Low allelic diversity in vaccine candidates genes from different locations sustain hope for Fasciola hepatica immunization

Dominquez, M. F., Gonzalez-Miguel, J., Carmona, C., Dalton, J., Cwiklinski, K., Tort, J. F., & Siles-Lucas, M. (2018). Low allelic diversity in vaccine candidates genes from different locations sustain hope for Fasciola hepatica immunization. *Veterinary Parasitology*, 1-7. Advance online publication. https://doi.org/10.1016/j.vetpar.2018.06.011

Published in:
Veterinary Parasitology

Document Version:
Peer reviewed version

Low allelic diversity in vaccine candidates genes from different locations sustain hope for *Fasciola hepatica* immunization

Author list

Maria Fernanda Dominguez[1], Javier González-Miguel[2], Carlos Carmona[1], John P. Dalton[3], Krystyna Cwiklinski[3], José Tort[1], Mar Siles-Lucas[2*]

Affiliations

[1]Departamento de Genética, Facultad de Medicina, Universidad de la Republica, UDELAR, Montevideo, Uruguay; [2]Parasitology Unit, IRNASA-CSIC, Salamanca, Spain; [3]School of Biological Sciences, Queen's University Belfast, United Kingdom

*Corresponding author (mmar.siles@irnasa.csic.es)

Abstract

Fasciola hepatica is a trematode parasite that causes fasciolosis in animals and humans. Fasciolosis is usually treated with triclabendazole, although drug-resistant parasites have been described in several geographical locations. An alternative to drug treatment would be the use of a vaccine, although vaccination studies that have been performed mainly in ruminants over the last 30 years, show high variability in the achieved protection and are not yet ready for commercialisation. Since *F. hepatica* exhibits a high degree of genomic polymorphism, variation in vaccine efficacy could be attributed, at least partially, to phenotypic differences in vaccine candidate sequences amongst parasites used in the challenge infections. To begin to address this issue, a collection of *F. hepatica* isolates from geographically dispersed regions, as well as parasites obtained from vaccination trials performed against a field isolate from Uruguay and the experimentally maintained South Gloucester isolate (Ridgeway Research, UK), were compiled to establish a *F hepatica* Biobank. These collected isolates were used for the genetic analysis of several vaccine candidates that are important in host-parasite interactions and are the focus of the H2020 PARAGONE vaccine project (https://www.paragoneh2020.eu/), namely FhCL1, FhCL2, FhPrx, FhLAP and FhHDM. Our results show that *F. hepatica* exhibits a high level of conservation in the sequences encoding each of these proteins. The consequential low variability in these vaccine candidates amongst parasites from different geographical regions reinforces the idea that they would be suitable immunogens against liver fluke isolates worldwide.

Keywords: *Fasciola hepatica*; Genetic variation; Phenotypic variation; Vaccines

1. Introduction

Fasciola hepatica is the causative agent of fasciolosis, a parasitic disease of ruminants that seriously affects farm productivity worldwide as a result of livestock morbidity and mortality, as well as being an important zoonotic parasite of man (Cwiklinski et al., 2016; Carmona and Tort, 2017; Mehmood et al., 2017). The annual losses related to pathologies caused by fasciolosis have been reported to be 3 billion USD (Spithill et al., 2012). While triclabendazole is the most effective drug treatment, *F. hepatica* has rapidly developed drug resistance resulting in the widespread threat to livestock production systems (Kelley et al., 2016). Therefore, the development of an effective vaccine is paramount and would represent the most appropriate and sustainable way forward in the control of fasciolosis (Dalton et al., 2013).

Since the early 1990s, a growing number of vaccine trials in livestock have evaluated the efficacy of candidate antigens from *F. hepatica*. Among them, different parasite secreted antigens such as cathepsin L peptidases 1 and 2 (FhCL1, FhCL2), the antioxidant peroxiredoxin (FhPrx) and the gut-associated exopeptidase leucine aminopeptidase (FhLAP) have been selected as vaccine candidates due to their importance in host-parasite interactions (Dalton et al., 2013; Toet et al., 2014). These studies highlighted that high levels of variability in vaccine efficacy occur between trials which likely results from differences between the antigen source, the adjuvants used and the host species vaccinated. Variation between animals from the same or different breeds was also observed. In general, native antigens were more effective than recombinant vaccines, and with the exception of FhLAP combination of antigens perform better than single antigen formulations (see an overview of vaccine data in Table 1).

Following the sequencing of the *F. hepatica* genome, high levels of genetic polymorphism were observed, particularly in the chemosensory and neurodevelopmental pathways which might account for adaptations to the host environment and the capacity for rapid evolution (Cwiklinski et al., 2015). Further genome sequencing of liver flukes from two American locations also found polymorphisms between the *F. hepatica* isolates (McNulty et al., 2017). Both studies were based on

sequencing several individual parasites but opened the path to population genetic approaches, a much needed follow-up of the helminth genomic era (Wit and Gilleard, 2017). Analysis of UK isolates based on neutral markers (microsatellites) confirmed substantial variation within *F. hepatica* populations (Beesley et al., 2017), which complements the population genetics studies of liver fluke populations carried out using ribosomal and mitochondrial markers to unravel geographical variations (reviewed by Teofanova et al., 2012). Although marked genetic heterogeneity between liver fluke populations is now well recognised, an association between *F. hepatica* haplotypes and specific phenotypic traits has yet to be made (reviewed in Zintl et al., 2015).

To further the development of vaccines against a range of economically important parasitic pathogens of livestock, the EU H2020 funded consortium PARAGONE (https://www.paragoneh2020.eu/) has brought together liver fluke researchers with an aim to develop a multi-valent vaccine against *F. hepatica*. Our current vaccine candidates include those molecules that have shown potential in previous studies, including cathepsin L proteases (FhCL1), leucine aminopeptidase (FhLAP) and peroxiredoxin (FhPrx), in addition to assessing the *F. hepatica* helminth defence molecule (FhHDM) as a vaccine candidate. Specifically, this study evaluates FhHDM as a recombinantly expressed protein, which complements recent sheep vaccination trials by the Prof. Ubeira group using native and synthetically synthesised FhHDM (Martínez-Sernández et al., 2017; Orbegozo-Medina et al., 2018).

