Queen's University Belfast

# Substantial genetic divergence between morphologically indistinguishable populations of Fasciola suggests the possibility of cryptic speciation.

Walker, S. M., Prodöhl, P., Hoey, E. M., Fairweather, I., Hanna, R. E. B., Brennan, G., & Trudgett, A. (2012). Substantial genetic divergence between morphologically indistinguishable populations of Fasciola suggests the possibility of cryptic speciation. International Journal for Parasitology, 42(13-14), 1193–1199. DOI: 10.1016/j.ijpara.2012.10.007

**Published in:**
International Journal for Parasitology

**Document Version:**
Peer reviewed version

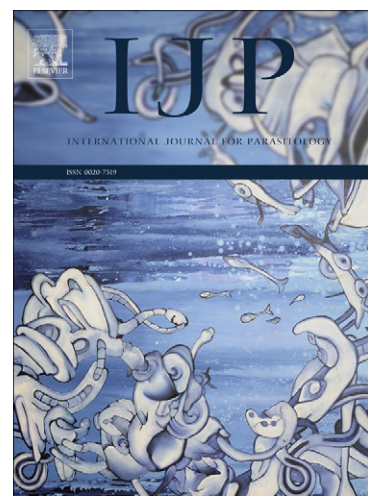
**Queen's University Belfast - Research Portal:**
Link to publication record in Queen's University Belfast Research Portal

**General rights**
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Download date:16. Feb. 2017

Accepted Manuscript




Substantial genetic divergence between morphologically indistinguishable populations of *Fasciola* suggests the possibility of cryptic speciation

S.M. Walker, P.A. Prodöhl, E.M. Hoey, I. Fairweather, R.E.B. Hanna, G. Brennan, A. Trudgett

PII: S0020-7519(12)00252-4
DOI: http://dx.doi.org/10.1016/j.ijpara.2012.10.007
Reference: PARA 3464

To appear in: *International Journal for Parasitology*

Received Date: 13 July 2012
Revised Date: 15 October 2012
Accepted Date: 16 October 2012

Please cite this article as: Walker, S.M., Prodöhl, P.A., Hoey, E.M., Fairweather, I., Hanna, R.E.B., Brennan, G., Trudgett, A., Substantial genetic divergence between morphologically indistinguishable populations of *Fasciola* suggests the possibility of cryptic speciation, *International Journal for Parasitology* (2012), doi: http://dx.doi.org/10.1016/j.ijpara.2012.10.007

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

1 **Substantial genetic divergence between morphologically indistinguishable**
2 **populations of *Fasciola* suggests the possibility of cryptic speciation**☆
3

4 SM Walker[a], PA Prodöhl[a], EM Hoey[a], I Fairweather[a], REB Hanna[b], G Brennan[a], A
5 Trudgett[a,]*
6
7
8
9
10
11
12 [a]*School of Biological Sciences, The Queen's University of Belfast, Belfast, N. Ireland*
13
14 [b]*Veterinary Sciences Division, Agri-Food and Biosciences Institute, Stormont, Belfast,*
15 *N. Ireland*
16
17 *Corresponding author. Tel.: +44 28 9079 2125; fax: +44 28 9097 5877.
18 *E-mail address*: a.trudgett@qub.ac.uk
19

20 ☆This study contains GenBank submissions **JX236561-JX236579**; **JX236580-**
21 **JX236590**; **JX236591-JX236639**; **JX236643-JX236655**; **JX236508-JX236560**; and
22 **JX236640-JX236642**.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

23 **Abstract**

24 The liver flukes, *Fasciola hepatica* and *Fasciola gigantica,* are considered to be sister
25 species and between them present a major threat worldwide to livestock production.
26 In this study sequence data have been employed from informative regions of the
27 nuclear and mitochondrial genomes of over 200 morphologically *F. hepatica* –like or
28 *F. gigantica* –like flukes from Europe, sub-Saharan Africa and South Asia to assess
29 genetic diversity. Evidence is presented for the existence of four well-separated
30 clades: African *gigantica*-like flukes, Indian *gigantica*-like flukes, European *hepatica*-
31 like flukes and African high-altitude *hepatica*-like flukes. Application of the
32 Biological Species Concept to trematodes is problematic; however, the degree of
33 separation between these groups was sufficient for them to be considered as distinct
34 species using the four times rule for speciation.

35

36 *Keywords*: *Fasciola*, Genetic diversity, Cryptic species

37

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

38 **1. Introduction**

39 Fasciolosis, a food-borne infection by fasciolid trematodes (most commonly
40 *Fasciola hepatica* or *Fasciola gigantica*) causes losses to agriculture estimated at US
41 $2,000 million per annum (McManus and Dalton, 2006). It is also an emerging
42 zoonotic disease with up to 17 million people thought to be infected worldwide
43 (Hopkins, 1992) and more than 90 million (Keiser and Utzinger, 2005) at risk of
44 infection. In the developing world the economic impact of fasciolosis may be
45 especially severe as there is a heavy reliance on buffalo and oxen as draught animals.
46 The spread and increased incidence of fasciolosis due to climate change, drug
47 resistance and intensification of agriculture have already been noted with *F. hepatica*
48 in Europe and may be expected to occur with *F. gigantica* in the developing world
49 (Mas Coma et al., 2009).

50 The evolution of *F. hepatica* and *F. gigantica*, together with other members of
51 the Fasciolidae, has recently been studied in detail (Lotfy et al., 2008) and the analysis
52 presented by these authors resulted in *F. hepatica* and *F. gigantica* being grouped as
53 sister species. *Fasciola hepatica* has a cosmopolitan distribution, being reported from
54 Europe, the Americas, Australasia and the more temperate regions of Africa and Asia,
55 whereas *F. gigantica* appears to be more restricted in its range, being reported from
56 the tropical regions of Asia and Africa, and Hawaii (introduced in the 19$^{th}$ century)
57 (Torgerson and Claxton, 1999). It is possible that outside Eurasia, the distribution of
58 *F. hepatica* is the result of anthropogenic introductions in historical times to the
59 Americas and Oceania (Mas Coma et al., 2009). Irrespective of the region, the
60 successful establishment of either *F. hepatica* or *F. gigantica* is dependent on the
61 distribution of its preferred intermediate host (Walker et al., 2008). Although
62 domestic bovine and ovine herbivores sustain the bulk of the fasciolid population,



