

1 A high throughput deep amplicon sequencing method to show the emergence  
2 and spread of *Calicophoron daubneyi* in United Kingdom cattle herds.

3 Neil D. Sargison<sup>1</sup> <sup>⊖</sup>, Kashif Shahzad<sup>2</sup> <sup>⊖</sup>, Stella Mazeri<sup>1</sup>, Umer Chaudhry<sup>1\*</sup>

4

5 1. University of Edinburgh, Royal (Dick) School of Veterinary Studies and Roslin Institute, Easter  
6 Bush Veterinary Centre, Midlothian, Scotland, EH25 9RG, UK

7

8 2. Department of Infection Biology, University of Skovde, Sweden

9

10 <sup>⊖</sup> Authors with equal contribution

11 \* Corresponding author: Umer Chaudhry Email: [uchaudhr@exseed.ed.ac.uk](mailto:uchaudhr@exseed.ed.ac.uk)

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30 **Abstract**

31 The prevalence of *C. daubneyi* infection in the United Kingdom has increased, but despite the  
32 potential for rumen flukes to cause production loss in ruminant livestock, understanding of their  
33 emergence and spread is poor. Here we describe the development of a method to explore the  
34 multiplicity of *C. daubneyi* infection and patterns of the parasite's emergence and spread, based on  
35 Illumina MiSeq deep sequencing of meta barcoded amplicons of a fragment of the mt-COX-1 locus.  
36 Our results show high levels of genetic diversity per infection and between populations of 10 to 47 of  
37 adult *C. daubneyi*, each from a total of 32 finished prime cattle consigned to slaughter from northern  
38 United Kingdom; with 18 unique mt-COX-1 haplotypes. This has implications for the adaptability of  
39 environmental and intermediate host stages of the parasite to changing climatic and animal  
40 management conditions, or of parasitic stages to exposure to anthelmintic drugs; potentially allowing  
41 for greater pathogenicity, or the development of anthelmintic resistance, respectively. Our results  
42 illustrate the impact of high levels of animal movements in the United Kingdom, whereby multiple  
43 common mt-COX-1 haplotypes were identified in 26 populations in the absence of geographical  
44 clustering of clades.

45

46 Keywords: *Calicophoron daubneyi*; rumen fluke; deep amplicon sequencing; genetic diversity.

47

48

49

50

51

52

53

54

55

56

57

58

59

## 60 **1. Introduction**

61 Digenean flukes belonging to the family Paramphistomidae are common parasites of livestock that  
62 are kept in a wide geographical range of environments supporting the intermediate mud, or water snail  
63 hosts (Huson et al., 2017). They are regarded as causes of severe disease and production loss in  
64 tropical and subtropical regions (Rolfe et al., 1991; Rangel-Ruiz et al., 2003). Historically,  
65 paramphistomes were considered to be uncommon in temperate regions (Wilmott, 1950; Deiana and  
66 Arru, 1963), but over the past two decades rumen flukes, now known to be *Calicophoron daubneyi*  
67 (Gordon et al., 2013), have become commonplace in wetter parts of Europe, notably in Spain and  
68 Portugal (Diaz et al., 2007; Arias et al., 2011; González-Warleta et al., 2013; Ferreras et al., 2014),  
69 France (Abrous et al., 2000; Mage et al., 2002; Rieu et al., 2007), Belgium (Malrait et al., 2015),  
70 Ireland (Murphy et al., 2008; Zintl et al., 2014; Toolan et al., 2015; Martínez-Ibeas et al., 2016) and  
71 the United Kingdom (Foster et al., 2008; Millar et al., 2012; Jones et al., 2017). Prior to the mid  
72 2000s, there were neither anecdotal reports of adult rumen flukes in cattle or sheep forestomachs, nor  
73 evidence of paramphistome eggs in faecal samples in the United Kingdom and Republic of Ireland;  
74 suggesting the new introduction and subsequent spread of the parasites. Pathogenesis and production  
75 loss due to *C. daubneyi* in Europe is mostly associated with the activity of immature flukes in the  
76 intestine (Foster et al., 2008; Mason et al., 2012), while there is little compelling evidence for adult  
77 flukes in the forestomach causing production loss (Sargison et al., 2016); putatively associated with a  
78 switch from feeding on host tissue in the intestine to microbial contents of the forestomach (reviewed  
79 by Huson et al., 2017).

80 A number of factors may have influenced the emergence and spread of *C. daubneyi* infection in  
81 Europe, such as: the abundance of suitable *Galba truncatula* intermediate hosts (Abrous et al., 1999;  
82 Jones et al., 2017); climatic change favouring the completion of the parasite's lifecycle (Skuce et al.,  
83 2013); the intensity of infection; and movement of free-living or parasitic stages between regions  
84 (Blouin et al., 1995). However, despite the pervasiveness of *C. daubneyi*, these epidemiological  
85 factors are poorly understood.

86 Abattoir-based studies have been used to explore the phylogenetics of *C. daubneyi* in cattle in the  
87 Republic of Ireland and in Northern Ireland, showing high levels of genetic diversity in the two  
88 countries and no evidence of geographical clustering of mitochondrial cytochrome c oxidase subunit I  
89 (mt-COX-1) DNA haplotypes (Zintl et al., 2014). The Median Joining Network had a linear rather  
90 than a star-like appearance, with the most commonly identified haplotype resembling the only  
91 contemporary NCBI reference published sequence (accession number: JQ815200), originating from  
92 Spain (Martínez-Ibeas et al., 2013). This implied either recent expansion of a small dispersed  
93 established population, or more probably multiple introductions from a larger mainland European  
94 parent population. These studies were based on generating individual two direction Sanger sequence  
95 data from 885 bp mt-COX-1 DNA fragments (Martínez-Ibeas et al., 2013) derived from 97 flukes (25

96 each from two farms and 20 from other cattle from the Republic of Ireland and 27 from other cattle  
97 presumed from Northern Ireland). Similar approaches, using high throughput technologies to allow  
98 the study of a larger and more diverse fluke population could potentially show if *C. daubneyi*  
99 infection emerged in United Kingdom cattle at a single time in a single cattle population, or  
100 repeatedly in different cattle populations, before spreading as a result of animal movement.

101 In this paper we describe a study using *C. daubneyi* adult parasites collected from prime cattle  
102 slaughtered in a large red meat abattoir, with the aims of: i) confirming the species identity of  
103 recovered rumen flukes; ii) identifying the presence of multiple genotypes per infection (multiplicity  
104 of infection) and iii) demonstrating the emergence and spread of *C. daubneyi* haplotypes. The species  
105 identity of rumen flukes was confirmed by deep amplicon sequencing of a 282 bp fragment of second  
106 internal transcribed spacer (ITS-2) rDNA. Haplotype diversity in 32 *C. daubneyi* populations, each  
107 derived from different single infected cattle consigned from different locations in the northern United  
108 Kingdom was shown by deep amplicon sequencing of a 333 bp fragment of the mt-COX-1 locus.  
109 Split and network trees of the mt-COX-1 haplotypes were examined to show patterns of emergence  
110 and spread of *C. daubneyi*, providing proof of concept for a novel approach to the study of parasite  
111 epidemiology.