An important task of the project is to investigate the potential genetic variability of these vaccine targets within different *F. hepatica* isolates that may be a cause of variability in efficacy results between vaccine trials. Accordingly, we sourced *F. hepatica* from geographically dispersed liver fluke populations for genetic analysis and compared these results with an analysis of the liver fluke isolates that we used in our vaccine trials, namely a field isolate from Uruguay and the laboratory-maintained South Gloucester isolate from Ridgeway Research, UK. The collection of *F.*

hepatica isolates sourced as part of this study is now housed in a *F. hepatica* Biobank and is publically available.

2. Materials and methods

2.1. Parasite material and sample processing

A comprehensive biobank for *F. hepatica* was established at the Institute of Natural Resources and Agrobiology of Salamanca (IRNASA-CSIC) (Salamanca, Spain), with the aim of collecting representative samples of *F.hepatica* from different geographical locations, hosts and variable drug susceptibility/resistance. Samples were sourced from geographically dispersed regions collected at local abattoirs and stored individually in RNAlater (Sigma Aldrich). Samples from the PARAGONE *F. hepatica* vaccine trials were also included in the biobank from cattle experimentally infected with a field isolate from Uruguay and sheep experimentally infected with the laboratory maintained South Gloucester isolate (Ridgeway Research). The samples used for this study are detailed in Table 2.

All samples were processed upon arrival to the biobank following the European rules for handling and traceability of biologicalsamples. In brief, adult liver flukes stored in RNAlater were transversally sliced in half and used for (a) extraction of genomic DNA (Nucleospin Tissue, Magerey Nagel), and (b) extraction of RNA and subsequent cDNA synthesis (PureLink RNA Minikit, Thermo Fischer Scientific; Kit First Strand cDNA Synthesis, Roche), according to the manufacturer's instructions. The extracted genomic DNA and RNA were assessed for quality and quantity by OD at 260/280 nm and by gel electrophoresis. The nucleic acids from each sample were stored at -80°C in 2D-coded tubes of 500 µl volume (Wilmut, Spain). The code on each tube was automatically associated with a submission form (Supplementary file 1) containing the sample information, including isolate number, origin (country and locality), animal host, date of isolation, and type of isolate. The unique code given to each sample includes the Nomenclature of Territorial

112 Units for Statistics (NUTs) country code, and two successive ordinal numbers identifying the
113 country region and the sample number.

114 **2.2. PCR Amplification and sequencing**

115 Primers were designed to amplify the whole coding sequence of each gene, including the
116 start and stop codons (Table 3), based on the following *F. hepatica* sequences publicly available in
117 GenBank: FhCL1 (U62288), FhCL2 (U62289), FhPrx (FJ168037) and FhLAP(AY644459). Due to
118 the small size of the FhHDM (FR848429) gene transcript, this gene was directly amplified from
119 genomic DNA of samples from Uruguay and Argentina, resulting in a 490 bp product that included
120 the coding sequence and the corresponding introns. PCR reactions were performed in 25 µl volume
121 containing 0.1 µl Taq Platinum (5 U/µl; Thermo Fischer Scientific, Spain), 2.5 µl 10x Taq Platinum
122 buffer, 1 µl 25 mM $MgSO_4$, 1 µl each dNTP (10 mM), 0.5 µl of each primer (10 µM) and 1 µl
123 cDNA. The PCR cycling conditions consisted of an initial denaturation at 94°C for 30s, followed by
124 35 cycles at 94°C for 30s, 48°C (or 60°C for FhHDM) for 1min and 72°C for 2 min, followed by a
125 final extension at 72°C for 5 min. PCR products were analysed by gel electrophoresis (1% agarose
126 gel) and the bands of the expected molecular weight were excised and the DNA extracted according
127 to the manufacturer's instructions (GeneJET Gel Extraction and DNA Cleanup Kit, ThermoFisher
128 Scientific, Spain). The purified products were added to the corresponding forward and reverse
129 primers used in the PCR reaction in a 96 well plate and sent for automatic sequencing (Standard-
130 Seq; Macrogen, The Netherlands) by traditional sequencing in a 96-capillary 3730xl DNA Analyzer
131 (Thermo Fischer Scientific).

132 **2.3. Data analysis and variant calling**

133 Chromatogram files were processed with CLC Main Workbench (Qiagen). All sequences
134 were imported and trimmed by quality, removing regions and/or sequences with poor quality, and
135 aligned to the reference consisting of the complete coding sequences (CDS) for each vaccine target
136 gene. All variant sites compared with the reference sequence were automatically detected and

termed “conflicts”, which were manually inspected to confirm whether or not these differences represented true variants. For each target gene the coverage was determined for all positions (the number of times a good quality read was found for this particular position), and a global variation list generated. Finally a short representation of each sample was generated keeping only those positions that show variation in at least one sample or each target gene (Supplementaryfile 2).

Specific analysis was carried out where non-synonymous substitutions resulted in a change of amino acid residues critical for protein structure and/or function, as determined by previous biochemical analysis of our target molecules and analysis using Conserved Domain Database (https://www.ncbi.nlm.nih.gov/cdd/).

3. Results and Discussion

In order to examine the sequence variability in our *F. hepatica* vaccine candidate genes (FhLAP, FhPrx, FhHDM and FhCL1), we analysed samples included in the *F. hepatica* BioBank from 87 individual adult parasites obtained from naturally and experimentally infected cattle and sheep isolated in Uruguay, Argentina, Belgium, Ireland and UK (laboratory; from sheep infected in Spain).

We detected low levels of variation for all the target genes. Almost half of the variations detected were homozygous sites found in all the samples analysed. The conservation at these positions, despite the diverse geographical origins of the samples and the fact that they are shared also by the two available genome drafts, strongly suggest that they represent errors from invalid base calling in the reference sequences available in GenBank rather than true variable sites.

Analysis of the 1,575bp of the FhLAP coding sequence identified 14 variable sites. Four homozygous variants were observed compared with the reference sequence (Acosta et al., 2008) in all the individuals sequenced showing good quality sequences (see Supplementary file 2) and in the corresponding sequence in both genomic drafts; namely three non-synonymous changes – at base

pairs C1210G (Pro to Ala), C1337T (Thr to Ile), A1489G (Thr to Ala), and one synonymous variation at position 78 (C78A, Table 4, Figure 1). Several geographical differences were observed for the remaining 10 variable sites. In particular, three fluke samples from Uruguay were heterozygous at position 93 (G/A) in contrast to the sequences amplified from the rest of the samples and the corresponding genome sequences, where a synonymous change (G to A) was observed (Table 4). A change at position 1464 was detected in a single adult parasite from Uruguay. Interestingly, we observed 3 heterozygous variants at positions A1039C, G1057T, A1166T in one of the Belgian field isolates. Specifically the variation at nucleotide position 1057 results in an amino acid change from Ala to Ser situated between 2 metal co-ordination residues adjacent to the active site that may subsequently have an effect on protein function (Acosta et al. 2008).