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

63 both *F. hepatica* and *F. gigantica* appear to be generalist parasites with regard to their
64 mammalian hosts and, thus, they are capable of infecting a wide range of species
65 (Torgerson and Claxton, 1999). This lack of specialisation with regard to the
66 definitive host, in which sexual reproduction takes place, might be expected to restrict
67 their evolutionary divergence (Maynard Smith, 1966). The phylogenetic proximity of
68 *F. hepatica* and *F. gigantica* has been recently underscored by the finding that the
69 transcriptomes of *F. gigantica* and *F. hepatica* showed homology (Blastx, $E < 1E^{-05}$)
70 for almost 90% of the protein sequences examined (Young et al., 2011). Where
71 intermediate hosts are present which can support both *F. hepatica* and *F. gigantica*,
72 such as in eastern Asia, hybrids between the species have been reported (reviewed in
73 Mas Coma et al., 2009). Such hybrids may be diploid, mixoploid or triploid and are
74 generally aspermic. Some diploid isolates of *F. hepatica* may also be aspermic,
75 suggesting that parthenogenic reproduction is relatively common in wild populations
76 of fasciolids (Itagaki et al., 2009; Hanna et al., 2008). In studies where both
77 mitochondrial and nuclear genes have been examined (Agatsuma et al., 2000; Itagaki
78 et al., 2005a, b) it has been shown that the mitochondrial genomes of hybrid flukes
79 from Japan and Korea may be of either *F. hepatica* or *F. gigantica* in origin as may
80 their nuclear genes. In contrast, with flukes from Vietnam the mitochondrial
81 sequences were exclusively of *F. gigantica* origin (Le et al., 2008). These differences
82 indicate that hybridisation events are relatively common and have occurred
83 independently on several occasions. Recently infra-populations of fasciolid flukes
84 from Egypt and Iran have also been shown to contain individuals bearing nuclear
85 genes derived from both *F. hepatica* and *F. gigantica* (Amer et al., 2011; Amor et al.,
86 2011). Taken together, this body of work seems to indicate that where *F. hepatica*
87 and *F. gigantica* are found in the same infra-population, hybridisation is a common

ACCEPTED MANUSCRIPT

88 occurrence and the ability of these hybrids to reproduce parthenogenically allows the
89 establishment of essentially clonal hybrid field populations. The aim of much of this
90 work has been to determine whether flukes from a particular region should be
91 regarded as *F. hepatica* or *F. gigantica* as this distinction is of importance with regard
92 to epidemiology (Mas-Coma et al., 2005). However, in view of the range of
93 reproductive strategies which may be operating in fasciolid populations (Fletcher et
94 al., 2004) it is questionable how well the species concept (as defined by Mayr, 1963)
95 can be applied to these trematodes (Kunz, 2002; de Meeûs et al., 2003). Moving from
96 theoretical to practical matters, control and therapy of fasciolosis in the foreseeable
97 future is likely to remain dependent on anthelmintic chemotherapy (Fairweather,
98 2011) or the development of a vaccine (McManus and Dalton, 2006). For these
99 strategies to be applied successfully in the different regions where fasciolosis is a
100 problem it is important that we have as full an understanding as possible of the
101 variability inherent in the target liver fluke populations.
102 Working with mitochondrial genomic material it has been shown that *F.*
103 *hepatica* populations can exhibit a high level of "intraspecific" diversity/divergence
104 within a relatively confined geographic region (Walker et al., 2007, 2011) and that *F.*
105 *hepatica*-like flukes can be found in geographical proximity to *F. gigantica* in
106 highland regions of eastern Africa where the local climatic conditions favour the
107 establishment of its preferred intermediate host, *Lymnaea (Galba) truncatula* (Walker
108 et al., 2008). The most parsimonious explanation for the origin of the highland
109 eastern African flukes is that they have been introduced relatively recently with
110 livestock of European origin. In order to determine more precisely the relationship of
111 these flukes to other populations of *F. hepatica* and *F. gigantica* a portion of the of
112 their lsrRNA (nLSU/28S rRNA) (Teofanova et al., 2011) was sequenced as an

ACCEPTED MANUSCRIPT

113 example of nuclear genomic material. This has been supplemented with sequences
114 from the highly informative mitochondrial genome region coding for Cox III, tRNA-
115 His and Cob (Walker et al., 2007) to determine mitochondrial lineages. Phylogenetic
116 tools have been applied to determine whether the fasciolids in our African samples
117 and in additional material from Europe and India could be characterised as either *F.*
118 *hepatica* or *F. gigantica*. We believe that this is the first time that such an extensive
119 dataset, in terms of geographic origin and number of samples ($n$ = 200) has been
120 analysed in this manner. Both nuclear and mitochondrial datasets indicate that the
121 division of these fasciolids into two species is simplistic and may conceal the
122 occurrence of cryptic speciation.
123
124 **2. Materials and methods**
125
126 *2.1. Sources of fasciolid specimens*
127 All samples of adult fasciolid worms were collected from abattoirs and thus
128 represent "wild-type" specimens. The African material was obtained from Iringa
129 (Tanzania), Mbeya (Tanzania), Kitulo (Tanzania), Njombe (Tanzania) and Amin el
130 Ghaysu (Egypt). Details of the Tanzanian collection sites have been presented
131 elsewhere (Walker et al., 2008). Indian samples were collected from Chennai
132 (southern India) and Aligarh (northern India). Flukes of (ultimately) European origin
133 were obtained from Ireland, The Netherlands, Greece and Australia. All flukes from
134 regions below 2,500 metres in Tanzania originated from cattle: three from Iringa,
135 three from Mbeya and one from Njombe. Highland (>2,500 m) Tanzanian flukes
136 were from a single infrapopulation found in a cow from the Kitulo region. All adult
137 flukes collected from India originated from water buffalo; one from Aligarh and one



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

138 from Chennai. Only one infrapopulation was sampled from Egypt, which was
139 harboured in a donkey. Flukes from Ireland were obtained from a cow whereas fluke
140 samples from The Netherlands (one), Greece (two) and Australia (one) were from
141 sheep. Where sufficient numbers of flukes were available, up to 24 flukes from each
142 infra-population were used for subsequent molecular analysis. All material was
143 initially washed in distilled water prior to storage in 99% molecular grade ethanol.
144

145 *2.2. Identification of fasciolid specimens*

146 Flukes were initially classified on a morphological basis and their body length
147 and width recorded (Walker et al., 2008). Flukes morphologically similar to *F.*
148 *gigantica* had a mean body width to body length ratio of 2.85 whereas those
149 morphologically similar to *F. hepatica* had a ratio of 1.59. Fasciolid species were
150 later identified according to the PCR-Restriction Fragment Length Polymorphism
151 protocol of Marcilla et al. (2002), based on the 28S rDNA gene.
152

153 *2.3. DNA extraction from adult liver flukes*

154 Approximately 25 $mm^2$ of fluke tissue were placed into 500 µl of 10%
155 Chelex® (Fluka) solution incorporating 10 µl of Proteinase K (Sigma) at a
156 concentration of 20 mg/ml. This was then heated in a heat block at 55˚C for 1 h,
157 followed by gentle vortexing and a further incubation at 95˚C for 30 min. The
158 mixture was then gently vortexed and centrifuged at 13,000 *g* for 10 s. A 200 µl
159 sample of the supernatant was removed and diluted 1:10 in deionised water before
160 storage at -20˚C.
161

