart represents the number of samples collected.

720 **Supplementary material**
721
722 **Figure S1.** Sample-size-based rarefaction and extrapolation sampling curves per subregion in Alberta,
723 Canada.
724
725 **Figure S2.** Genetic diversity of EmsB assemblages of *Echinococcus multilocularis* among sub-regions in
726 Alberta, Canada from 2012 to 2017.
727
728 **Figure S3.** Analysis of the spatial distribution of EmsB profiles of *Echinococcus multilocularis* collected in
729 Western Canada from 2012 to 2017.
730
731 **Table S1.** EmsB profiles in *Echinococcus multilocularis* worms isolated from coyotes and red foxes in
732 three provinces in Western Canada from 2012 to 2017.
733
734 **Table S2.** Conformity between European EmsB profiles and those in *Echinococcus multilocularis* worms
735 isolated from coyotes and red foxes from Western Canada collected from 2012 to 2017.

736 **Table S3.** Diversity analysis of EmsB profiles of *Echinococcus multilocularis* from the province of
737 Alberta, compared to some endemic regions in Europe.
738
739


(a)
N2
N1
SK1
ESK
EAB
ESK2
BC1
ECA
A1
A3
A4
A2
A10
E4
E5
E2
A6
A5
A8
E3
A9
A7
E1


(b)
EG (outgroup)
O1
99
N2
N1
North America
100
SK1
BC1
ESK2
100
E5
100
E4
E3
Canada and Europe
98
99
E2
E1
EAB
98
ECA
100
ESK
A9
A8
100
A7
A6
A5
Asia
88
A10
100
A4
A3
A2
A1
0.002


(a)
0.0
0.1
0.2
0.3
0.4
E.granulosus-1
E.granulosus-2
P1
P2
P3
P4
P5
P6
P7
P8
P9
P10
P11
P12
P13
P14
P15
P16
100
71
97
93
100
75
100
85
97
85
92
94
Province
AB
SK
BC
(b)
0.0
0.1
0.2
0.3
0.4
E.granulosus_1
E.granulosus_2
Alaska_1
Svalbard_2
Svalbard_1
Svalbard_5
Svalbard_3
Svalbard_4
Alaska_4
Alaska_2
Alaska_3
P1
P2
P3
Japan_2
Japan_4
Japan_3
Japan_5
Canada_SK
Alaska_5
Japan_1
Kyrgyztan_1
Kyrgyztan_2
P5
France_92
France_90
France_287
France_91
France_232
France_236
France_234
P4
France_438
France_436
P16
P8
P14
P15
France_574
P6
Poland_13
Germany_82
Slovakia_20
Austria_8
Switzerland_10
Denmark_7
Czech_Rep_6
France_98
P13
Switzerland_30
P9
France_608
France_341
P7
P11
Denmark_35
Switzerland_88
Austria_45
Germany_68
Germany_70
Austria_46
Austria_43
P10
Poland_40
P12
France_433
France_320
France_330
France_88
France_86
France_132
France_213
France_215
France_131
France_134
France_259
France_85
France_87
France_299
France_429
France_430
France_252
France_251
France_425
France_82
France_241
France_163
France_302
France_432
France_1
France_301
Arctic
Asia
Europe
100
99
90
99
85
96
99
99
99
76
98
100
89
94
87
94
98
99
76
92
83
78
94
87
90

(*a*)