Only 6 variants were observed throughout the FhPrx sequence (Table 5, Supplementary file 3). Three of them correspond to non-synonymous variations observed in one individual fluke from Ireland (T227A and A234C, resulting in Phe to Ile and Asp to Ala changes, respectively) and in one Belgian fluke that has a Val to Glu change due to the substitution T498A. These variable sites were not at positions expected to affect the active site residues important for peroxiredoxin function (McGonigle et al 1997). The remaining three changes were synonymous and showed variable amounts of homozygous and heterozygous individuals in samples from different locations. Interestingly, geographical differences can be detected at some positions; for example nucleotides at position 82 are predominantly homozygous in the Irish samples with thymine being more frequent (82T; 23/33), while the UK samples preferably display a homozygous cytosine (82C; 9/12 samples). Samples from Belgium show a predominance of heterozygotes in this position, while in Uruguay three possible genotypes at this site are detected (Table 5 and Supplementary file 2).

A more complex scenario pertains for FhCL1 as six genes in the *F. hepatica* genome encode isoforms of this protease (Cwiklinski et al., 2015, McNulty et al 2017). The similarity between these six genes does not allow for the design of PCR primers that can differentiate between them. As such, the FhCL1 primers used in this study amplify at least three of the FhCL1 genes, which

188 makes defining SNPs in the sequences challenging (Table 6, Supplementary file 4). Consequently,
189 the variants that might exist in each of them are confounded by the different FhCL1 sequences.
190 Clear evidence of this is the lower quality of the FhCL1 sequences, since they represent a pool of
191 highly similar but divergent molecules from each individual. In addition, 45 potential variants were
192 observed, interspersed throughout the gene sequence. This number of variants is in contrast to the
193 variation observed in genes of similar size analyzed as part of this study, namely FhLAP and
194 FhCL2, that show lower levels of variation. Nine positions (495, 602, 624, 663, 666, 697, 715, 804,
195 911) differ from the FhCL1 originally described (Roche et al., 1997) but are homozygous in all the
196 samples analyzed despite their geographical origin and the available genomic sequences, suggesting
197 base call errors in the original sequences. The remaining 36 variations might correspond to either
198 true heterozygous sites or variant sites between FhCL1 sequences. While changes at 13 of these
199 positions do not alter the resulting amino acid, non-synonymous substitutions occur at the
200 remaining 23 variable positions. Only one of the non-synonymous changes (V267L) corresponds to
201 an amino acid position potentially important for the CL1 functionality (the S2 pocket). In fact while
202 Val is present in FhCL1 at this position, the presence of Leu is characteristic of FhCL2. However, a
203 recent study demonstrates that a single change in this position generated by site directed
204 mutagenesis does not affect the specificity of the enzyme (Corvo et al, 2018). These observations
205 reinforce the idea that a diversity of FhCL1 isoforms with subtle differences might be produced
206 simultaneously by the parasite. The only experimental methods to further explore the complexity
207 and diversity of this *Fasciola* cathepsin L subfamily would be to consider digital PCR sequencing
208 or next generation sequencing.

209 Since assessing the variability of the FhCL1 vaccine candidate was challenging due to the
210 presence of several genes, we evaluated the potential for variation within the complex *F. hepatica*
211 cathepsin gene family by analyzing a related protease of the same family, FhCL2, which is encoded
212 by a single copy gene. In contrast to that found for FhCL1, 12 changes were observed throughout
213 the FhCL2 sequence in comparison to the reference sequence (Table 7, Supplementary file 5). Four

variants, comprised of non-synonymous changes at positions 41, 302, 490 and 501, were found in both of the *F. hepatica* genome sequences and in all the individuals analysed showing good quality sequences (see Supplementary file 2), indicating only variation compared with the reference sequence. The remaining variants were represented by six non-synonymous changes and two silent mutations at positions 273 and 378 (G273A, Ser; G378A, Gln) (Table 7). The synonymous change at position 378 within four adult parasites obtained in Ireland is of particular note, as this position corresponds to the amino acid residue that is part of the active site S1 binding subsite and therefore a non-synonymous change at this position may have affected the enzymatic activity of the protein. A predominance of the heterozygote state (A/G) at two base pair positions (273 and 373) was also observed in most of individuals analysed, which in the case of position 373 that corresponds to the first position of the codon resulted in two amino acid variants, namely Asn and Asp. In addition, we observed variations that were specific to geographical locations. Specifically, variations at positions 649 and 670 were only been found in the Irish and UK (Ridgeway) isolates, indicating this may be a variation restricted to European fluke isolates. In contrast, variability at position 259 was only observed in the samples from Uruguay and Argentina suggesting a distribution restricted to South America.

To date our analysis of the FhHDM gene in samples from South America has detected six variations within the intron sequences and two synonymous substitutions in the coding sequence that consequently do not affect the protein sequence (Table 8, Supplementary file 6). Further analysis of other isolates is required to determine whether this high level of conservation is observed across all *F. hepatica* isolates. As the molecular function of FhHDM relies on its distinct secondary protein structure that includes an amphipathic α-helix (Robinson et al., 2011), and is encoded by a single copy gene, the lack of variation we observe indicates that the function of this molecule needs to be maintained by maintaining the secondary structure of this molecule.

Antigenic polymorphism is considered an important mechanism of immune evasion used by a large number of pathogens, including helminths (Maizels and Kurniawan-Atmadja, 2002). In

addition, when allelic diversity affects immune targets, it might pose a challenge for vaccine approaches. As a result, the development of multivalent vaccines may require the incorporation of allelic variants in order to cover most of the isolates/strains responsible for an infection. This type of approach has recently been used for the development of a vaccine against malaria (Terheggen et al., 2014). Similarly, *Haemonchus contortus* vaccination studies using cathepsin B-like cysteine proteases have identified parasite populations with defined allelic profiles suggesting that immunisation imposes a genetic selection on these genes and that specific alleles may play a immunoprotective role (Martin et al., 2015).