112

## 113 **2. Materials and methods**

### 114 *2.1. Study resources*

115 Populations of adult *C. daubneyi* were collected from a total of 32 finished prime cattle  
116 slaughtered during 2014 in a large red meat abattoir in central Scotland during the winter (on 5  
117 sampling days between 13<sup>th</sup> January and 3<sup>rd</sup> March 2014) and summer/autumn (on 5 sampling days  
118 between 25<sup>th</sup> August and 6<sup>th</sup> October 2014). The sample collection strategy was designed both for  
119 logistical reasons, and to reduce the risk of disproportionate collection of multiple *C. daubneyi*  
120 populations from the same holdings, thereby allowing a representative distribution of cattle origins of  
121 consignment. All cattle slaughtered in United Kingdom abattoirs are uniquely identified and traceable  
122 under the auspices of the British Cattle Movement Services ([https://www.gov.uk/cattle-tracing-](https://www.gov.uk/cattle-tracing-online)  
123 [online](https://www.gov.uk/cattle-tracing-online)), showing each animal's origin prior to slaughter. Briefly, the identification of every tenth  
124 animal slaughtered on each sampling day was checked and recorded at the point of slaughter and the  
125 corresponding forestomachs were tagged to allow samples to be collected from the correct animals  
126 and matched to BCMS data. Forestomachs were incised along the greater curvature of the rumen and  
127 everted to remove their contents, as a standard part of the abattoir's tripe preparation process. The  
128 total numbers of adult *C. daubneyi* were enumerated and between 10 and 50 rumen flukes from each  
129 parasitised animal, depending on the numbers of flukes present (Supplementary Table 1), were fixed  
130 in 70% ethanol and archived.

131 *2.2. Genomic DNA extraction*

132 In order to avoid contamination with any progeny DNA, a ~2 mg section of each parasite's head  
133 was taken and rinsed twice for 5 min in a petri dish in ddH<sub>2</sub>O. The tissue sections were then lysed in  
134 lysis buffer and Protinease K (10mg/ml, New England BioLabs). The lysis buffer contained 50 mM  
135 KCL, 10 mM Tris (pH 8.3), 2.5 mM MgCl<sub>2</sub>, 0.045% Nonidet p-40, 0.45% Tween-20, 0.01% gelatin  
136 and ddH<sub>2</sub>O in 50ml volumes (Chaudhry et al., 2015a). Samples were lysed in 50 µl for 98 minutes at  
137 60°C followed by 15 minutes at 94°C then stored at -80°C until PCR.

138

139 *2.3. Confirmation of fluke DNA*

140 1 µl of each lysate containing fluke DNA was used to make population pools representing between  
141 10 and 47 parasites from each of the sampled cattle. A final dilution of 1:20 of pooled lysate: ddH<sub>2</sub>O  
142 was used as PCR template.

143

144 *2.4. Deep amplicon sequencing of r-ITS-2 and mt-COX-1 DNA*

145 Deep amplicon sequencing was performed from a 282 bp fragment of ITS-2 rDNA using  
146 previously published universal trematode primer sets (Adlard et al., 1993; Chaudhry et al., 2015a) and  
147 from a 333 bp fragment of mt-COX-1 using newly developed primers (Supplementary Table 2A)  
148 amplifying a locus within the 885 bp fragment used by Zintl et al. (2014). Adapters were added to  
149 allow the successive annealing of the primers and N is the number of random nucleotides included  
150 between each specific primer to increase the variety of generated amplicons. Four forward and four  
151 reverse primers were mixed in equal proportion in the first round PCR reaction made under the  
152 following conditions: 5X KAPA HiFi Fidelity buffer, 10 mM dNTPs, 10 µM forward and reverse  
153 adapter primer, 0.5 U KAPA HiFi Fidelity Polymerase (KAPA Biosystems, USA), 14µl ddH<sub>2</sub>O and  
154 1µl of worm lysate. The thermocycling conditions of the PCR were 95°C for 2 minutes, followed by  
155 35 cycles of 98°C for 20 seconds, 65°C for 15 seconds for ITS-2 and 70°C for 15 seconds for COX-1,  
156 72°C for 15 seconds and a final extension 72°C for 5 minutes. PCR products were purified with  
157 AMPure XP Magnetic Beads (1X) (Beckman coulter, Inc.) using a special magnetic stand and plate  
158 according to the protocols described by Beckman coulter, Inc.

159 A second round PCR was performed using eight forward and twelve reverse barcoded primers  
160 (Supplementary Table 2B), avoiding repetitions of the same barcoded primer combinations in  
161 different samples. The second round PCR conditions were; 5X KAPA HiFi Fidelity buffer, 10 mM  
162 dNTPs, 10 µM barcoded forward (N501 to N508) and reverse (N701 to N712) primers, 0.5 U KAPA  
163 HiFi Fidelity Polymerase (KAPA Biosystems,USA), 14 µl ddH<sub>2</sub>O and 2 µl of first round PCR  
164 product as DNA template. The thermocycling conditions of the PCR were 98°C for 45 seconds,  
165 followed by 7 cycles of 98°C for 20 seconds, 63°C for 20 seconds, and 72°C for 2 minutes. PCR  
166 products were purified with AMPure XP Magnetic Beads (1X) according to the protocols described  
167 by Beckman coulter, Inc. A pooled library was prepared from 10 µl of bead purified product from

168 each population and checked with KAPA qPCR library quantification kit (KAPA Biosystems, USA);  
169 before being run on an Illumina MiSeq Sequencer using a 500-cycle pair end reagent kit (MiSeq  
170 Reagent Kits v2, MS-103-2003) at a concentration of 15 nM with addition 25% Phix Control v3  
171 (Illumina, FC-11-2003).

172 The MiSeq separates all sequence by sample during post-run processing using recognised indices  
173 to generate FASTAQ files. These data were analysed with our own adapted pipeline. Briefly, a  
174 NCBI BLASTN search was used to generate and build consensus sequences of the *C. daubneyi* ITS-2  
175 rDNA and mt-COX-1 loci from FASTA files using Geneious Pro 5.4 software (Drummond et al.,  
176 2012). Data analysis was performed using Mothur v1.39.5 software (Schloss et al., 2009) and the  
177 Illumina Mi-seq standard operating procedures (Kozich et al., 2013). Overall, about 100,000 ITS-2  
178 rDNA and mt-COX-1 reads were generated from the MiSeq data set of 32 *C. daubneyi* populations.  
179 In summary, raw paired-end reads were made into contigs, and those that were too long or had  
180 ambiguous bases were removed. The sequence data were trimmed to the region amplified by the  
181 primers and aligned to the consensus sequence to only include this region. Any sequence data that did  
182 not hit with the ITS-2 rDNA and mt-COX-1 consensus sequences were discarded as being trace  
183 amplicon contamination.

184 ITS-2 rDNA and Mt-COX-1 sequence reads were then aligned in Geneious software (Kearse et al.,  
185 2012) before removing polymorphisms only occurring once, considered to be artefacts due to  
186 sequencing errors (Chaudhry et al., 2015b). The aligned sequences were then imported into the CD-  
187 HIT software (Huang et al., 2010) to calculate the frequency of the haplotypes present in each  
188 population, and those sequences showing 100% base pair similarity were collapsed into a single  
189 haplotype (cd-hit.org/).

190

### 191 2.5. Haplotypic diversity of the *C. daubneyi* mt-COX-1 locus

192 Within population haplotypic diversity was examined using all of the generated sequences; and  
193 between population haplotypic diversity was examined by dividing the number of sequence reads for  
194 each haplotype in each population by the number of flukes from which the sequences were derived.