Location
AB
SK
BC

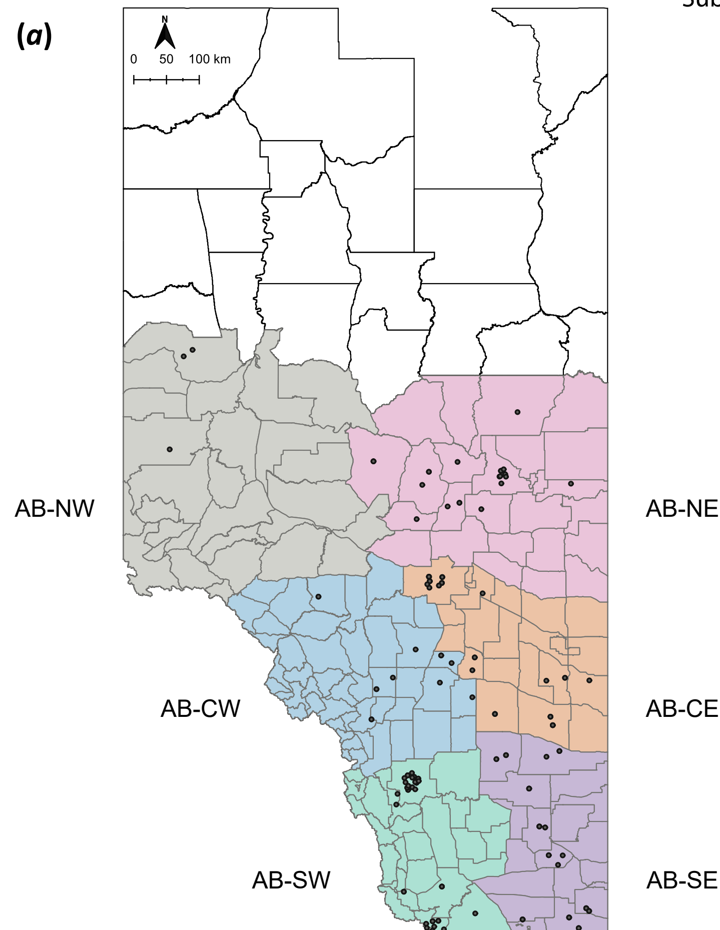
Distance
0.08
0.1
0.2
0.3

(*b*)