Our study highlights a low level of allelic variability within prime vaccine candidate gene targets (FhCL1, FhPrx, FhHDM and FhLAP) in geographically dispersed *F. hepatica* isolates as well as the laboratory maintained isolate used in our sheep vaccine trials. These results suggest that the variability among vaccine studies (see Table 1) are not related to heterogeneity in these genes. As such host-related effects and the level of animal susceptibility status to *F. hepatica* infection should be investigated in the future that may provide a more plausible explanation for the variability observed in *F. hepatica* vaccine trials. Adjuvant selection is also an important issue for liver fluke vaccine trials. There is a need to find an adjuvant as potent as Freund's Complete Adjuvant, without the side effects, that will elicit appropriate potent immune responses in a broad range of hosts to overcome animal variability. Furthermore, to date native fluke antigens have been more effective than recombinant vaccines, which may imply that post-translational modifications such as glycosylation may be important in eliciting protective immune responses, which needs to be taken into consideration.

To date, investigation of vaccine candidate antigen variability in trematodes has only been previously examined in schistosomes (Gleichsner et al., 2015). Consistent with our study of *F. hepatica*, analysis of the cathepsin B sequence in *Schistosoma mansoni* also identified low levels of that most probably will not impact on protein-antibody interaction and binding (Simões et al., 2007). Studies of candidate antigens tetraspanins both in *S.mansoni* and *S.japonicum* show more

variability, however it should be taken into account that these membrane proteins constitute a large gene family, so it is not clear if the variability is allelic or due to multiple genes, similarly to what we found in FhCL1 (Cupit et al., 2011; Young et al., 2015; Zhang et al., 2011). Taken together these data suggest that in spite of the variability that might be detected at the genomic level, trematodes show conservation in the sequence of antigens important for key biological functions. This limited variability reinforces the idea that the selected proteins could be good immunogens for liver fluke worldwide. In addition to our study of antigen variability, we have sourced 87 *F. hepatica* isolates that are now housed in a publically available *F. hepatica* biobank that provides the liver fluke research community with access to a variety of isolates, which in particular are suitable for ongoing population genetic studies of this parasite. This biobank can be accessed upon request to M. Siles-Lucas at IRNASA-CSIC and is open to new submissions.

Acknowledgements

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 635408 (PARAGONE; https://www.paragoneh2020.eu/). MSL and JGM acknowledge the financial support from the Ministry MINECO, Spain (Project nb. AGL2015-67023-C2-2-R). Thanks are specially given to Prof. José Perez and Prof. Alvaro Martínez (Veterinary Faculty, University of Córdoba, Spain), Prof. Peter Geldhof (Faculty of Veterinary Medicine, Ghent University, Belgium), and Prof. Grace Mulcahy and Dr. Laura Garza-Cuartero (School of Veterinary Medicine, University College Dublin, Republic of Ireland) for providing the parasite isolates used in this study. Thanks are also given to María González for technical support.

292 **References**

293 Acosta, D., Cancela, M., Piacenza, L., Roche, L., Carmona, C., Tort, J.F., 2008. *Fasciola hepatica*
294 leucine aminopeptidase, a promising candidate for vaccination against ruminant fasciolosis.
295 Mol. Biochem. Parasitol. 158, 52–64.

296 Beesley, N.J., Williams, D.J.L., Paterson, S., Hodgkinson, J., 2017. *Fasciola hepatica* demonstrates
297 high levels of genetic diversity, a lack of population structure and high gene flow: possible
298 implications for drug resistance. 47, 11–20.

299 Buffoni, L., Martínez-Moreno, F.J., Zafra, R., Mendes, R.E., Pérez-Écija, A., Sekiya, M., Mulcahy,
300 G., Pérez, J., Martínez-Moreno, A., 2012. Humoral immune response in goats immunised with
301 cathepsin L1, peroxiredoxin and Sm14 antigen and experimentally challenged with *Fasciola*
302 *hepatica*. Vet. Parasitol. 185, 315–321.

303 Carmona, C., Tort, J.F., 2017. Fasciolosis in South America: epidemiology and control challenges.
304 J. Helminthol. 91, 99-109.

305 Cupit, P.M., Steinauer, M.L., Tonnessen, B.W., Eric Agola, L., Kinuthia, J.M., Mwangi, I.N.,
306 Mutuku, M.W., Mkoji, G.M., Loker, E.S., Cunningham, C., 2011. Polymorphism associated
307 with the *Schistosoma mansoni* tetraspanin-2 gene. Int. J. Parasitol. 41, 1249–1252.

308 Cwiklinski, K., O'Neill, S.M., Donnelly, S., Dalton, J.P., 2016.A prospective view of animal and
309 human fasciolosis. Parasite Immunol.38, 558-568.

310 Cwiklinski, K., Dalton, J.P., Dufresne, P.J., La Course, J., Williams, D.J., Hodgkinson, J., Paterson,
311 S., 2015. The *Fasciola hepatica* genome: gene duplication and polymorphism reveals
312 adaptation to the host environment and the capacity for rapid evolution. Genome Biol. 16, 71.

313 Corvo, I., Ferraro, F., Merlino, A., Zuberbuhler, K., O'Donoghue, A.J., Pastro, L., Pi-Denis, N.,
314 Basika, T., Roche, L., McKerrow, J.H., Craik, C.S., Caffrey, C.R., Tort, J.F., 2018. Substrate

315 Specificity of Cysteine Proteases Beyond the S2 Pocket: Mutagenesis and Molecular
316 Dynamics Investigation of *Fasciola hepatica* Cathepsins L. Front. Mol. Biosci. 5, 40.

317 Dalton, J.P., McGonigle, S., Rolph, T.P., Andrews, S.J., 1996. Induction of Protective Immunity in
318 Cattle against Infection with *Fasciola hepatica* by Vaccination with Cathepsin L Proteinases
319 and with Hemoglobin. Infect. Immun. 64, 5066–5074.

320 Dalton, J.P., Robinson, M.W., Mulcahy, G., O'Neill, S.M., Donnelly, S., 2013. Immunomodulatory
321 molecules of *Fasciola hepatica*: Candidates for both vaccine and immunotherapeutic
322 development. Vet. Parasitol. 195, 272–285.

323 Gleichsner, A.M., Thiele, E.A., Minchella, D.J., 2015. It's all about those bases: The need for
324 incorporating parasite genetic heterogeneity into the development of schistosome vaccines.
325 PLoS Negl. Trop. Dis. 9, 1–5.

326 Golden, O., Flynn, R.J., Read, C., Sekiya, M., Donnelly, S.M., Stack, C., Dalton, J.P., Mulcahy, G.,
327 2010. Protection of cattle against a natural infection of *Fasciola hepatica* by vaccination with
328 recombinant cathepsin L1 (rFhCL1). Vaccine 28, 5551–5557.