162 *2.4. mtDNA analysis*



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

163 Details of the region used for the identification of liver fluke mitochondrial
164 haplotypes, their amplification by PCR and the subsequent sequencing of the
165 amplicons have been given elsewhere (Walker et al., 2011). Briefly, this comprised
166 1,400 bp of contiguous mtDNA enclosing the regions coding for cytochrome c
167 oxidase subunit III (cox III), transfer RNA histidine (tRNA-His) and cytochome b
168 (cob). Two primer sets were used to generate overlapping fragments in PCR: Primer
169 set 1: Fhmt1.1F 5′-gcttgtgggttttcttaggg-3′, Fhmt1.1R 5′-caaccaaacctcaacaacct-3′;
170 Primer set 2: Fhmt1.2F 5′-tgtggtgtcggagagttctg-3′ and , Fhmt1.2R 5′-
171 taaccataggatccgcctga-3. Fragment 1 consisted of nucleotides 77 to 881 of the
172 complete *F. hepatica* mitochondrial sequence (Le et al., 2001), whilst fragment two
173 ran from nucleotides 681 to 1,480. As the second primer set failed to amplify the
174 morphologically *F. gigantica*-like flukes, an additional primer set was designed, Fgmt
175 1.2F 5'-ggtgtcggagagttctgttg-3' and Fgmt1.2R 5'-accaaatcaggaaacaccaa-3'. PCR
176 products were purified as detailed elsewhere (Walker et al., 2011) and sequenced
177 commercially by Macrogen (Korea). The sequences comprising the non-African *F.*
178 *hepatica* dataset have been published previously (Walker et al., 2011; Teofanova et
179 al., 2011) and accession numbers may be obtained from those sources. The Indian
180 and African datasets have the following accession numbers: Aligarh, **JX236561-**
181 **JX236579**; Chennai, **JX236580-JX236590**; Iringa, **JX236591-JX236639**; Kitulo,
182 **JX236643-JX236655**; Mbeya, **JX236508-JX236560**; Njombe, **JX236640-**
183 **JX236642**.
184
185 *2.5. rDNA analysis*
186 Primers 28SF 5’- agctgattacccgctgaact-3’ and 28SR 5’-
187 ctgagaaagtgcactgacaag-3’were used for amplification of the region from 15 bp to 632



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

188 bp (618 bp in length), with GenBank accession number **AJ440788** (*F. hepatica* partial
189 28SrRNA gene, Bolivia: Northern altiplano) (Marcilla et al., 2002) using the PCR
190 conditions described by Teofanova et al. (2011). PCR products were purified and
191 sequenced as above.
192
193 *2.6. Data Analysis*
194 Raw sequencing data (both directions in each case) were initially assembled
195 using ChromasPro (Technelysium Pty. Ltd, Australia) software and subsequent
196 alignments were carried out in BioEdit (Hall, 1999). All sequencing variants were
197 double checked on chromatogram traces before the derived sequences were allowed to
198 progress for further analysis. DNAsp (Rozas et al., 2003) was used to calculate the
199 haplotype diversity, average number of nucleotide differences between sequences and
200 nucleotide diversity (Nei, 1987) for each sample. To investigate the phylogenetic
201 relationships and relative frequency among the resulting mtDNA haplotypes, taking
202 into consideration the geographic origin of the samples, a haplotype network was
203 constructed using the median-joining method (Bandelt et al., 1999), implemented in
204 Network 4.5.0.2 (fluxus-engineering.com, Fluxus Technology Ltd., UK, 2004). A
205 Bayesian phylogenetic analysis on the resulting mtDNA haplotypes was also
206 performed using MrBayes version 3.2 (Ronquist and Huelsenbeck, 2003) using the
207 GTR+G model of DNA substitution, estimated using MODELTEST v3.0, Akaike
208 Information Criterion (Posada and Crandall, 1998). Four replicates of the Markov
209 chain Monte Carlo (MCMC) search were run with four chains of 10,000,000
210 generations each. Trees were sampled every 1,000 generations following a burn-in of
211 25,000 generations for each replicate. A *Fasciola jacksoni* sequence was used as an
212 outgroup for the analysis. In an attempt to provide a quantitative assessment of



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

213 whether the differentiation between the mtDNA groups/clades identified from the
214 phylogenetic analyses was sufficient to confirm the existence of distinct species, the
215 data were analysed according to the "4x rule" species criterion suggested by Birky et
216 al. (2010). This approach is designed to identify clusters that are separated by $t \geq 4N_e$
217 generations (where $t$ is the time to most recent common ancestor and $N_e$ is the
218 effective population size), which is equivalent to the upper 95% confidence limit of
219 the coalescent time and the depth of separation formed by random drift. For two such
220 clades the ratio of divergence between individuals from each clade ($K$) and θ, the
221 Watterson estimator of population mutation rate, is given by the equation $K/\theta =$
222 $2t\mu/2N_e\mu \geq 8N_e\mu/2N_e\mu = 4$. Although we do not have an accurate measure of $N_e$ or
223 the mutation rate ($\mu$), these can be assumed to be the same for each clade and can thus
224 be eliminated from the equation. Speciation may therefore be considered to have
225 occurred (by this criterion) when the mean sequence divergence ($K$) between
226 individuals in two candidate clades is greater than 4 * θ, where θ [$\theta = \pi/(1-4\pi/3)$] is
227 derived from the mean sequence difference between individuals within a clade, π; this
228 parameter can be calculated from our sequence dataset.
229
230 **3. Results**
231
232 *3.1. Differentiation according to morphology and 28S rDNA*
233 Morphological criteria assigned all the Indian flukes to *F. gigantica*. All of
234 the European flukes were classified as *F. hepatica* as were the African flukes from
235 Kitulo (Tanzania) and Egypt. All other African flukes were of the *F. gigantica*
236 phenotype.

ACCEPTED MANUSCRIPT

237 The analysis of the 618 bp fragment from 28S rDNA is shown in Table 1.
238 There were four positions at which nucleotide variation was observed. These
239 occurred at positions 105, 130, 283 and 547. This variation was not in complete
240 accordance with the morphological determination of species. At position 105, a
241 substantial minority of the European and Egyptian flukes carried a G, as did all of the
242 Indian and most of the Tanzanian flukes. At locus 130, all but one of the European,
243 Australian and Egyptian flukes carried an A as did the flukes from Kitulo whilst the
244 Indian and other African flukes had a G at this position. However, one European and
245 one Tanzanian fluke were heterogeneous at this site. Position 283 distinguished
246 between African "*F. gigantica*" and Indian "*F. gigantica*", the former having a G as
247 opposed to the A seen in the other populations. Again, two heterogeneous flukes
248 were observed. Position 547 gave the best agreement with morphological
249 classification with the *F. hepatica*-like flukes bearing a C at this position and the *F.*
250 *gigantica*-like samples a T, although again two aberrant samples were noted. The
251 heterogeneity seen at positions 130, 283 and 547 in the European and Iringa
252 populations was due in each case to an individual fluke showing two alleles at each
253 position.
254
255 *3.2. Mitochondrial haplotype analysis*
256 The fluke samples which provided mitochondrial sequence data are shown in
257 Table 2 together with statistical data relating to the sequences. In general, this region
258 of the mitochondrial genome was extremely variable with the number of haplotypes
259 observed being approximately one-third of the sample size. The population from the
260 highland region of Tanzania (Kitulo) was the most diverse with a very high average
261 number of nucleotide differences between paired sequences and the highest levels of