195 The genetic diversity estimation was calculated from all of the sequences generated using the  
196 DnaSP 5.10 software package (Librado and Rozas, 2009). Briefly, sequence polymorphism was  
197 estimated through the haplotype frequency ( $H_f$ ), variance of haplotype diversity ( $H_v$ ), nucleotide  
198 diversity ( $\pi$ ), the mean number of pairwise differences ( $k$ ), the number of segregating sites ( $S$ ) and the  
199 mutation parameter based on an infinite site equilibrium model and the number of segregating sites  
200 ( $\theta_s$ ).

201

### 202 2.6. Split and network tree analysis of mt-COX-1 haplotypes

203 A split tree of the mt-COX-1 haplotypes was constructed by HKY+G model of substitution using  
204 the Maximum Likelihood method in the SplitTrees4 software (Huson and Bryant, 2006). The program

205 jModeltest 12.2.0 was used to select the appropriate model of nucleotide substitutions for Maximum  
206 Likelihood analysis (Posada, 2008). Branch supports were obtained by 1000 bootstraps of the data.

207 A network tree of 18 mt-COX-1 haplotypes generated this study and other haplotypes at the same  
208 locus published on NCBI GenBank was produced based on a neighbour joining algorithm using  
209 Network 4.6.1 software (Fluxus Technology Ltd), built on a sparse network with the epsilon  
210 parameter is set to zero default. The tree was rooted with the corresponding mt-COX-1 sequence of  
211 *Fasciola gigantica* (AJ853848).

212

### 213 2.7. Ethical approval

214 The abattoir based study was approved by the Veterinary Ethical Review Committee of the Royal  
215 (Dick) School of Veterinary Studies (VERC 21 14).

216

## 217 3. Results

### 218 3.1. Confirmation of species identity by ITS-2 rDNA sequence

219 The Illumina MiSeq ITS-2 rDNA data confirmed the species identity of *C. daubneyi* in each of the  
220 32 populations. In all cases, the gross morphology and the size of each worm was typical of *C.*  
221 *daubneyi*, hence the results supported the validity of the rDNA ITS-2 as a genetic marker to identify  
222 *C. daubneyi* (Gordon et al., 2013). Two unique haplotypes of the rDNA ITS-2 were identified among  
223 the 32 *C. daubneyi* populations, with a single nucleotide polymorphism at position 84 (A/G), when  
224 compared with NCBI GeneBank reference sequences derived from Ireland (accession number  
225 AB973394.1).

226

### 227 3.2. Within population haplotype frequency of mt-COX-1 locus

228 A total of eighteen unique haplotypes of the mt-COX-1 locus was identified among the 32 *C.*  
229 *daubneyi* populations, each derived from a single bovine host. In twelve populations, a single  
230 haplotype predominated at a frequency of 0.9 – 1.0 (Fig. 1). These comprised of five populations  
231 (Pop8, Pop16, Pop19, Pop20, Pop21) of between 10 and 22 parasites that contained a single  
232 haplotype, six populations (Pop1, Pop14, Pop25, Pop26, Pop29, Pop31) of between 16 and 35  
233 parasites that contained two haplotypes; and one population (Pop12) of 23 parasites that contained  
234 three haplotypes (Table 1, Fig. 1).

235 In contrast, twenty populations had high frequencies of multiple mt-COX-1 haplotypes (Fig. 1).  
236 These comprised of four populations (Pop7, Pop9, Pop13, Pop24) of between 11 and 37 parasites that  
237 contained six haplotypes; three populations (Pop3, Pop11, Pop32) of between 31 and 47 parasites that  
238 contained five haplotypes; seven populations (Pop4, Pop5, Pop17, Pop22, Pop23, Pop27, Pop30) of  
239 between 17 and 43 parasites that contained four haplotypes; and six populations (Pop2, Pop6,  
240 Pop10, Pop15, Pop18, Pop28) of between 12 and 28 parasites that contained a maximum of 3  
241 haplotypes (Table 1, Fig. 1).

242 Overall, the genetic diversity at both haplotype and nucleotide level indicated that the 32  
243 populations were highly diverse, with collapsed overall values of 0.949 and 0.011, respectively (Table  
244 1). There were no patterns in the geographical distribution of the farms from which cattle with more  
245 or less diverse populations were consigned to slaughter.

246 The results would be consistent with a single introduction of *C. daubneyi* infection to some of the  
247 farms where the cattle had been grazed during their life history and multiple time introductions to  
248 most.

249

### 250 3.3. *Between population phylogenetic analysis of mt-COX-1 locus*

251 Haplotypes HA8 and HA6 accounted for 35.9% and 13.6% of all of the sequence reads generated  
252 per fluke, being present in 10 and 9 populations, respectively. Five haplotypes (HA1, HA7, HA12,  
253 HA13 and HA16) accounted for between 5% and 10%; six haplotypes (HA2, HA3, HA5, HA9,  
254 HA17, HA18) accounted for between 1% and 5%; and 5 rare haplotypes (HA4, HA10, HA11, HA14,  
255 HA15) accounted for less than 1% of the sequence reads generated per fluke. Each of the rare  
256 haplotypes was present in between one and four populations.

257 The split tree of eighteen distinct mt-COX-1 haplotypes in 32 populations (Fig. 2) showed that  
258 nine haplotypes (HA1, HA6, HA8, HA9, HA12, HA13, HA15, HA16, HA17) were shared between  
259 different *C. daubneyi* populations derived from cattle consigned for slaughter from holdings in SW  
260 Scotland/NW England, East Scotland/NE England and Northern Ireland. Three haplotypes (HA2,  
261 HA7, HA18) were present in the populations derived from cattle consigned for slaughter from  
262 holdings in SW Scotland/NW England and East Scotland/NE England. Three haplotypes  
263 (HA5, HA10, HA11) were shared between populations derived from cattle consigned for slaughter  
264 from holdings in SW Scotland/NW England and Northern Ireland. The remaining three haplotypes  
265 (HA3, HA4, HA14) were only present in a single population derived from cattle consigned for  
266 slaughter from a holding in East Scotland, consistent with the emergence, but no spread of these  
267 haplotypes.

268 The split tree shows all eighteen haplotypes located in three distinct parts of the tree (Fig. 2). Four  
269 haplotypes (HA5, HA6, HA9, HA16) are present in a clade I; nine haplotypes (HA1, HA2, HA3,  
270 HA4, HA8, HA10, HA12, HA13, HA15) in a clade II; and five haplotypes (HA7, HA11, HA14,  
271 HA17, HA18) in a clade III of the network. Fig. 2 shows that the haplotypes are more closely related  
272 to each other within, rather than between clades, indicating at least three independent origins of *C.*  
273 *daubneyi* infection in the United Kingdom. The split tree showed no clear pattern of geographical  
274 distribution of haplotypes, or clades; hence the data are consistent with the spread of haplotypes  
275 across several locations after their emergence in the United Kingdom.

276 Twenty-three unique haplotypes of mitochondrial COX-1 locus were generated in the final  
277 filtration for the phylogenetic analyses. The network tree (Fig. 3) shows the 18 *C. daubneyi* mt-COX-  
278 1 haplotypes identified in northern United Kingdom, alongside twelve collapsed NCBI GenBank Irish



279 haplotypes, a Spanish haplotype and two collapsed NCBI GenBank haplotypes from South Africa and  
280 Zimbabwe. Five of the twelve NCBI GenBank Irish haplotypes were present in SW Scotland/NW  
281 England, East Scotland/NE England and Northern Ireland; three in SW Scotland/NW England and  
282 East Scotland/NE England; and one in SW Scotland/NW England and Northern Ireland.