329 Kelley, J.M., Elliott, T.P., Beddoe, T., Anderson, G., Skuce, P., Spithill, T.W., 2016. Current Threat
330 of Triclabendazole Resistance in *Fasciola hepatica*. Trends Parasitol. 32, 458–469.

331 Maggioli, G., Acosta, D., Silveira, F., Rossi, S., Giacaman, S., Basika, T., Gayo, V., Rosadilla, D.,
332 Roche, L., Tort, J., Carmona, C., 2011. The recombinant gut-associated M17 leucine
333 aminopeptidase in combination with different adjuvants confers a high level of protection
334 against *Fasciola hepatica* infection in sheep. Vaccine 29, 9057–9063.

335 Maizels, R.M., Kurniawan-Atmadja, A., 2002.Variation and polymorphism in helminth parasites.
336 Parasitology 125, S25-37.

337 Martín, S., Molina, J.M., Hernández, Y.I., Ferrer, O., Muñoz, M.C., López, A., Ortega, L., Ruiz, A.,
338 2015. Influence of immunoprotection on genetic variability of cysteine proteinases from
339 *Haemonchus contortus* adult worms.Int. J. Parasitol. 45, 831-840.

340 Martínez-Sernández, V., Perteguer, M.J., Mezo, M., González-Warleta, M., Gárate, T., Valero,
341 M.A., Ubeira, F.M., 2017. *Fasciola* spp: Mapping of the MF6 epitope and antigenic analysis
342 of the MF6p/HDM family of heme-binding proteins. PLoS One 12, e0188520.

343 McGonigle, S., Curley, G.P, Dalton J.P., 1997. Cloning of peroxiredoxin, a novel antioxidant
344 enzyme, from the helminth parasite *Fasciola hepatica*. Parasitology 115, 101-104.

345 McNulty, S.N., Tort, J.F., Rinaldi, G., Fischer, K., Rosa, B.A., Smircich, P., Fontenla, S., Choi,
346 Y.J., Tyagi, R., Hallsworth-Pepin, K., Mann, V.H., Kammili, L., Latham, P.S., Dell'Oca, N.,
347 Dominguez, F., Carmona, C., Fischer, P.U., Brindley, P.J., Mitreva, M., 2017. Genomes of
348 *Fasciola hepatica* from the Americas Reveal Colonization with *Neorickettsia* Endobacteria
349 Related to the Agents of Potomac Horse and Human Sennetsu Fevers. PLoS Genet. 13,
350 e1006537.

351 Mehmood, K., Zhang, H., Sabir, A.J., Abbas, R.Z., Ijaz, M., Durrani, A.Z., Saleem, M.H., Ur
352 Rehman, M., Iqbal, M.K., Wang, Y., Ahmad, H.I., Abbas, T., Hussain, R., Ghori, M.T., Ali,
353 S., Khan, A.U., Li, J., 2017.A review on epidemiology, global prevalence and economical
354 losses of fasciolosis in ruminants. Microb. Pathog. 109, 253-262.

355 Mendes, R.E., Pérez-Écija, R.A., Zafra, R., Buffoni, L., Martínez-Moreno, Á., Dalton, J.P.,
356 Mulcahy, G., Pérez, J., 2010. Evaluation of hepatic changes and local and systemic immune
357 responses in goats immunized with recombinant Peroxiredoxin (Prx) and challenged with
358 *Fasciola hepatica*. Vaccine 28, 2832–2840.

359 Mulcahy, G., O'Connor, F., McGonigle, S., Dowd, A., Clery, D.G., Andrews, S.J., Dalton, J.P.,
360 1998. Correlation of specific antibody titre and avidity with protection in cattle immunized
361 against *Fasciola hepatica*. Vaccine 16, 932–939.

362 Mulcahy, G., O'Connor, F., Clery, D., Hogan, S.F., Dowd, A.J., Andrews, S.J., Dalton, J.P., 1999.
363 Immune responses of cattle to experimental anti-*Fasciola hepatica* vaccines. Res. Vet. Sci. 67,
364 27–33.

365 Orbegozo-Medina, R.A., Martínez-Sernández, V., González-Warleta, M., Castro-Hermida, J.A.,
366 Mezo, M., Ubeira, F.M., 2018. Vaccination of sheep with Quil-A® adjuvant expands the
367 antibody repertoire to the *Fasciola* MF6p/FhHDM-1 antigen and administered together impair
368 the growth and antigen release of flukes. Vaccine 36, 1949-1957.

369 Pacheco, I.L., Abril, N., Morales-Prieto, N., Bautista, M.J., Zafra, R., Escamilla, A., Ruiz, M.T.,
370 Martínez-Moreno, A., Pérez, J., 2017. Th1/Th2 balance in the liver and hepatic lymph nodes of
371 vaccinated and unvaccinated sheep during acute stages of infection with *Fasciola hepatica*.
372 Vet. Parasitol. 238, 61–65.

373 Pérez-Ecija, R.A., Mendes, R.E., Zafra, R., Buffonni, L., Martínez-Moreno, A., Pérez, J., 2010.
374 Pathological and parasitological protection in goats immunised with recombinant cathepsin L1
375 and challenged with *Fasciola hepatica*. Vet. J. 185, 351–353.

376 Piacenza, L.A., Acosta, D., Basmadjian, I., Dalton, J.P., Carmona, C., 1999. Vaccination with
377 Cathepsin L Proteinases and with Leucine Aminopeptidase Induces High Levels of Protection
378 against Fascioliasis in Sheep. Infect. Immun. 67, 1954–1961.

379 Robinson, M.W., Donnelly, S., Hutchinson, A.T., To, J., Taylor, N.L., Norton, R.S., Perugini,
380 M.A., Dalton, J.P., 2011. A family of helminth molecules that modulate innate cell responses
381 via molecular mimicry of host antimicrobial peptides. PLoS Pathog. 7, e1002042.

382 Roche, L., Dowd, A.J., Tort, J., McGonigle, S., MacSweeney, A., Curley, G.P., Ryan, T., Dalton,
383 J.P., 1997. Functional expression of *Fasciola hepatica* cathepsin L1 in *Saccharomyces*
384 *cerevisiae*. Eur. J. Biochem. 245, 373-380.

385 Simões, M., Bahia, D., Zerlotini, A., Torres, K., Artiguenave, F., Neshich, G., Kuser, P., Oliveira,
386 G., 2007. Single nucleotide polymorphisms identification in expressed genes of *Schistosoma*
387 *mansoni*. Mol. Biochem. Parasitol. 154, 134–140.