ACCEPTED MANUSCRIPT

262 nucleotide diversity; this was in contrast to the population from the lowland regions of
263 Tanzania which was the least diverse.
264 A Bayesian tree constructed using the sequences of the unique mitochondrial
265 haplotypes and rooted using the homologous region from *F. jacksonii* is shown in Fig.
266 1. The *F. gigantica*- and *F. hepatica*- like flukes were clearly separated into two
267 deeply rooted clades. Within the "*gigantica*" clade there was a further division
268 between the *F. gigantica*-like flukes from Africa and those from India. A division of
269 comparable or greater depth existed between a subset of the highland (Kitulo) flukes
270 and the rest of the *F. hepatica*-like population. There was strong posterior probability
271 support (P=1) for each of these groupings. Within the *F. gigantica* African clade and
272 the "European" *F. hepatica* clade there was little geographical structuring.
273 In order to substantiate these findings through the use of an alternative
274 analytical tool, a Median Joining network was constructed using the mitochondrial
275 dataset (Fig. 2). This employs a different algorithm and allows for the simultaneous
276 existence of ancestral and derived haplotypes in the network – this is not possible
277 using trees. This analysis indicated that there were four clearly separated clades based
278 on geographic origin but with a subset of the highland Tanzanian flukes being more
279 closely associated with the European population than others from the Kitulo region.
280 The two *F. hepatica*-like clades were separated by almost three times the number of
281 nucleotide substitutions separating the Indian and African *F. gigantica*-like clades.
282
283 *3.3. Application of the 4x rule for speciation*
284 The population statistics used to test for speciation in accordance with the 4x
285 rule model are presented in Table 3. With the exclusion of the subset of highland



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

286 Tanzanian flukes, which clustered with the European *F. hepatica* population, it can be
287 seen that each clade meets this criterion for speciation.
288
289 **4. Discussion**
290 This study was prompted by our interest in the morphologically *F. hepatica*-
291 like liver flukes found in association with *Lymnaea truncatula*, the favoured
292 intermediate host for *F. hepatica*, in areas of the Tanzanian highlands above 2,500 m
293 (Walker et al., 2008). rDNA genes have been extensively employed as
294 representatives of the nuclear genome in studies of speciation in digeneans (Lotfy et
295 al., 2008) and the partial 28S subunit has been reported to be suitable for use in
296 distinguishing between *F. hepatica* and *F. gigantica* (Marcilla et al., 2002). Liver
297 flukes from eastern Europe are known to contain two 28S genotypes (Tefanova et al.,
298 2011) and the relative proportions of these has been shown to correlate with
299 environmental temperature when used in ecological niche modeling (Kantzoura et al.,
300 2011). Analysis of the partial 28S sequences (Table 1) at the 130 position gave good
301 discrimination between morphologically *hepatica*-like and *gigantica*-like flukes with
302 *hepatica*-like flukes carrying the 130A genotype whilst the *gigantica*-like flukes
303 carried a 130G transition. Position 283 was of interest in that it distinguished between
304 the *gigantica*-like flukes of African origin and those from India, and position 547 was
305 comparable to position 130 in its ability to discriminate between *hepatica*-like flukes
306 and *gigantica*-like flukes. Taken in toto the results from the analysis of the variable
307 positions in 28S rDNA indicate that the nuclear genome of the highland Tanzanian
308 flukes is related to that of the subset of *F. hepatica* of European origin which has been
309 associated with tolerance of high temperatures (Kantzoura et al., 2011) but they also
310 reveal that *F. gigantica* from Africa and India may differ in their nuclear genomes.



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

311 This may be of importance as to date most research on *F. gigantica* has been
312 undertaken with material from Asia with its applicability to African *F. gigantica*
313 being assumed.

314 As expected, the mitochondrial genome region analysed was highly
315 informative with regard to intra-population differences. The statistics presented in
316 Table 2 indicate that the *F. hepatica*-like flukes from the Tanzanian highland region
317 were by far the most diverse genetically. This was in contrast to the *F. gigantica*-like
318 Tanzanian flukes which provided the least genetically diverse group. It has been
319 proposed that *F. gigantica* originated in the east African lowlands and spread to Asia
320 following the domestication of bovids (Mas Coma et al., 2009). In general, more
321 recently colonised areas have lower genetic variability (Hewitt, 2000). The ranking
322 order of liver fluke genetic diversity seen in Table 2 (European>Indian>African
323 lowland) might be taken to imply that African populations of *F. gigantica* are less
324 ancient than those of India but our results may be influenced more by the size of the
325 geographic region from which the samples were collected than their relative age. A
326 confounding factor may be the frequency with which population bottle-necks may
327 have occurred – for example, Europe, southern Africa and India differ greatly in their
328 history of glacial cycles. An explanation of the high level of genetic diversity seen in
329 the highland Tanzanian flukes was provided by the analyses presented in Figs. 1 and 2
330 which show that this population is made up of two distinct clades. One of these was
331 20 nucleotide differences removed from the European *F. hepatica* clade whereas the
332 other was over 70 substitutions distant (Fig. 2).

333 The mitochondrial data supported the distinction seen with the nuclear 28S
334 rDNA gene in the African and Indian *F. gigantica*-like flukes in that these populations
335 were separated into two well supported clades. There was a deep separation between



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

336 the morphologically *F. hepatica*–like and the *F. gigantica*–like flukes, with the former
337 clade showing further sub-division into the flukes of ultimately European origin
338 (Netherlands/Ireland/Greece/Australia/Egypt (Mas Coma et al., 2009)) compared with
339 those from the Tanzanian highlands. Four of the five haplotypes present in the
340 Tanzanian highland population formed a clade separated from the European flukes by
341 at least 71 nucleotide substitutions (Fig. 2). Relating such data to a molecular clock is
342 hazardous for reasons discussed in detail elsewhere (Walker et al., 2011), but
343 assuming a rate of change for *F. hepatica* comparable with that reported for
344 *Schistosoma mansoni* (Morgan et al., 2005) of 4% per million years, these 71
345 substitutions would be equivalent to approximately 1 million years, ruling out the
346 possibility that they could have been introduced as parasites present in domesticated
347 herbivores. The remaining haplotype present in the highland Tanzanian fluke
348 population was represented by four flukes and was within the range of nucleotide
349 diversity seen in European *F. hepatica*. As such it may have been introduced together
350 with the cattle and sheep known to have been brought into this area in recent times
351 (Walker et al., 2008).