283

#### 284 **4. Discussion**

285 There is concern about the probable recent introduction, increased prevalence and potential  
286 economic impact of rumen fluke infection of United Kingdom cattle. A study of 339 cattle  
287 slaughtered in a Scottish red meat abattoir described a high prevalence of rumen fluke infection in  
288 29% of cattle consigned from Scotland, Northern Ireland and northern England. The study described  
289 the geographical distribution of rumen fluke infection, estimated the minimal impact of adult flukes  
290 on production, and demonstrated the value of faecal egg counts as a tool to diagnose infection in live  
291 animals and to study the epidemiology of the disease (Sargison et al., 2016). Parasites collected from  
292 the infected cattle were used as template for the current study.

293 Deep amplicon sequencing of meta barcoded mt-COX-1 DNA afforded a practical and relatively  
294 inexpensive method, when compared to conventional Sanger sequencing (Zintl et al., 2014), to  
295 investigate the haplotype diversity between and within populations of *C. daubneyi* rumen flukes  
296 derived from individual cattle consigned to slaughter from different locations in SW Scotland/NW  
297 England, East Scotland/NE England and Northern Ireland. The Illumina MiSeq platform allows for  
298 relatively error-free reads of up to 400 bp (Avramenko, R.W. et al. 2017; Ali et al., 2018), hence the  
299 requirement for *de novo* primer design to amplify a 333 bp fragment of the mt-COX-1 locus in this  
300 study. The 885 bp fragment (Martínez-Ibeas et al., 2013) used in the Irish cattle phylogenetic study  
301 (Zintl et al., 2014) was too large, hence unsuitable for the Illumina MiSeq platform, albeit our primers  
302 were designed to amplify a locus within the larger fragment to allow for genomic comparisons. Our  
303 method enabled the generation of haplotype sequences from a representative total of 721 parasites,  
304 showing the proportions of each haplotype present in populations in individual cattle consigned to  
305 slaughter from 32 different farms. The study design allowed the proportions of haplotypes present to  
306 be mapped to precise geographical locations from where cattle were consigned to slaughter.

307 However, infection of final hosts is considered to be cumulative, hence our observations may be  
308 confounded by both recent infection and burdens established over previous years (González-Warleta  
309 et al., 2013), depending on the production system and previous treatments effective against rumen  
310 flukes. Some of the cattle in our study may have spent their entire life on a single farm, while others  
311 may have had several movements prior to slaughter, hence their *C. daubneyi* burdens at slaughter  
312 could have been acquired on other farms, depending on their grazing history and the availability of  
313 intermediate snail hosts.

314 *C. daubneyi* was the only rumen fluke species identified by ITS-2 rDNA sequence data in our  
315 study. Data analysis of the full sequence of the ITS-2 rDNA locus only revealed intraspecific

316 variation at a single position, hence this locus was not useful for phylogenetic comparisons. The  
317 northern United Kingdom *C. daubneyi* were genetically indistinguishable from Irish populations  
318 (Zintl et al., 2014) at the ITS-2 rDNA locus. In contrast, our results showed high levels of haplotype  
319 diversity in the mt-COX-1 locus, with 18 unique haplotypes in populations derived from the northern  
320 United Kingdom. The nucleotide diversity was also high, highlighting large differences in nucleotide  
321 sequences at the same locus. These results are similar to those described in Irish cattle (Zintl et al.,  
322 2014), and are consistent with parasite introduction from multiple sources, and a subsequent high  
323 mutation rate. This may have important implications for the adaptability of the parasite to potentially  
324 favourable factors such as the availability of different intermediate hosts, illustrated by the  
325 exploitation of *Galba truncatula* snails by *C. daubneyi* and *Fasciola hepatica* in the United Kingdom  
326 (Jones et al., 2015) and changing climatic and management conditions for environmental stages  
327 (Skuce et al., 2013); or to potentially adverse factors such as the exposure of parasitic stages to  
328 anthelmintic drugs, allowing for the development of anthelmintic resistance.

329 An aim of the present study was to use novel molecular genetic approaches to gain insight into the  
330 emergence and spread of *C. daubneyi* infection in United Kingdom. At farm level, infection may  
331 emerge singly or on multiple occasions through animal movement, or translocation of intermediate  
332 hosts, or free-living egg, miracidia, cercaria or metacercaria stages. The presence of a single mt-  
333 COX-1 haplotype in five populations suggests a single emergence of *C. daubneyi* infection on the  
334 farms from which the parasite populations were derived, putatively associated with a combination of  
335 geographical isolation and infrequent animal introductions. In contrast, the identification of one  
336 population with three unique haplotypes is consistent with multiple introductions, but no subsequent  
337 spread of *C. daubneyi* infection, putatively in a cattle-finishing herd with frequent introduction of  
338 animals but no onward movements other than to the abattoir. Multiple common mt-COX-1  
339 haplotypes in 26 populations in the absence of geographical clustering of clades is consistent with  
340 multiple introductions on the farms from which the parasite populations were derived and subsequent  
341 spread of *C. daubneyi* infection. The theoretical alternative explanation of recent expansion of a small  
342 divergent established population is unlikely given the high variation in proportions of different  
343 haplotypes and the parasite's reproductive strategy; and due to the absence of reports of rumen flukes  
344 in United Kingdom livestock between 1950 (Wilmott, 1950) and 2008 (Foster et al., 2008). Hence,  
345 our results illustrate the potential impact of high levels of animal movements in the United Kingdom  
346 (Vernon, 2011), for example involving trade in weaned suckler calves and store cattle, on the spread  
347 of infection.

348 In summary, our genetic data provide first insights in the emergence and the spread of *C. daubneyi*  
349 infection in United Kingdom. Our findings suggest both single and multiple independent emergence  
350 of *C. daubneyi* infection at farm level. Most common mt-COX-1 haplotypes were identified in  
351 several populations across a range of geographic locations, highlighting the role of animal movements  
352 in the spread of infectious disease. This understanding is relevant to the educational dissemination

353 and implementation of sustainable parasite control strategies. The study provides proof of concept of  
354 a method that could be used in the study of host-parasite relationships to examine influences of host  
355 movement, animal management, availability of intermediate hosts, and climate change on the  
356 epidemiology of parasitic diseases.

357

### 358 **Acknowledgments**

359 We acknowledge Scotbeef Limited who fully funded this project and allowed us to carry out the  
360 sampling and collect material. We acknowledge our colleagues who helped with sampling as well as  
361 the slaughterhouse workers and the Meat Hygiene Service for making this work possible. Work at the  
362 Roslin Institute uses facilities funded by the Biotechnology and Biological Sciences Research Council  
363 (BBSRC).