388 Spithill, T.W., Carmona, C., Piedrafita, D., Smooker, P.M., 2012. Prospects for immunoprophylaxis
389 against *Fasciola hepatica* (Liver Fluke). In: Caffrey, C.R. (Ed.), Parasitic Helminths: Targets,
390 Screens, Drugs, and Vaccines, First ed. Wiley-VCH Verlag GmbH and Co., KGaA, Weinheim,
391 pp. 467–486.

392 Teofanova, D., Hristov, P., Yoveva, A., Radoslavov, G., 2012. Issues Associated with Genetic
393 Diversity Studies of the Liver Fluke, *Fasciola hepatica* (Platyhelminthes, Digenea,
394 Fasciolidae), in: Caliskan, M. (Ed.), Genetic Diversity in Microorganisms. InTech, pp. 251–
395 274.

396 Terheggen, U., Drew, D.R., Hodder, A.N., Cross, N.J., Mugyenyi, C.K., Barry, A.E., Anders, R.F.,
397 Dutta, S., Osier, F.H., Elliott, S.R., Senn, N., Stanisic, D.I., Marsh, K., Siba, P.M., Mueller, I.,
398 Richards, J.S., Beeson, J.G., 2014. Limited antigenic diversity of *Plasmodium falciparum*
399 apical membrane antigen 1 supports the development of effective multi-allele vaccines. BMC
400 Med. 12, 183.

401 Toet, H., Piedrafita, D.M., Spithill, T.W., 2014. Liver fluke vaccines in ruminants: strategies,
402 progress and future opportunities. Int. J. Parasitol. 44, 915–927.

403 Villa-Mancera, A., Reynoso-Palomar, A., Utrera-Quintana, F., Carreón-Luna, L., 2014. Cathepsin
404 L1 mimotopes with adjuvant QuilA induces a Th1/Th2 immune response and confers

405 significant protection against *Fasciola hepatica* infection in goats. Parasitol. Res. 113, 243–
406 250.

407 Wit, J., Gilleard, J.S., 2017. Re-sequencing helminth genomes for population and genetic studies.
408 Trends Parasitol. 33, 388-399.

409 Young, N.D., Chan, K.-G., Korhonen, P.K., Min Chong, T., Ee, R., Mohandas, N., Koehler, A. V.,
410 Lim, Y.-L., Hofmann, A., Jex, A.R., Qian, B., Chilton, N.B., Gobert, G.N., McManus, D.P.,
411 Tan, P., Webster, B.L., Rollinson, D., Gasser, R.B., 2015. Exploring molecular variation in
412 *Schistosoma japonicum* in China. Sci. Rep. 5, 17345.

413 Zafra, R., Pérez-Ecija, R.A., Buffoni, L., Moreno, P., Bautista, M.J., Martínez-Moreno, A.,
414 Mulcahy, G., Dalton, J.P., Pérez, J., 2013. Experimentally induced disease Early and Late
415 Peritoneal and Hepatic Changes in Goats Immunized with Recombinant Cathepsin L1 and
416 Infected with *Fasciola hepatica*. J. Comp. Pathol. 148, 373–384.

417 Zhang, W., Li, J., Duke, M., Jones, M.K., Kuang, L., Zhang, J., Blair, D., Li, Y., McManus, D.P.,
418 2011. Inconsistent protective efficacy and marked polymorphism limits the value of
419 *Schistosoma japonicum* tetraspanin-2 as a vaccine target. PLoS Negl. Trop. Dis. 5, e1166.

420 Zintl, A., Talavera, S., Sacchi-Nestor, C., Ryan, M., Chryssafidis, A., Mulcahy, G., 2015.
421 Comparison of *Fasciola hepatica* genotypes in relation to their ability to establish patent
422 infections in the final host. Vet. Parasitol. 210, 145-150.

Figure legends

Figure 1. SNPs variants detected along the region amplified of the FhLAP gene. The C-terminal catalytic domain (Interpro IPR000819) is highlighted in light green. The active site residues are shown in the boxes and the residues involved in metal binding are shown above the schematic, indicated by the lines. The variable positions in different geographical populations of liver fluke are shown at the bottom of the figure. The numbers in brackets indicate number of heterozygote individuals.