352 Little evidence was seen that would indicate the presence of hybrid flukes in
353 our populations. In all but two individuals the mitochondrial haplotype was
354 consistent with the nuclear 28S rDNA genotype. These aberrant flukes from the
355 European and Iringa populations were heterozygous at 28S rDNA positions 130, 283
356 and 547 which could indicate hybridization between an *F. hepatica*- like fluke and an
357 African *F. gigantica*- like fluke in previous generations. Whilst this is possible for the
358 Tanzanian fluke, it seems less likely in the case of the European fluke as it was from
359 an infrapopulation present in a Dutch sheep and would thus be distant from known
360 sources of *F. gigantica.*



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

361 The results presented in this study show that the four geographical populations
362 of *Fasciola* spp. can be assigned to four clades which are well separated genetically.
363 Application of the 4x rule for speciation (Table 3) suggested that these clades are
364 sufficiently divergent to be regarded as species under the definition forming the basis
365 of this model (Birky et al., 2010). However, it has been proposed that in the absence
366 of an internationally agreed definition of a species that is applicable to parasites we
367 should refrain from naming new species "just on the basis of a certain number of base
368 exchanges within their ribosomal DNA sequence" (Kunz, 2002). *Fasciola gigantica*
369 and *F. hepatica* are well established as separate species and can be distinguished by a
370 number of characteristics - morphologically (Periago et al., 2006), by their
371 preferences for intermediate hosts (*Radix natalensis* does not appear to be capable of
372 supporting the growth of *F. hepatica* (Boray, 1985)), the minimal temperature
373 required for egg hatching (Grigoryan, 1958; Ross and McKay, 1929) and by the
374 ability of *F. hepatica* to evade immune responses in the Indonesian Thin-Tailed sheep
375 (Roberts et al., 1997). If the Tanzanian *F. hepatica*-like flukes and African *F.*
376 *gigantica*-like flukes are separated from these populations due to the process of
377 speciation we might expect to discover differences of comparable importance to their
378 biology on closer examination (Nicolalde-Morejon et al., 2009; Nadler and Perez-
379 Ponce de Leon, 2011). Indeed, prior to 1965 (when it was synonymised with *F.*
380 *gigantica* (Kendall, 1965)) liver flukes from India and elsewhere in south Asia were
381 classified as *Fasciola indica* (Varma, 1953). This distinction was made on
382 morphological grounds, with their spines and eggs in particular showing distinct
383 differences from both European *F. hepatica* and African *F. gigantica*. With further
384 research on the four clades identified in this study it may become apparent that there

ACCEPTED MANUSCRIPT

385 are distinct differences in their physiology, behaviour or morphology. If this is the
386 case then it may be pragmatic at that stage to designate them as separate species.
387 In conclusion, it has been shown that four populations of *Fasciola spp*. from
388 hosts located in geographically and climatically different environments form
389 divergent lineages. It is to be presumed that the selective pressures which have
390 brought about this situation may also have led to the evolution of other adaptive
391 differences between the lineages. These may be expressed as biological
392 characteristics such as host preference, tolerance to high or low temperatures,
393 sensitivity to anthelmintics and immunoevasive mechanisms, all of which could be of
394 importance to the future spread and control of fasciolosis.
395

396 **Acknowledgements**
397 We would like to thank Professor P Ramasamy (University of Alagappa,
398 India) and Professor SMA Abidi (Aligarh Muslim University, India) for their
399 assistance in the collection of flukes. This study was supported in part by the DeLiver
400 consortium of the European Union VII Framework Programme (FOOD-CT-200X-
401 023025) and a British Society for Parasitology Ann Bishop Award (2006) to SMW.
402 The funders had no role in study design, data collection and analysis, decision to
403 publish or preparation of the manuscript.
404

405 **References**
406

407 Agatsuma, T., Arakawa, Y., Iwagami, M., Honzako, Y., Cahyaningshi, U., Kang,
408 S.Y., Hong, S.J., 2000. Molecular evidence of natural hybridization
409 between *Fasciola hepatica* and *F. gigantica*. Parasitol. Int. 49, 231–238.



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

410

411 Amer S., Dar Y., Ichikawa M., Fukuda Y., Tada C., Itagaki T., Nakai Y., 2011.
412 Identification of *Fasciola* species isolated from Egypt based on sequence
413 analysis of genomic (ITS1 and ITS2) and mitochondrial (NDI and COI)
414 gene markers. Parasitol Int. 60, 5-12.

415

416 Amor N., Halajian A., Farjallah S., Merella P., Said K., Slimane B.B., 2011,
417 Molecular characterization of *Fasciola* spp. from the endemic area of
418 northern Iran based on nuclear ribosomal DNA sequences. Exp Parasitol.
419 128, 196-204.

420

421 Birky C.W. Jr, Adams J., Gemmel M., Perry J., 2010, Using population genetic
422 theory and DNA sequences for species detection and identification in
423 asexual organisms. PLoS One. 5(5) e10609.

424

425 Boray, J.C., 1985. Flukes of domestic animals. In: Gaafar, Howard, Marsh (Eds.),
426 Parasites, Pests and Predators, Elsevier, New York, NY, 179–278.

427

428 Fairweather, I. 2011, Reducing the future threat from (liver) fluke: realistic
429 prospect or quixotic fantasy? Vet Parasitol. 180, 133-143.

430

431 Fletcher, H.L., Hoey, E.M., Orr, N., Trudgett, A., Fairweather, I. & Robinson,
432 M.W., 2004, The occurrence and significance of triploidy in the liver fluke,
433 *Fasciola hepatica*. Parasitology 128, 69–72.

434



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

435 Grigoryan G.A., 1958, Experimental study of *Fasciola gigantica* infestation in
436 sheep. Trudy Armyansk Inst Zhivotnovod i Veterinariya 3, 155-168.
437
438 Hall T.A., 1999, BioEdit: a user-friendly biological sequence alignment editor and
439 analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 95-98.
440
441 Hanna R.E., Edgar H., Moffett D., McConnell S., Fairweather I., Brennan G.P.,
442 Trudgett A., Hoey E.M., Cromie L., Taylor S.M., Daniel R., 2008,
443 *Fasciola hepatica*: Histology of the testis in egg-producing adults of several
444 laboratory-maintained isolates of flukes grown to maturity in cattle and
445 sheep and in flukes from naturally infected hosts. Vet Parasitol. 157, 222-
446 34.
447
448 Hewitt G., 2000, The genetic legacy of the Quaternary ice ages. Nature. 405, 907-
449 13.
450
451 Itagaki, T., Kikawa, M., Sakaguchi, K., Shimo, J., Terasaki, K., Shibahara, T.,
452 Fukuda, K., 2005a. Genetic characterization of parthenogenetic *Fasciola*
453 sp. in Japan on the basis of the sequences of ribosomal and mitochondrial
454 DNA. Parasitology 131, 679–685.
455
456 Itagaki, T., Kikawa, M., Terasaki, K., Shibahara, T., Fukuda, K., 2005b. Molecular
457 characterization of parthenogenic *Fasciola* sp. in Korea on the basis of
458 DNA sequence of ribosomal ITS1 and mitochondrial NDI gene. J. Vet.
459 Med. Sci. 67, 1115–1118.



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

460

461 Itagaki T., Sakaguchi K., Terasaki K., Sasaki O., Yoshihara S., Dung T.V.., 2009.
462 Occurrence of spermic diploid and aspermic triploid forms of *Fasciola* in
463 Vietnam and their molecular characterization based on nuclear and
464 mitochondrial DNA. Parasitol. Int. 58, 81-85.

465

466 Kantzoura V., Kouam M.K., Feidas H., Teofanova D., Theodoropoulos G., 2011,
467 Geographic distribution modelling for ruminant liver flukes (*Fasciola*
468 *hepatica*) in south-eastern Europe. Int. J. Parasitol. 41 747-53.