364

### 365 **References**

- 366 Abrous, M., Rondelaud, D., Dreyfuss, G., Cabaret, J., 1999. Infection of *Lymnaea truncatula* and  
367 *Lymnaea glabra* by *Fasciola hepatica* and *Paramphistomum daubneyi* in farms of central France.  
368 *Vet. Res.* 30, 113-118.
- 369 Abrous, M., Rondelaud, D., Dreyfuss, G., 2000. A field study of natural infections in three freshwater  
370 snails with *Fasciola hepatica* and/or *Paramphistomum daubneyi* in central France. *J. Helminthol.* 74,  
371 189-194.
- 372 Adlard, R.D., Barker, S.C., Blair, D., Cribb, T.H., 1993. Comparison of the second internal transcribed  
373 spacer (ribosomal DNA) from populations and species of Fasciolidae (Digenea). *Int. J. Parasitol.* 23,  
374 423-425.
- 375 Ali, Q., Rashid, I., Shabbir, M.Z., Shahzad, K., Ashraf, K., Sargison, N.D., Chaudhry, U., 2018.  
376 Emergence and the spread of the F200Y benzimidazole resistance mutation in *Haemonchus*  
377 *contortus* and *Haemonchus placei* from buffalo and cattle in Pakistan. *Vet. Parasitol.*
- 378 Arias, M., Lomba, C., Dacal, V., Vázquez, L., Pedreira, J., Francisco, I., Piñeiro, P., Cazapal-Monteiro,  
379 C., Suárez, J.L., Díez-Baños, P., Morrondo, P., Sánchez-Andrade, R., Paz-Silva, A., 2011. Prevalence of  
380 mixed trematode infections in an abattoir receiving cattle from northern Portugal and north-west  
381 Spain. *Vet. Rec.* 168, 408.
- 382 Avramenko, R.W., Redman, E.M., Lewis, R., Bichuette, M.A., Palmeira, B.M., Yazwinski, T.A., Gilleard,  
383 J.S., 2017. The use of nemabiome metabarcoding to explore gastro-intestinal nematode species  
384 diversity and anthelmintic treatment effectiveness in beef calves. *Int. J. Parasitol.* 47, 893-902.
- 385 Blouin, M.S., Yowell, C.A., Courtney, C.H., Dame, J.B., 1995. Host Movement and the genetic  
386 structure of populations of parasitic nematodes. *Genetics* 141, 1007-1014.
- 387 Chaudhry, U., Paridon, B. van, Shabbir, M.Z., Shafee, M., Ashraf, K., Yaqub, T., Gilleard, J.S., 2015a.  
388 Molecular evidence shows that the liver fluke *Fasciola gigantica* is the predominant *Fasciola* species  
389 in ruminants from Pakistan. *J. Helminthol.* 11, 1-8.

- 390 Chaudhry, U., Redman, E.M., Muthusamy, R., Gilleard, J.S., 2015b. Genetic evidence for the spread  
391 of a benzimidazole resistance mutation across southern India from a single origin in the parasitic  
392 nematode *Haemonchus contortus*. *Int. J. Parasitol.* 45, 721-728.
- 393 Deiana, S., Arru, E., 1963. Further enquiries on paramphistomosis (*Paramphistomum cervi*) in  
394 Sardinia. *Ric. Sci. 2. Ser. Pt.2. Rend. B.* 35, 449.
- 395 Diaz, P., Pedreira, J., Sánchez-Andrade, R., Suárez, J.L., Arias, M.S., Francisco, I., Fernández, G., Díez-  
396 Baños, P., Morrondo, P., 2007. Risk periods of infection by *Calicophoron daubneyi* (Digenea:  
397 Paramphistomidae) in cattle from oceanic climate areas. *Parasitol. Res.* 101, 339-342.
- 398 Drummond, A.J., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., Heled, J., Kearse, M.,  
399 Markowitz, S., Moir, R., Stones-Havas, S., Sturrock, S., Thierer, T., Wilson, A., 2012. Geneious v5.6.
- 400 Ferreras, M.C., González-Lanza, C., Pérez, V., Fuertes, M., Benavides, J., Mezo, M., González-Warleta,  
401 M., Giráldez, J., Martínez-Ibeas, A.M., Delgado, L., Fernández, M., Manga-González, M.Y., 2014.  
402 *Calicophoron daubneyi* (Paramphistomidae) in slaughtered cattle in Castilla y León (Spain). *Vet.*  
403 *Parasitol.* 199, 268-271.
- 404 Foster, A.P., Otter, A., O'Sullivan, T., Cranwell, M.P., Twomey, D.F., Millar, M.F., Taylor, M.A., 2008.  
405 Rumen fluke (paramphistomosis) in British cattle (letter). *Vet. Rec.* 162, 528.
- 406 González-Warleta, M., Lladosa, S., Castro-Hermida, J.A., Martínez-Ibeas, A.M., Conesa, D., Muñoz, F.,  
407 López-Quílez, A., Manga-González, M.Y., Mezo, M., 2013. Bovine paramphistomosis in Galicia  
408 (Spain): Prevalence, intensity, aetiology and geospatial distribution of the infection. *Vet Parasitol.*  
409 191, 252-263.
- 410 Gordon, D.K., Roberts, L.C.P., Lean, N., Zadoks, R.N., Sargison, N.D., Skuce, P.J., 2013. Identification  
411 of the rumen fluke *Calicophoron daubneyi*, in GB livestock: possible implications for liver fluke  
412 diagnosis. *Vet. Parasitol.* 195, 65-71.
- 413 Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. *Mol.*  
414 *Biol. Evol.* 23, 254-267.
- 415 Huson, K.M., Oliver, N.A.M., Robinson, M.W., 2017. Paramphistomosis of ruminants: an emerging  
416 parasitic disease in Europe. *Trends Parasitol.* 33, 836-844.
- 417 Jones, R.A., Williams, H., Dalesman, S., Brophy, P., 2015. Confirmation of *Galba truncatula* as an  
418 intermediate host snail for *Calicophoron daubneyi* in Great Britain, with evidence of alternative snail  
419 species hosting *Fasciola hepatica*. *Parasites Vectors* 8, 656.
- 420 Jones, R.A., Brophy, P.M., Mitchell, E.S., Williams, H.W., 2017. Rumen fluke (*Calicophoron daubneyi*)  
421 on Welsh farms: prevalence, risk factors and observations on co-infection with *Fasciola hepatica*.  
422 *Parasitol.* 144, 237-247.
- 423 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A.,  
424 Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic:  
425 an integrated and extendable desktop software platform for the organization and analysis of  
426 sequence data. *Bioinformatics* 28, 1647-1649.
- 427 Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism  
428 data. *Bioinformatics* 25, 1451-1452.