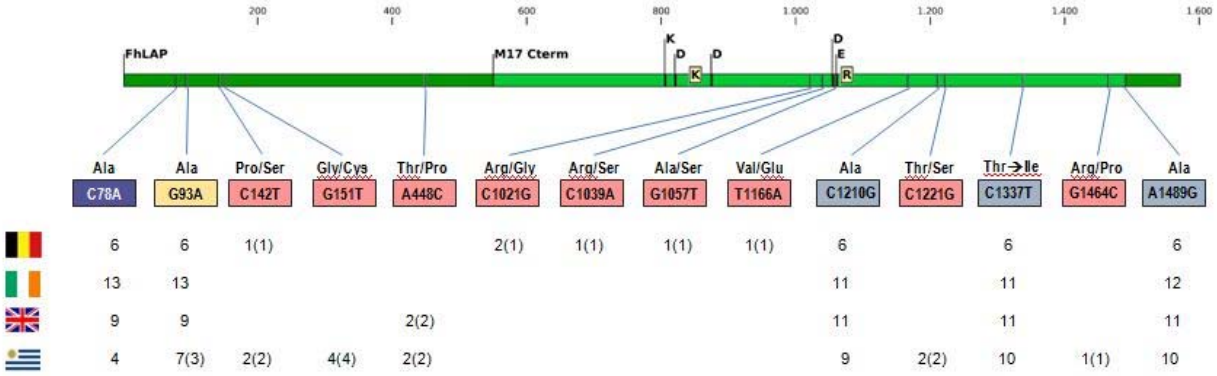

Table 1. Efficacy of single or combination vaccines against *Fasciola hepatica.*

Antigen	Source	Host	Schedule[1]	Adjuvant	Efficacy[2]	Reference
FhCL1	Adult E/S	Cattle	10-500 µg X 3	FCA/FIA	38.2-69.5%	Dalton et al., 1996
			200 µg X 3		42.5%	
		Sheep	100 µg X 2		33%	Piacenza et al., 1999
	Recombinant	Cattle	200 µg X 2	Montanide[3]	47.2%	Golden et al., 2010
				Montanide[4]	49.2%	
		Goat	100 µg X 2	Quil A	0%	Pérez-Ecija et al., 2010
					38.7%	Buffoni et al., 2012
					0%	Zafra et al., 2013
				Montanide[3]	0%	Pacheco et al., 2017
	Mimotope	Goat	1×10^{13}pp	Quil A	46.9-79.5 %	Villa-Mancera et al., 2014
FhCL1+Hb	Adult E/S	Cattle	200 µg X 3	FCA/FIA	51.9%	Dalton et al., 1996
FhCL2		Sheep	100 µg X 2	FCA/FIA	34%	Piacenza et al., 1999
FhCL2+Hb		Cattle	200 µg X 3	FCA/FIA	72.4%	Dalton et al., 1996
				FCA/FIA	72.4%	Mulcahy et al., 1998
				FIA	11.2%	
				FCA/FIA	29%	Mulcahy et al., 1999
FhCL1+CL2		Sheep	100 µg X 2	FCA/FIA	60%	Piacenza et al., 1999
		Cattle	200 µg X 3	FCA/FIA	55%	Mulcahy et al., 1999
FhCL1+CL2 +LAP	Adult E/S, SOM	Sheep	100 µg X 2	FCA/FIA	79%	Piacenza et al., 1999
FhCL1+Prx +Sm14	Recombinant	Goat	100 µg X 2	Quil A	10.1%	Buffoni et al., 2012
FhLAP	Adult SOM	Sheep	100 µg X 2	FCA/FIA	89.6%	Piacenza et al., 1999
	Recombinant	Rabbit	100 µg X 2	FCA/FIA	78%	Acosta et al., 2008
		Sheep	100 µg X 2	FCA/FIA	83.8%	Maggioli et al., 2011
				Alum	86.7%	
				Adyuvac 50	74.4%	
				DEAE-D	49.8%	
				Ribi	49.5%	
FhPrx	Recombinant	Goat	100 µg X 2	Quil A	33.1%	Mendes et al., 2010
					33.1%	Buffoni et al., 2012

Adult E/S or SOM, *F. hepatica* adult worm excreted/secreted or somatic products. FCA/FIA, Freund's complete/incomplete adjuvant. [1]Vaccination dose (µg, micrograms; pp, phage particles number) and boosts number; [2]Only referred to the reduction of worm numbers in vaccinated and infected animals, compared with infected and non-vaccinated animals.[3]Montanide ISA70VG. [4]Montanide ISA206VG.

Table 2. *Fasciola hepatica* isolates from the Biobank used in this study.

Country	Region	Isolate type	Host	Nb. of samples
Uruguay	Colonia	Field	Cattle	2
Uruguay	Florida	Field	Cattle	3
Uruguay	Montevideo	Field	Cattle	14[1]
Argentina	Córdoba	Field	Cattle	1
Ireland	East Ireland	Field	Cattle	34
Belgium	Lendelede Lapscheure Zuienkerke Nieuwmunster Aartrijke	Field	Cattle	6
Uruguay	Canelones	Experimental infection	Cattle	15[2]
UK	Laboratory	Experimental infection	Sheep	12[3]

[1]Samples used only for FhHDM amplification and sequencing. [2] Field isolate from Uruguay used to infect Polled Hereford-breed cattle as part of the PARAGONE vaccine trial using FhCL1, FhLAP, FhPRX and FhHDM in Montanide 61VG. [3] Laboratory maintained South Gloucester isolate (Ridgeway Research) used to infect Merino sheep-breed in Cordoba, Spain as part of the PARAGONE vaccine trial using FhCL1, FhLAP, FhPRX and FhHDM in Montanide 61VG.

Table 3. Primers used for amplification.

Primer	Sequence
FhCL1Fwd	5′-ATGCGATTATTCGTATTAGCCG-3′
FhCL1Rev	5′-TCACGGATATTGTGCCACC-3′
FhCL2Fwd	5′-ATGCGGTGCTTCGTATTAGC-3′
FhCL2Rev	5′-TCACGGAAATCGTGCCACC-3′
FhPrxFwd	5′-ATGTGTGATCGCGATCAGTGCTC-3′
FhPrxRev	5′-CTAGTTGGCTGAGGAGAAATATG-3′
FhLAPFwd	5′-ATGGCGGCGTTGGCTGTGGG-3′
FhLAPRev	5′-CTATTTGAATCCCAGTCGTGG-3′
FhHDMFwd	5’-GTCTTGCTGTGGTCCTTCTT-3′
FhHDMRev	5′-TTTCCCGCGTATTTCTCCAA-3′

Table 4. Sequence changes detected in the FhLAP coding sequence of the analysed *Fasciola hepatica* isolates, showing the number of heterozygote individuals in brackets.

FhLAP			**Geographical origin –NUTs code- and number of flukes**				
Nucleotide change	**Amino acid change**	**Change type***	**BE, 6**	**IE, 30**	**UK, 12**	**UY, 11**	**AR, 1**
C78A	26A	Synonymous	6	13	9	4	
G93A	31A	Synonymous	6	13	9	7 (3)	
C142T	P46S	Non-synonymous	1 (1)			2 (2)	
G151T	G51C	Non-synonymous				4 (4)	
A448C	T150P	Non-synonymous			2 (2)	2 (2)	
C1021G	R341G	Non-synonymous	2 (1)				
C1039A	R347S	Non-synonymous	1 (1)				
G1057T	A353S	Non-synonymous	1 (1)				
T1166A	V389E	Non-synonymous	1 (1)				
C1210G	P404A	Non-synonymous	6	11	11	9	
C1221G	T407S	Non-synonymous				2 (2)	
C1337T	T446I	Non-synonymous	6	11	11	10	
G1464C	R488P	Non-synonymous				1 (1)	
A1489G	T497A	Non-synonymous	6	12	11	10	

Table 5. Sequence changes detected in the FhPrx coding sequence of the analysed *Fasciola hepatica* isolates, showing the number of heterozygote individuals in brackets.

FhPrx			**Geographical origin –NUTs code- and number of flukes**				
Nucleotide change	**Amino acid change**	**Change type**	**BE, 6**	**IE,33**	**UK, 12**	**UY, 11**	**AR, 1**
C82T	30N	Synonymous	5 (4)	30 (7)	3 (3)	8 (5)	
T227A	F79I	Non-synonymous		1			
A234C	D81A	Non-synonymous		1			
A418G	142E	Synonymous	4 (4)	9 (8)	12 (3)	7 (7)	1 (1)
T498A	V169E	Non-synonymous	1				
A514G	174E	Synonymous	2 (2)	1 (1)		4 (3)	1 (1)

Table 6. Sequence changes detected in the FhCL1 coding sequence of the analysed *Fasciola hepatica* isolates, showing the number of heterozygote individuals in brackets.