469

470 Kendall, S.B., 1965. Relationships between the species of Fasciola and the
471 molluscan hosts. Adv. Parasitol. 3, 59–98.

472

473 Kunz, W., 2002. When is a parasite species a species? Trends Parasitol. 18, 121-
474 124.

475

476 Le, T.H., De, N.V., Agatsuma, T., Nguyen, T.G.T., Nguyen, Q.D., McManus,
477 D.P., Blair, D., 2008. Human fascioliasis and the presence of
478 hybrid/introgressed forms of *Fasciola hepatica* and *Fasciola gigantica* in
479 Vietnam. Int. J. Parasitol. 38, 725–730.

480

481 Lotfy W.M., Brant S.V., DeJong R.J., Le T.H., Demiaszkiewicz A., Rajapakse
482 R.P., Perera V.B., Laursen J.R., Loker E.S., 2008. Evolutionary origins,
483 diversification, and biogeography of liver flukes (Digenea, Fasciolidae).
484 Am J Trop Med Hyg. 79, 248-55



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

485

486 Marcilla A., Bargues M.D., Mas-Coma S., 2002. A PCR-RFLP assay for the
487 distinction between *Fasciola hepatica* and *Fasciola gigantica*. Mol. Cell.
488 Probes 16, 327-33.

489

490 Mas-Coma S., Bargues M.D., Valero M.A., 2005. Fascioliasis and other plant-
491 borne trematode zoonoses. Int. J. Parasitol. 35, 1255-78.

492

493 Mas-Coma, S., Valero, M.A., Bargues, M.D., 2009. Fasciola, lymnaeids and
494 human fascioliasis, with a global overview on disease transmission,
495 epidemiology, evolutionary genetics, molecular epidemiology and control.
496 Adv. Parasitol. 69, 41-146.

497

498 Maynard Smith, J., 1966. Sympatric speciation. Am. Nat. 100**,** 637 -650.

499

500 McManus D.P. and Dalton J.P., 2006. Vaccines against the zoonotic trematodes
501 *Schistosoma japonicum*, *Fasciola hepatica* and *Fasciola gigantica*.
502 Parasitology 133, S43-S61.

503

504 de Meeûs, T., Durand, P., Renáud, F., 2003. Species concepts: what for? Trends
505 Parasitol. 19, 425-427.

506

507 Morgan J.A., Dejong R.J., Adeoye G.O., Ansa E.D., Barbosa C.S., Brémond P.,
508 Cesari I.M., Charbonnel N., Corrêa L.R., Coulibaly G., D'Andrea P.S., De
509 Souza C.P., Doenhoff M.J., File S., Idris M.A., Incani R.N., Jarne P.,

ACCEPTED MANUSCRIPT

510 Karanja D.M., Kazibwe F., Kpikpi J., Lwambo N.J., Mabaye A.,
511 Magalhães L.A., Makundi A., Moné H., Mouahid G., Muchemi G.M.,
512 Mungai B.N., Séne M., Southgate V., Tchuenté LA., Théron A., Yousif F.,
513 Zanotti-Magalhães E.M., Mkoji G.M., Loker E.S., 2005. Origin and
514 diversification of the human parasite *Schistosoma mansoni*. Mol. Ecol. 14,
515 3889-902.
516
517 Nadler S.A., Ponce de León G., 2011. Integrating molecular and morphological
518 approaches for characterizing parasite cryptic species: implications for
519 parasitology. Parasitology. 138, 1688-709.
520
521 Nicolalde-Morejón, F., Vergara-Silva, F., González-Astorga, J., Vovides, A.P.,
522 Espinosa de los Monteros, A., 2009. Reciprocal illumination of
523 morphological characters upon a molecular hypothesis supports the
524 proposal of a new species of cycad from Mexico. Systemat. Biodiver. 7,
525 73–79.
526
527 Nei, M., 1987. Molecular evolutionary genetics. (chapter 10) Columbia University
528 Press, New York.
529
530 Periago, M.V., Valero, M.A., Panova, M., Mas-Coma, S., 2006. Phenotypic
531 comparison of allopatric populations of *Fasciola hepatica* and *Fasciola*
532 *gigantica* from European and African bovines using a computer image
533 analysis system (CIAS). Parasitol. Res. 99, 368–378.
534

ACCEPTED MANUSCRIPT

535 Posada, D., Crandall, K.A. 1998. MODELTEST: testing the model of DNA
536 substitution. Bioinformatics. 14, 817-818.
537
538 Roberts, J.A., Estuningsih, E., Widjayanti, S., Wiedosari, E., Partoutomo, S.,
539 Spithill, T.W., 1997. Resistance of Indonesian thin tail sheep against
540 *Fasciola gigantica* and *F. hepatica.* Vet. Parasitol. 68, 69–78.
541
542 Ronquist, F., Huelsenbeck, J.P. 2003. MrBayes 3: Bayesian phylogenetic inference
543 under mixed models. Bioinformatics. 19, 1572-1574.
544
545 Ross I.C., McKay, A.C., 1929. The bionomics of *Fasciola hepatica* in New South
546 Wales and of the intermediate host *Limnaea brazieri*. Bulletin of the
547 Council of Scientific and Industrial Research Australia. No 43.
548
549 Rozas J., Sánchez-DeBarrio, J.C., Messeguer, X., Rozas, R. 2003. DnaSP. DNA
550 polymorphism analyses by the coalescent and other methods.
551 Bioinformatics. 19, 2496-2497.
552
553 Teofanova, D., Kantzoura, V., Walker, S., Radoslavov, G., Hristov, P.,
554 Theodoropoulos, G., I. Bankov, I., Trudgett, A., 2010. Genetic diversity of
555 liver flukes *(Fasciola hepatica)* from Eastern Europe. Infect. Genet. Evol.
556 11. 109-15.
557

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

558 Torgerson, P., Claxton, J., 1999. Epidemiology and control. In: Dalton (Ed.),
559 Fasciolosis, CAB International Publishing, Wallingford, United Kingdom,
560 pp. 113–149.
561
562 Varma, A.K., 1953. On *Fasciola indica* n. sp. with some observations on *F.*
563 *hepatica* and *F. gigantica*. J. Helminthol. 27, 185–198.
564
565 Walker, S.M., Prodöhl, P.A., Fletcher, H.L., Hanna, R.E.B., Kantzoura, V., Hoey
566 E.M., and Trudgett, A., 2007. Evidence for multiple mitochondrial lineages
567 of *Fasciola hepatica* (liver fluke) within infrapopulations from cattle and
568 sheep. Parasitol. Res. 101, 117-125.
569
570 Walker, S.M., Makundi,A.E., Namuba, F.V., Kassuku A.A, Keyyu, J., Hoey, E.,
571 Prödohl, P., Stothard, JR and Trudgett, A., 2008. The distribution of
572 *Fasciola hepatica* and *Fasciola gigantica* within southern Tanzania –
573 constraints associated with the intermediate host. Parasitology 135, 495-
574 503.
575
576 Walker, S.M, Johnston, C., Hoey E.M., Fairweather, I., Borgsteede, F., Gaasenbeek,
577 C., Prodöhl, P.A., A. Trudgett, A., 2011. Population dynamics of the liver
578 fluke, *Fasciola hepatica*: the effect of time and spatial separation on the
579 genetic diversity of fluke populations in the Netherlands. Parasitology 138,
580 215-23.
581