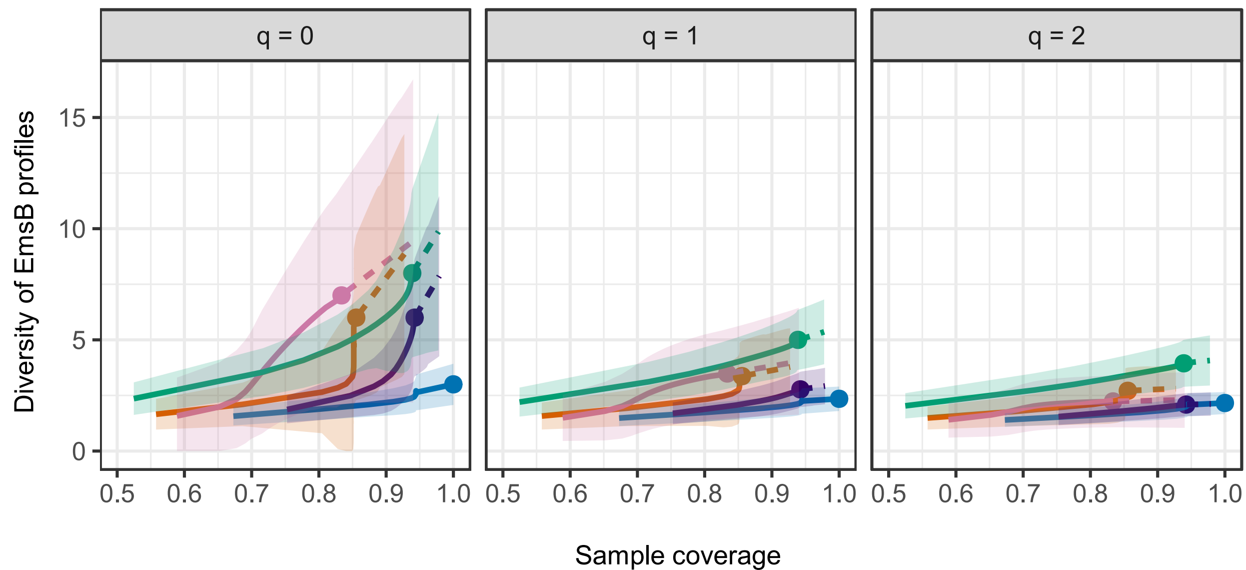
Haplotypes
ESK
ESK2
EAB
ECA
BC1
N2

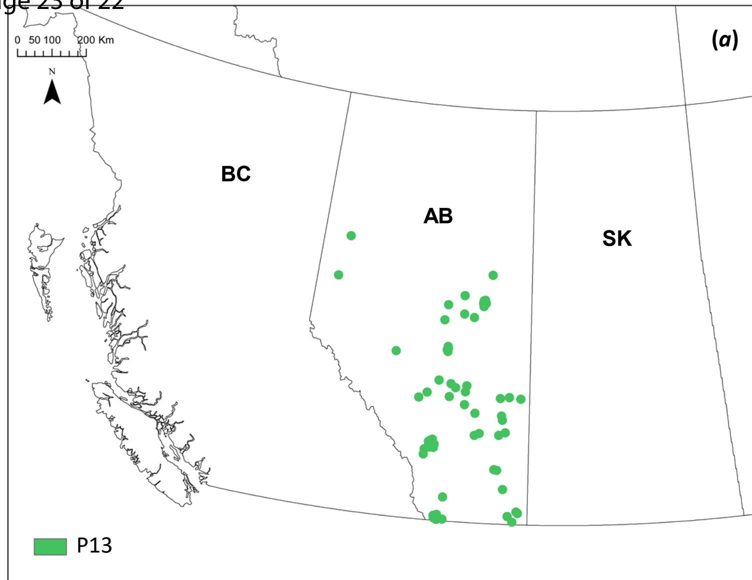
Distance
0.08
0.1
0.2
0.3

(*a*)

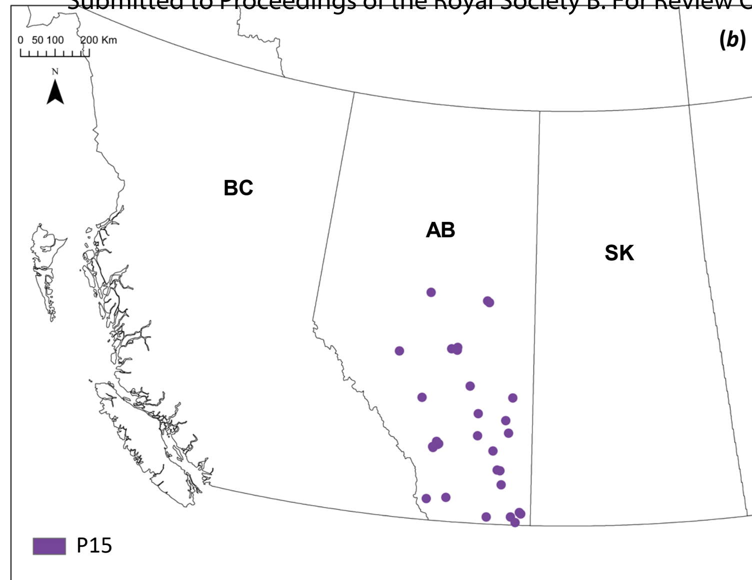


(*b*)