FhCL1			Geographical origin –NUTs code- and number of flukes				
Nucleotide change	**Amino acid change**	**Change type**	**BE, 6**	**IE, 34**	**UK, 12**	**UY, 11**	**AR, 1**
A30G	10T	Synonymous	2 (2)	3 (1)			
A69G	23Q	Synonymous	5 (3)	23 (10)	12 (2)	9	
C87T	29N	Synonymous	6 (6)	24 (21)	11 (11)	8 (8)	1 (1)
C104A	A35D	Non-synonymous	1 (1)	10 (10)	3 (3)		
C112G	Q38E	Non-synonymous	5 (5)	20 (20)	5 (5)		
A136G	K46E	Non-synonymous	6 (5)	21 (17)	12	11 (11)	1 (1)
T153C	51I	Synonymous	6 (5)	27 (17)	12	11	1
T162C	54H	Synonymous	5 (5)	11 (11)	12 (12)	3 (3)	1 (1)
C257T	T86I	Non-synonymous		3 (3)	4 (4)	1 (1)	
T265C	S89P	Non-synonymous	5 (5)	20 (19)	6 (6)	5 (5)	1 (1)
G295A	V99I	Non-synonymous	5 (5)	25 (24)	12 (12)	8 (8)	1 (1)
T315C	105N	Synonymous	5 (5)	19 (19)		3 (3)	
A365G	E122G	Non-synonymous	5 (5)	14 (14)	2 (2)	4 (4)	
C390T	130G	Synonymous	6 (2)	27 (7)	12 (3)	11 (8)	1
G449A	R150K	Non-synonymous	2 (2)	7 (7)			
C492T	164S	Synonymous	6 (6)	18 (17)	12 (12)	11 (11)	1 (1)
C493G	R165G	Non-synonymous	6 (5)	21 (7)	12 (6)	11 (9)	1
A495T	R165G	Non-synonymous	6	24	12	11	1
G517A	G173S	Non-synonymous	6 (6)	21 (11)	12 (12)	11 (11)	1 (1)
G602A	G202E	Non-synonymous	6	19	12	11	1
A623G	K208R	Non-synonymous	5 (5)	11 (8)	12 (12)	7 (7)	1 (1)
A624G	K208R	Non-synonymous	6	19	12	11	1
A630G	210L	Synonymous	4 (4)	17 (14)	12 (12)	9	1 (1)
T657C	219Y	Synonymous	6 (6)	18 (16)	12 (12)	11 (11)	1 (1)
T663G	221V	Synonymous	6	18	12	11	1
A666T	Q222H	Non-synonymous	6	17	12	11	1
G670T	G224T	Non-synonymous	3 (3)	4 (4)	5 (5)		
C694T	232L	Synonymous		6 (6)	11 (11)	1 (1)	
A697G	I233V	Non-synonymous	6	17	12	11	1
T703G	S235A	Non-synonymous	6 (6)	17 (15)	12 (5)	11 (11)	1 (1)
T715G	S239A	Non-synonymous	6	17	12	11	1
G721A	V241I	Non-synonymous	6 (5)	8 (8)		5 (5)	1 (1)
G760A	S254G	Non-synonymous	2 (2)	10 (10)	12 (12)	11 (11)	1 (1)
C788T	S263L	Non-synonymous	4 (4)	12 (12)	10 (10)	11 (11)	1 (1)
C793T	L265F	Non-synonymous	3 (3)	8 (6)	5 (5)	5 (5)	1 (1)
C796G	R266G	Non-synonymous	6 (5)	17 (13)	12 (12)	11 (11)	1 (1)
G799T	V267L	Non-synonymous	6 (5)	17 (16)	12 (12)	11 (11)	1 (1)
C804T	268N	Synonymous	6	17	12	11	1

T828C	276Y	Synonymous	3 (3)	14 (14)	12 (12)	11 (11)	1 (1)
G839A	G280D	Non-synonymous	5 (5)	15 (15)	12 (12)	11 (11)	1 (1)
T878C	L293S	Non-synonymous	4 (4)	11 (9)	12 (12)	11 (11)	1 (1)
C881A	S294Y	Non-synonymous	4 (4)	2 (2)	11 (11)	10 (10)	1 (1)
G882C	S294Y	Non-synonymous	3 (3)	1 (1)	11 (10)	11 (10)	
T911C	V304A	Non-synonymous	6	17	12	11	
C918T	366N	Synonymous	5 (3)	17 (17)	12 (7)	7	

Table 7. Sequence changes detected in the FhCL2 coding sequence of the analysed *Fasciola hepatica* isolates, showing the number of heterozygote individuals in brackets.

FhCL2			**Geographical origin –NUTs code- and number of flukes**				
Nucleotide change	**Amino acid change**	**Change type**	**BE, 6**	**IE, 34**	**UK, 12**	**UY,11**	**AR, 1**
A41T	Y14F	Non-synonymous	5	15	10	8	
G259A	E87K	Non-synonymous				3 (3)	1 (1)
G273A	91S	Synonymous	3 (3)	29 (29)	9 (9)	6 (5)	
T302A	F101Y	Non-synonymous	6	29	11	11	
T352G	Y118D	Non-synonymous		4 (4)			
A373G	N125D	Non-synonymous	6 (6)	28 (27)	11 (11)	8 (8)	
G378A	126Q	Synonymous		4			
C490A	P164T	Non-synonymous	6	29	11	11	1
G501T	L167F	Non-synonymous	6	29	11	11	1
G649T	G217C	Non-synonymous		5 (5)	4 (4)		
G670T	G224C	Non-synonymous		5 (5)	5 (5)		
G750C	M250I	Non-synonymous		1 (1)			

Table 8. Sequence changes detected in the FhHDM coding sequence of the analysed *Fasciola hepatica* isolates, showing the number of heterozygote individuals in brackets.

FhHDM			Geographical origin – NUTs code- and number of flukes	
Nucleotide change	**Amino acid change**	**Change type**	**UY (34)**	**AR(1)**
G13T		Intron 1	6 (4)	
C26T		Intron 1	2 (2)	
T35C		Intron 1	14 (11)	
A66G	51A	Synonymous	4 (4)	
A105C	63L	Synonymous	12 (9)	
A25G		Intron 4	11 (10)	
A37G		Intron 4	11 (8)	
T47C		Intron 4	6 (4)	