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

582 Young N.D., Jex A.R., Cantacessi C., Hall R.S., Campbell B.E., Spithill T.W.,
583 Tangkawattana S., Tangkawattana P., Laha T., Gasser R.B., 2011. A
584 portrait of the transcriptome of the neglected trematode, *Fasciola*
585 *gigantica*--biological and biotechnological implications. PLoS Negl Trop
586 Dis. 5(2):e1004.
587

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

588 **Figure legends**

589

590 **Fig. 1**. Bayesian phylogenetic tree of liver fluke (*Fasciola gigantica* and *Fasciola*
591 *hepatica*) mtDNA haplotypes (branch lengths are proportional to the number of
592 inferred changes). Different colours/shades represent the geographic origin of
593 clades/groups identified in this study. Nodal values represent posterior probability
594 support for particular groups. Only values higher than 0.52 are shown. Taxon name
595 codes: Fh, *Fasciola hepatica*; Fg, *Fasciola gigantica*; Du. Netherlands; In, India; S,
596 south; N, north; Tz, – Tanzania; I, Iringa; M, Mbeya; Eur, Europe (Ireland/Greece);
597 Ob, Australia. *Fasciola jacksonii* served as an outgroup.

598

599 **Fig. 2.** Median Joining Network for liver fluke (*Fasciola gigantica* and *Fasciola*
600 *hepatica*) mtDNA haplotypes. Different colours/shades represent the geographic
601 origins of clades/groups identified in this study. Sizes of nodes are proportional to
602 frequency of the haplotype within the dataset and distance between nodes within a
603 group is approximately proportional to the number of nucleotide changes between
604 haplotypes. For clarity of presentation the minimal number of nucleotide
605 substitutions between groups is indicated rather than displayed to scale.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

606 **Highlights**

607 *Fasciola hepatica* and *Fasciola gigantica* from Africa, Europe and India have been
608 compared

609 The analyses support the existence of four distinct clades

610 Cryptic speciation may have occurred, masking significant biological differences

ACCEPTED MANUSCRIPT

Figure 1 colour

ACCEPTED MANUSCRIPT

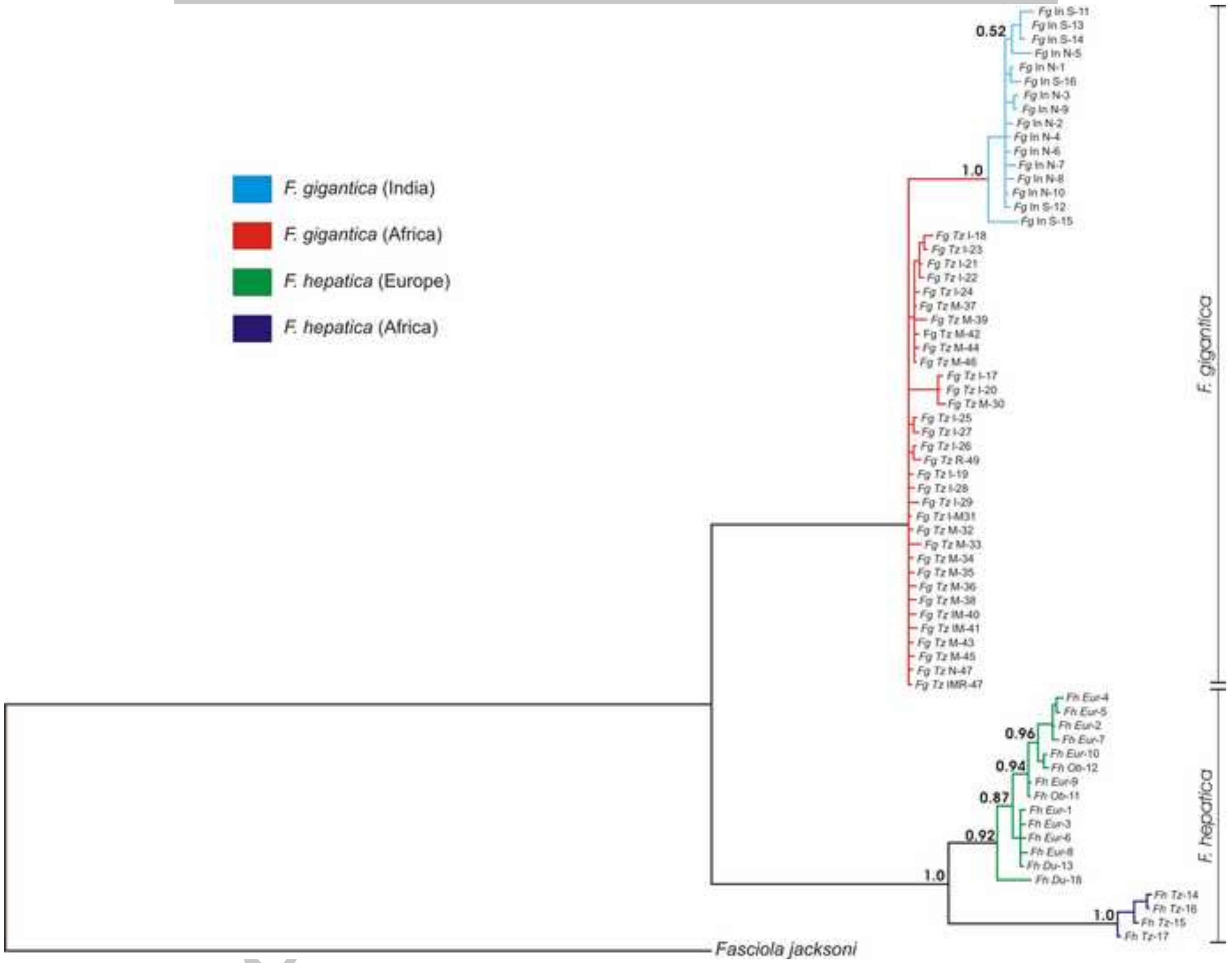
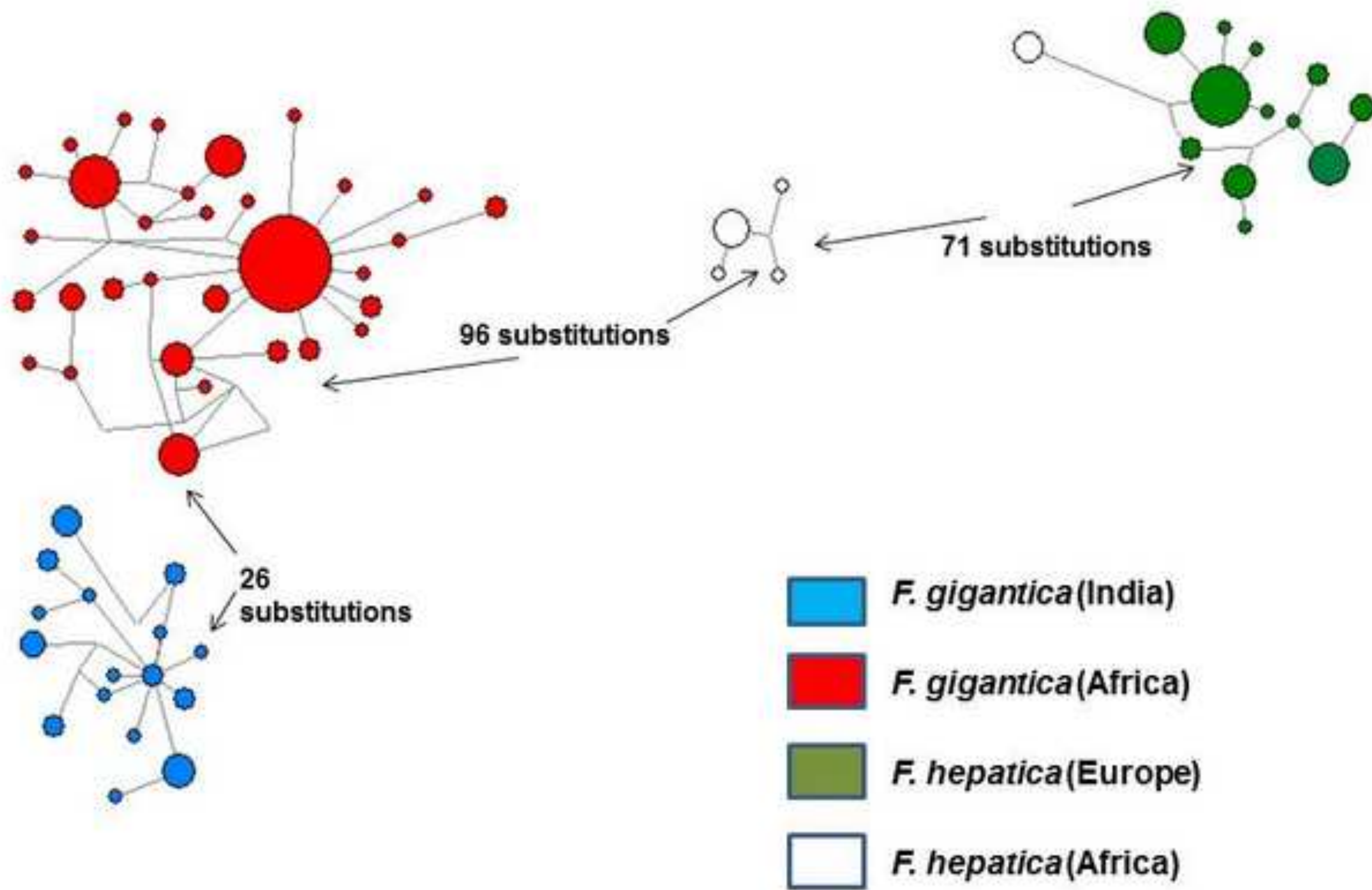