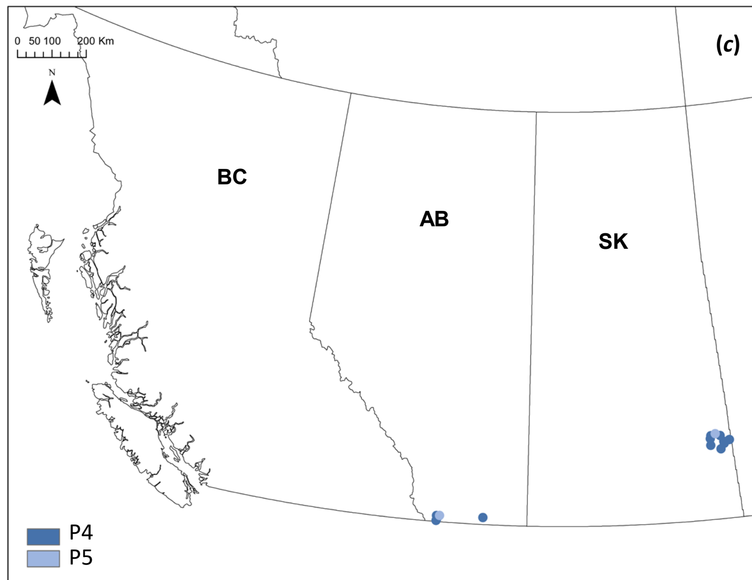




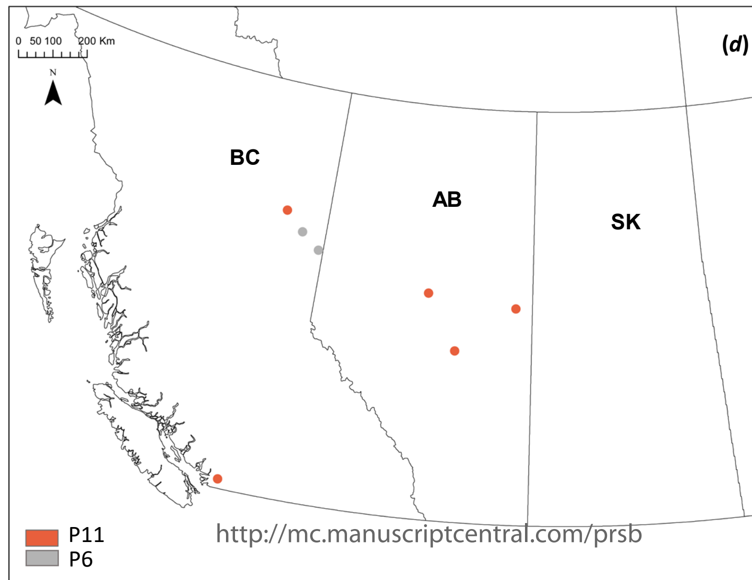
(a)
0 50 100 200 Km
N
BC
AB
SK
P13



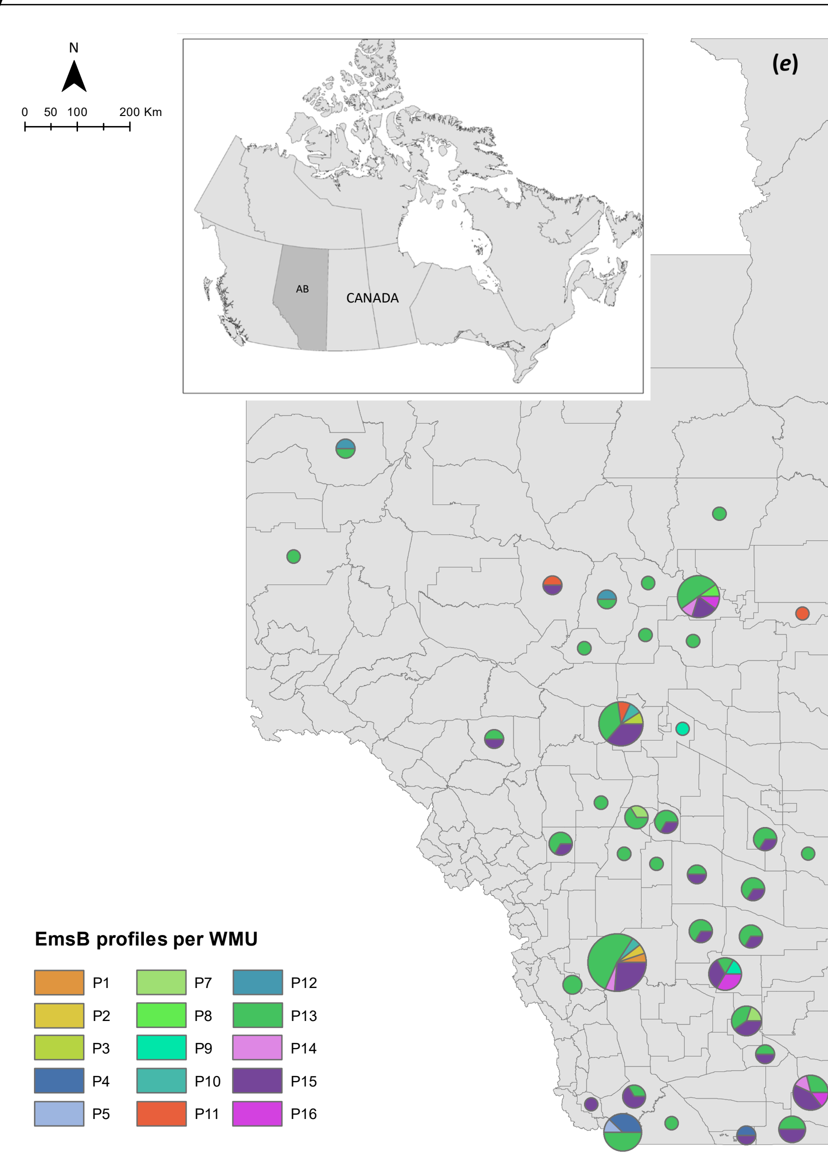
(b)
0 50 100 200 Km
N
BC
AB
SK
P15



(c)
0 50 100 200 Km
N
BC
AB
SK
P4
P5



(d)
0 50 100 200 Km
N
BC
AB
SK
P11
P6



(e)
N
0 50 100 200 Km
AB
CANADA
EmsB profiles per WMU
P1
P2
P3
P4
P5
P7
P8
P9
P10
P11
P12
P13
P14
P15
P16