- 429 Mage, C., Bourgne, H., Toullieu, J.M., Rondelaud, D., Dreyfuss, G., 2002. *Fasciola hepatica* and  
430 *Paramphistomum daubneyi*: changes in prevalences of natural infections in cattle and in *Lymnaea*  
431 *truncatula* from central France over the past 12 years. *Vet. Res.* 33, 439-447.
- 432 Malrait, K., Verschave, S., Skuce, P., Loo, H. van, Vercruysse, J., Charlier, J., 2015. Novel insights into  
433 the pathogenic importance, diagnosis and treatment of the rumen fluke (*Calicophoron daubneyi*) in  
434 cattle. *Vet. Parasitol.* 207, 134-139.
- 435 Martínez-Ibeas, A.M., González-Warleta, M., Martínez-Valladares, M., Castro-Hermida, J.A.,  
436 González-Lanza, C., Miñambres, B., Ferreras, C., Mezo, M., Manga-González, M.Y., 2013.  
437 Development and validation of a mtDNA multiplex PCR for identification and discrimination of  
438 *Calicophoron daubneyi* and *Fasciola hepatica* in the *Galba truncatula* snail. *Vet. Parasitol.* 195, 57-  
439 64.
- 440 Martínez-Ibeas, A.M., Munita, M.P., Lawlor, K., Sekiya, M., Mulcahy, G., Sayers, R., 2016. Rumen  
441 fluke in Irish sheep: prevalence, risk factors and molecular identification of two paramphistome  
442 species. *BMC Vet. Res.* 12, 143.
- 443 Mason, C., Stevenson, H., Cox, A., Dick, I., 2012. Disease associated with immature paramphistome  
444 infection in sheep. *Vet. Rec.* 170, 343-344.
- 445 Millar, M., Colloff, A., Scholes, S., 2012. Disease associated with immature paramphistome infection  
446 (letter). *Vet. Rec.* 171, 509-510.
- 447 Murphy, T.M., Power, E.P., Sanchez-Miguel, C., Casey, M.J., Toolan, D.P., Fagan, J.G., 2008.  
448 Paramphistomosis in Irish cattle (letter). *Vet. Rec.* 162, 831.
- 449 Posada, D., 2008. JModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25, 1253–1256.
- 450 Rangel-Ruiz, L.J., Albores-Brahms, S.T., Gamboa-Aguilar, J., 2003. Seasonal trends of  
451 *Paramphistomum cervi* in Tabasco, Mexico. *Vet. Parasitol.* 116, 217-222.
- 452 Rieu, E., Recca, A., Bénet, J.J., Saana, M., Dorchies, P., Guillot, J., 2007. Reliability of coprological  
453 diagnosis of *Paramphistomum* sp. infection in cows. *Vet. Parasitol.* 146, 249-253.
- 454 Rolfe, P.F., Boray, J.C., Nichols, P., Collins, G.H., 1991. Epidemiology of paramphistomosis in cattle.  
455 *Int. J. Parasitol.* 21, 813-819.
- 456 Sargison, N., Francis, E., Davison, C., Bronsvoort, B.M., Handel, I., Mazeri, S., 2016. Observations on  
457 the biology, epidemiology and economic relevance of rumen flukes (Paramphistomidae) in cattle  
458 kept in a temperate environment. *Vet. Parasitol.* 219, 7-16.
- 459 Skuce P.J., Morgan, E.R., Dijk, J. van, Mitchell, M., 2013. Animal health aspects of adaptation to  
460 climate change: beating the heat and parasites in a warming Europe. *Animal* 7(s2), 333-345.
- 461 Toolan, D.P., Mitchell, G., Searle, K., Sheehan, M., Skuce, P.J., Zadoks, R.N., 2015. Bovine and ovine  
462 rumen fluke in Ireland-Prevalence, risk factors and species identity based on passive veterinary  
463 surveillance and abattoir findings. *Vet. Parasitol.* 212, 168-174.
- 464 Vernon, M.C., 2011. Demographics of cattle movements in the United Kingdom. *BMC Vet. Res.* 7(31).  
465 <https://doi.org/10.1186/1746-6148-7-31>
- 466 Wilmott, S., 1950. On the species *Paramphistomum* Fischoeder, 1901 occurring in Britain and  
467 Ireland with notes on some material from the Netherlands and France. *J. Helminthol.* 24, 155-170.

468 Zintl, A., Garcia-Campos, A., Trudgett, A., Chryssafidis, A.L., Talavera-Arce, S., Fu, Y., Egan, S., Lawlor,  
469 A., Negredo, C., Brennan, G., Hanna, R.E., De Waal, T., Mulcahy, G., 2014. Bovine paramphistomes  
470 in Ireland. *Vet. Parasitol.* 204, 199-208.

471

## 472 **Figure Legends**

473 **Fig. 1.** Relative allele frequencies of eighteen individual haplotypes in thirty-two populations. The  
474 colours in the pie chart circles indicate the haplotype frequency in each individual population in the  
475 map.

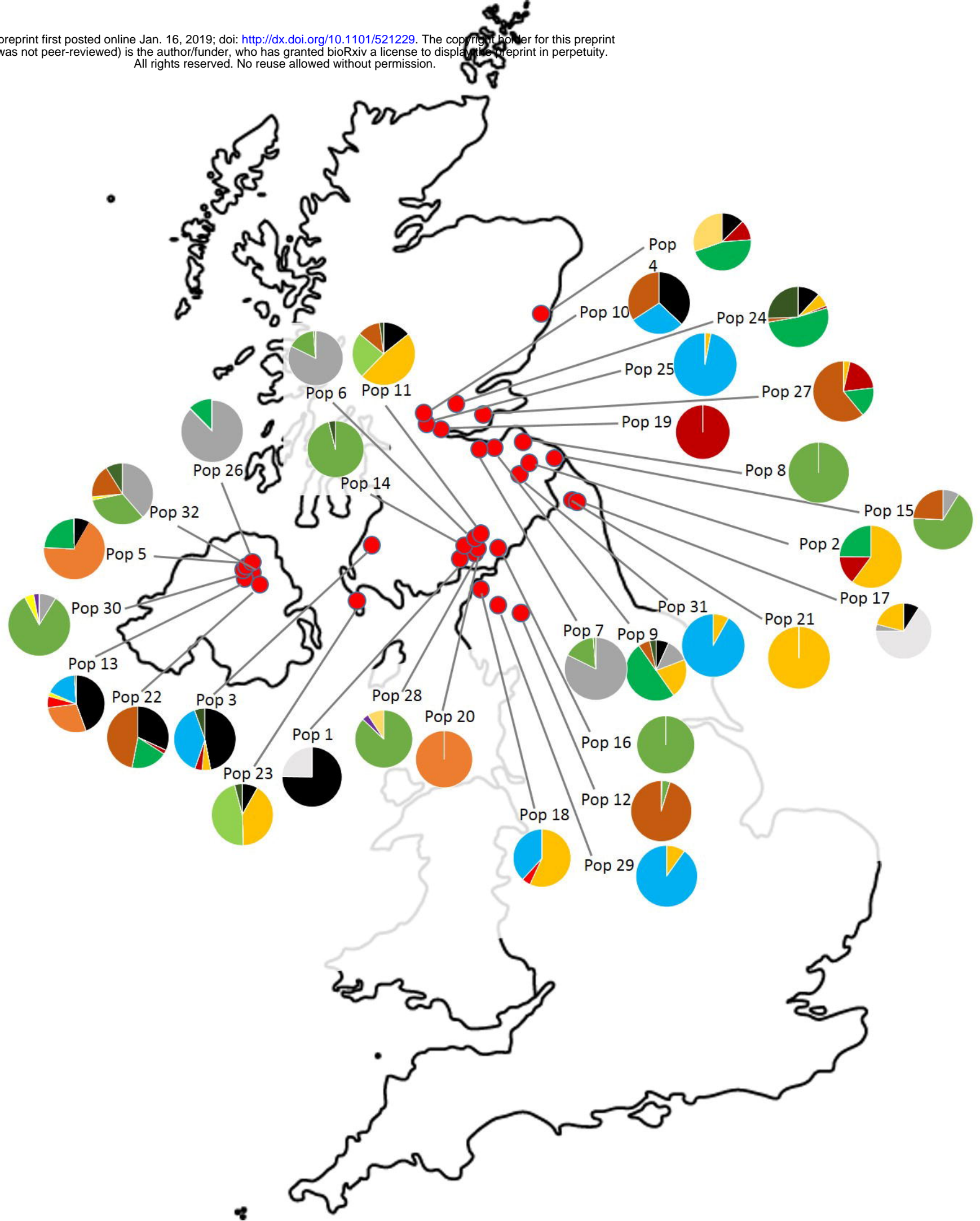
476 **Fig. 2.** Split tree of the 18 mt-COX-1 haplotypes generated with the median-net method in  
477 SplitsTrees4 software (Huson and Bryant, 2006). The colours in the pie chart circles replicate the  
478 haplotype frequency in the thirty-two populations indicated on the map. The size of each haplotype  
479 circle represents the percentage of sequence reads generated per fluke, shown by rows in the inset  
480 table.