Figure 2 colour

ACCEPTED MANUSCRIPT



611

612

613 **Table 1**

614 28S rDNA diversity for fasciolids from different geographic regions.

615

616

| 28S Position | Origin of Flukes | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | European | Australian | Egyptian | South Indian | North Indian | Tanzanian Iringa | Tanzanian Kitulo | Tanzanian Mbeya | Tanzanian Njombe |
| 105 | | | | | | | | | |
| G | 12/32 | 6/6 | 1/4 | 10/10 | 20/20 | 52/52 | | 61/61 | 4/4 |
| | 20/32 | | 3/4 | | | | 15/16 | | |
| A | | | | | | | 1/16 | | |
| G/A | | | | | | | | | |
| 130 | | | | | | | | | |
| G | | | | 10/10 | 20/20 | 51/52 | | 61/61 | 4/4 |
| | 31/32 | 6/6 | 4/4 | | | | 16/16 | | |
| A | 1/32 [a] | | | | | 1/52 [b] | | | |
| G/A | | | | | | | | | |
| 283 | | | | | | | | | |
| G | | | | | | 51/52 | | 60/61 | 4/4 |
| | 31/32 | 6/6 | 4/4 | 10/10 | 20/20 | | 16/16 | 1/61 | |
| A | 1/32 [a] | | | | | 1/52 [b] | | | |
| G/A | | | | | | | | | |
| 547 | | | | | | | | | |
| C | 31/32 | 6/6 | 4/4 | | | | 16/16 | | |
| | | | | 10/10 | 20/20 | 51/52 | | 61/61 | 4/4 |
| T | 1/32 [a] | | | | | 1/52 [b] | | | |
| C/T | | | | | | | | | |

617

618 Numbers indicate the number of flukes within a set (numerator) and the number from

619 the respective geographic region sampled (denominator).

620 [a] Due to the same Dutch fluke.

621 [b] Due to the same Iringa fluke.

622



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

623
624
625 **Table 2**

626 Population and mitochondrial genome genetic statistics for fasciolids from different geographic regions.

627

| Populations | No. sequences | No. polymorphic sites | No. haplotypes | Ave no. nucleotide differences | Nucleotide diversity | Standard deviation | Haplotype diversity |
|---|---|---|---|---|---|---|---|
| All African *Fasciola gigantica*-like | 109 | 49 | 32 | 3.947 | 0.00346 | 0.00041 | 0.854 |
| All Indian *F. gigantica*-like | 30 | 39 | 16 | 7.533 | 0.00662 | 0.00096 | 0.945 |
| All European *Fasciola hepatica* | 50 | 28 | 12 | 8.579 | 0.00748 | 0.00051 | 0.836 |
| Kitulo TZ *F. hepatica*-like | 13 | 80 | 5 | 37.641 | 0.03435 | 0.00705 | 0.731 |

628
629 Australian and Egyptian flukes are consolidated with European flukes.



ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

630
631 **Table 3**
632 Variables associated with the populations used to test compliance with the "4x rule"
633 for speciation.
634
635

| Population | Nucleotide diversity (π) | θ | θ x 4 | Sequence divergence between clades (K) | K ≥ 4θ ? |
|---|---|---|---|---|---|
| All *Fasciola* spp. | | | | 0.10151 | |
| European *Fasciola hepatica* | 0.00748 | 0.00755 | 0.0302 | | Yes |
| All *Fasciola gigantica* | 0.01283 | 0.01283 | 0.05128 | | Yes |
| All *F. hepatica*-like | | | | 0.06595 | |
| European *F. hepatica* | 0.00748 | 0.00755 | 0.0302 | | Yes |
| Highland Tanzanian *F. hepatica* [a] | 0.0029 | 0.0029 | 0.0116 | | Yes |
| All *F. gigantica* spp. | | | | 0.03052 | |
| African *F. gigantica* | 0.00346 | 0.00346 | 0.01384 | | Yes |
| Indian *F. gigantica* | 0.00662 | 0.00663 | 0.02652 | | Yes |

636
637 [a] Highland Tanzanian samples clustering with the European clade (*n* = 4) were
638 excluded.
639

ACCEPTED MANUSCRIPT

*Graphical Abstract (for review)

ACCEPTED MANUSCRIPT