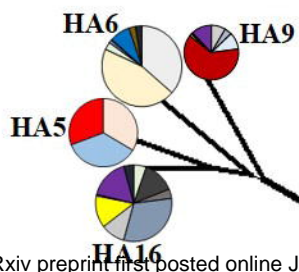
481 **Fig. 3.** Network tree of the 18 mt-COX-1 haplotypes of the present study aligned with 12 published  
482 NCBI GeneBank sequences of *C. daubneyi* from Ireland, Spain and Africa using Geneious Pro 5.4  
483 software (Kearse et al., 2012). Those sequences showing 100% base pair similarity are grouped into a  
484 single haplotype using the CD-HIT Suite software (Huang et al., 2010), resulting in 23 unique  
485 haplotypes in the final filtration for the phylogenetic analyses. The tree is rooted with the  
486 corresponding mt-COX-1 sequence of *Fasciola gigantica*.

487

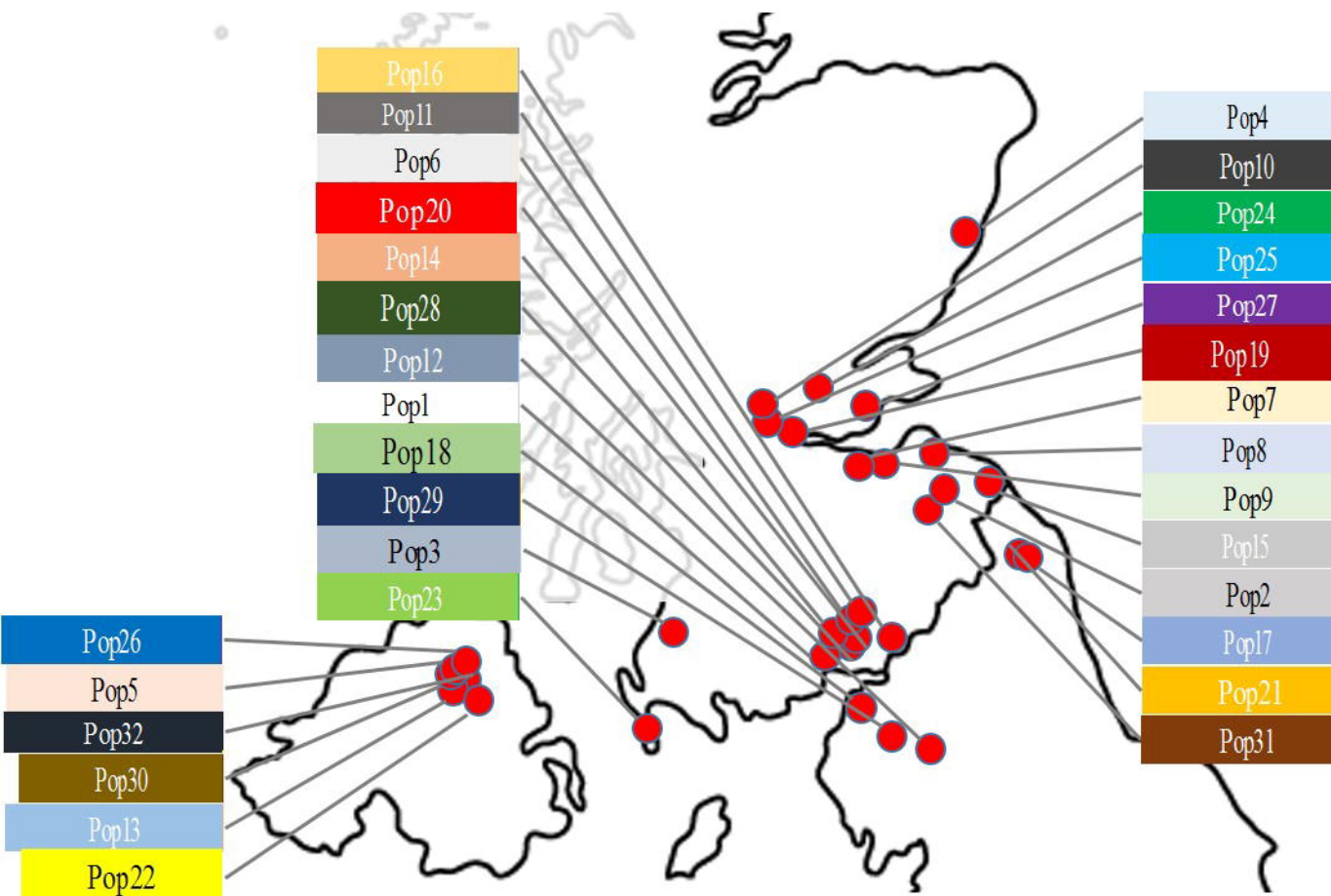
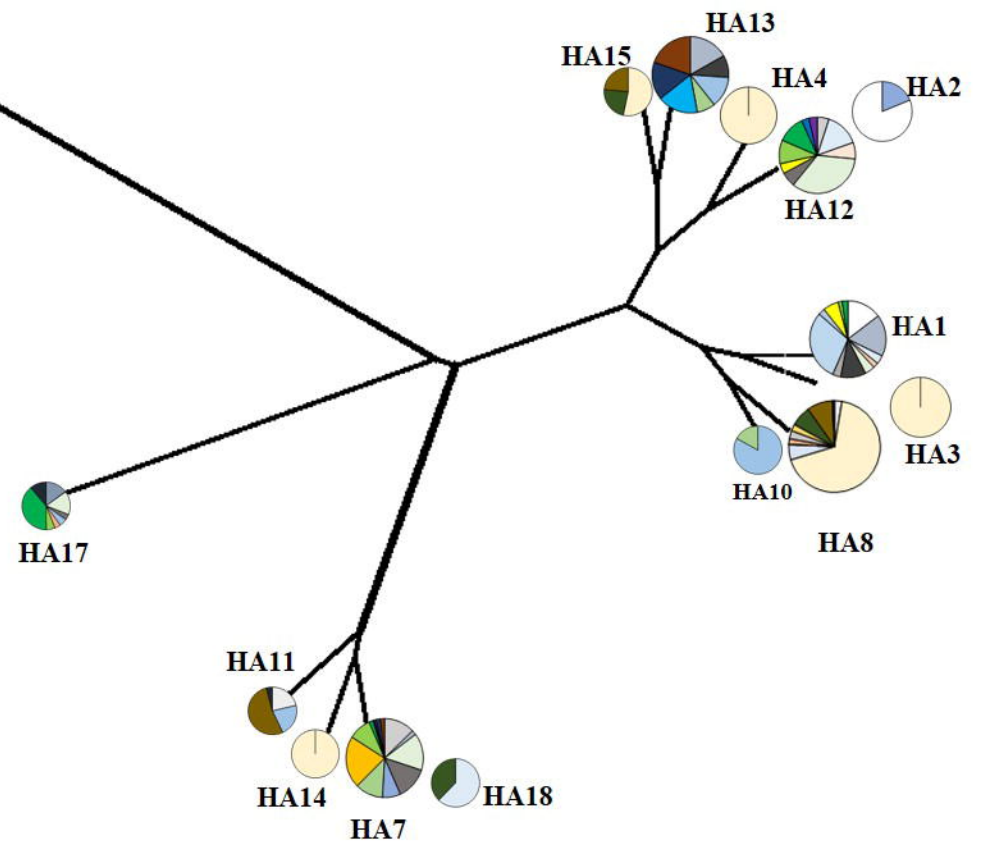
488

489



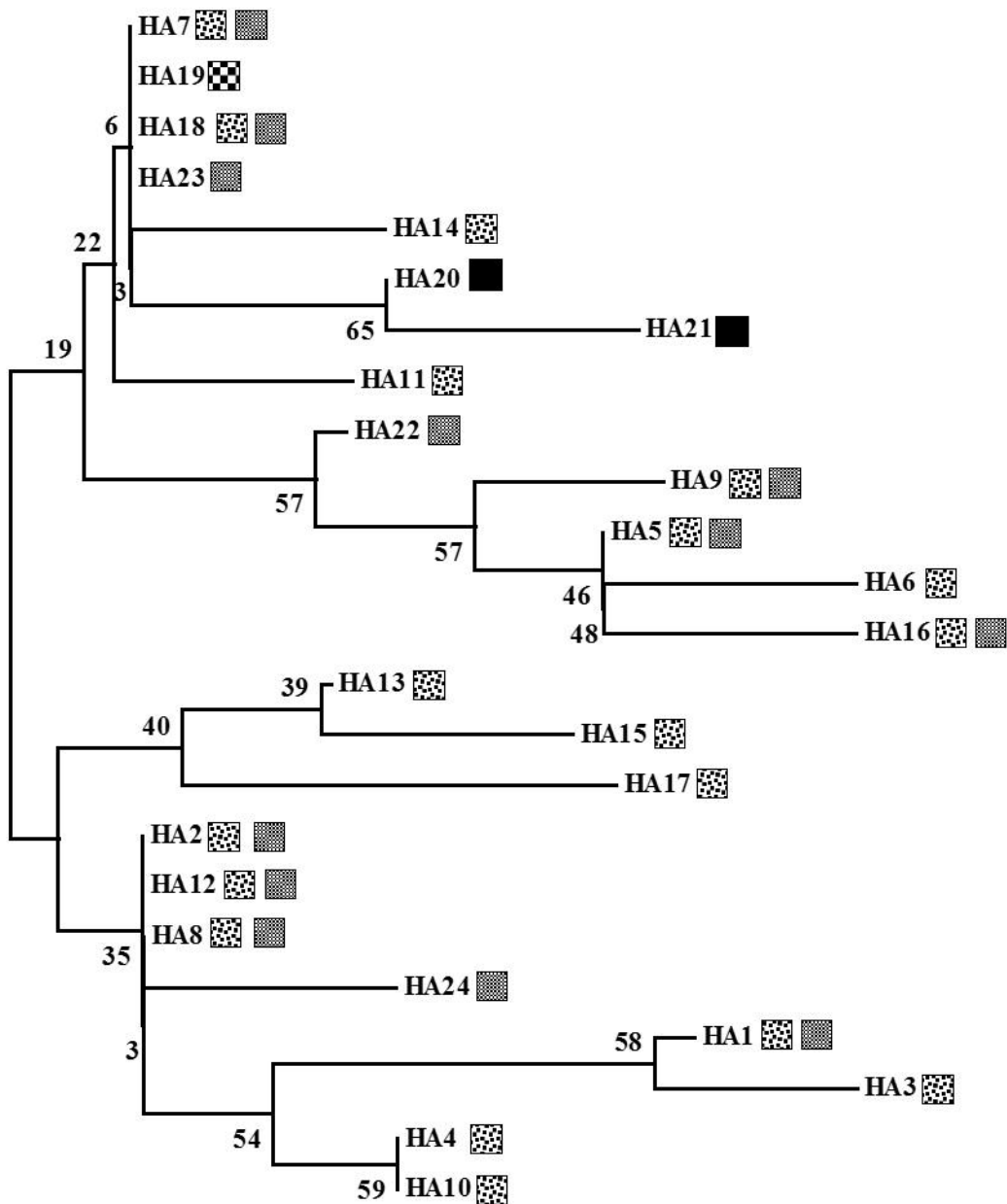


bioRxiv preprint first posted online Jan. 16, 2019; doi: <http://dx.doi.org/10.1101/521229>. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



Haplotype	Sequence reads generated per fluke (%)	Present in number of populations
HA10	0.3	2
HA11	0.4	4
HA15	0.5	3
HA14	0.5	1
HA4	0.7	1
HA18	1.0	2
HA3	1.2	1
HA17	1.3	9
HA2	1.4	2
HA5	3.6	3
HA9	3.9	7
HA16	5.8	9
HA7	6.7	13
HA13	6.8	7
HA1	8.0	12
HA12	8.4	10
HA6	13.6	9
HA8	35.9	10





0.001

UK   
 Ireland   
 Spain   
 Africa 

**Table 1:** Summary of genetic diversity of mt-COX-1 haplotypes identified from thirty-two populations of *C. daubneyi* in the United Kingdom.

Populations ID	Total number of Illumina MiSeq Reads	Total no of haplotype generated from each population	Haplotype diversity ( $H_d$ )	Segregating sites (S)	Nucleotide diversity ( $\Pi$ )	Mutation parameter based on S ( $\Theta_s$ )	Mean number of pairwise difference (k)
Pop1	1128	2	0.857	26	0.029	0.024	9.77
Pop2	1081	3	0.800	5	0.006	0.004	2.13
Pop3	2166	5	0.960	15	0.011	0.012	3.62
Pop4	1697	4	0.824	8	0.007	0.005	2.76
Pop5	1551	4	0.771	9	0.006	0.004	3.17
Pop6	6600	3	0.913	13	0.008	0.010	2.79
Pop7	3840	6	0.948	26	0.016	0.012	4.18
Pop8	2232	1			N/A		
Pop9	3668	6	0.965	21	0.012	0.022	4.08
Pop10	1675	3	0.824	10	0.008	0.003	4.31
Pop11	1419	5	0.952	16	0.011	0.012	3.79
Pop12	1381	3	0.824	8	0.008	0.007	2.94
Pop13	3896	6	0.977	34	0.021	0.021	4.22
Pop14	1819	2	0.812	7	0.006	0.005	2.26
Pop15	1783	3	0.935	43	0.016	0.013	5.29
Pop16	986	1			N/A		
Pop17	1824	4	0.957	16	0.012	0.013	4.01
Pop18	1039	3	0.913	9	0.006	0.007	2.21
Pop19	1190	1			N/A		
Pop20	1955	1			N/A		
Pop21	1085	1			N/A		
Pop22	1210	4	0.815	9	0.011	0.007	3.62
Pop23	1162	4	0.842	7	0.008	0.006	2.86
Pop24	1213	6	0.848	10	0.009	0.007	3.00
Pop25	1903	2	0.395	4	0.004	0.003	1.57
Pop26	1397	2	0.589	6	0.007	0.005	2.52
Pop27	1176	4	0.765	6	0.006	0.003	2.52
Pop28	3396	3	0.812	8	0.008	0.006	2.66
Pop29	1911	2	0.771	36	0.007	0.004	2.44
Pop30	4477	4	0.754	7	0.006	0.004	2.05
Pop31	1090	2	0.692	5	0.007	0.003	2.30
Pop32	1993	5	0.912	12	0.012	0.010	3.94
<b>Total</b>	<b>64943</b>	<b>18</b>	<b>0.949</b>	<b>76</b>	<b>0.011</b>	<b>0.038</b>	<b>3.597</b>