Veterinary Parasitology 275 (2019) 108933

Contents lists available at ScienceDirect

Veterinary Parasitology

journal homepage: www.elsevier.com/locate/vetpar

Research paper

Validation of ITS-2 rDNA nemabiome sequencing for ovine gastrointestinal nematodes and its application to a large scale survey of UK sheep farms

Elizabeth Redman[a], Camila Queiroz[a], David J. Bartley[b], Michel Levy[a], Russell W. Avramenko[a], John Stuart Gilleard[a,*]

[a] Department of Comparative Biology and Experimental Medicine, Host Parasite Interactions Program, Faculty of Veterinary Medicine, 3330, Hospital Drive, University of Calgary, Calgary, Alberta, T2N 4N1 Canada
[b] Moredun Research Institute, Pentlands Science Park, Bush Loan, Penicuik, EH26 0PZ, United Kingdom

ARTICLE INFO

Keywords:
Nemabiome
Deep-sequencing
Amplicon assay
ITS-2
PCR bias
Correction factors
Sheep

ABSTRACT

We have validated ITS-2 rDNA nemabiome next-generation amplicon sequencing to determine relative species abundance of gastrointestinal nematode species in ovine fecal samples. In order to determine species representation biases, ITS-2 rDNA amplicon sequencing was applied to mock communities or field populations with known proportions of L3 for eight of the major ovine gastrointestinal nematode species: *Teladorsagia circumcincta*, *Trichostrongylus vitrinus, Haemonchus contortus, Cooperia curticei*, *Trichostrongylus axei*, *Trichostrongylus colubriformis*, *Chabertia ovina* and *Oesophagostumum venulosum*. Correction factors, calculated from this data, were shown to reduce species representation biases when applied to an independent set of field samples of known composition. We compared ITS-2 rDNA amplicon sequencing data that was generated from harvested eggs, freshly hatched L1 or L3 larvae following fecal culture and no statistically significant differences were found for the more abundant parasite species. We then applied the validated ITS-2 rDNA nemabiome amplicon sequencing assay to a set of archived L1 gastrointestinal nematode populations, collected in 2008 from fecal samples from 93 groups of 20 ewes and 61 groups of 20 lambs derived from 99 UK sheep farms. The presence of the major gastrointestinal nematode species had previously been determined on this large sample set by species-specific PCR. We show how the ITS-2rDNA amplicon sequencing data provided much more detailed information on species abundance than the previous species-specific PCR. This new data represents the most comprehensive overview of the relative abundance of the major gastrointestinal nematode species across UK sheep farms to date. Substantial variance in the relative abundance of both *T. circumcincta* and *T. vitrinus* between farms was revealed with the former species being of statistically significantly higher abundance in all three regions sampled (England, Scotland and Wales). The data also revealed that the relative abundance of *T. circumcinta* was statistically significantly higher in ewes than in lambs with the opposite pattern being the case for *T. vitrinus*. The nemabiome sequencing data also clearly illustrated the sporadic nature and skewed distribution of *H. controtus* across UK sheep farms as well as a higher relative abundance on farms from England compared to Wales and Scotland. The nemabiome survey also provides the first widescale data on the relative abundance of the two major large intestinal nematodes *C. ovina* and *O. venulosum*. This work validates ITS-2 rDNA nemabiome sequencing for use in sheep and illustrates the power of the approach for large scale surveillance of ovine gastrointestinal nematodes.

1. Introduction

Gastrointestinal nematodes of grazing livestock generally exist in complex communities consisting of multiple co-infecting species (Vlassoff, 1976; Agneessens et al., 1997; Giudici et al., 1999; Burgess et al., 2012; Stromberg et al., 2015). In the case of sheep, the majority of the most common, and most pathogenic, species belong to the strongyle group (Sargison et al., 2007; Besier et al., 2016; O'Connor et al., 2006) within the Clade V group of the nematode phylogeny (Blaxter and Koutsovoulos, 2015). Different species vary considerably in their pathogenicity, clinical presentation, epidemiology, inherent sensitivity to different drugs and drug resistance status (Sangster et al.,

* Corresponding author.
E-mail addresses: libbyredman@hotmail.co.uk (E. Redman), camila.dequeiroz@ucalgary.ca (C. Queiroz), dave.bartley@moredun.ac.uk (D.J. Bartley), mlevy@ucalgary.ca (M. Levy), rwavrame@ucalgary.ca (R.W. Avramenko), jsgillea@ucalgary.ca (J.S. Gilleard).

https://doi.org/10.1016/j.vetpar.2019.108933
Received 17 July 2019; Received in revised form 29 September 2019; Accepted 30 September 2019
0304-4017/ © 2019 Elsevier B.V. All rights reserved.

1980; Besier et al., 2016). Consequently, it is critical to be able to determine the prevalence and relative abundance of each individual species rather than treating them as a homogenous group. This need is ever more important given the emergence of anthelmintic drug resistance, the impact of climate change on parasite ranges, distributions and epidemiology and the increasing demand for evidence-based use of pharmaceuticals in parasite control. Routine diagnostics typically involve microscopic examination and counting of parasite eggs following their separation from fecal debris using flotation techniques. Strongle eggs are essentially indistinguishable by visual microscopic inspection between the different species and further identification is most commonly achieved by culturing fecal samples and morphologically examining L3 (van Wyk et al., 2004). Limited or overlapping morphological features/morphometrics of some species generally only allows identification to the genus level (McMurtry et al., 2000; van Wyk et al., 2004). This procedure can also be labour intensive, and because of the specialist technical experience required, is prone to error and operator variance (Gasser et al., 2008). Molecular methods such as PCR-RFLP (Gasser et al., 1994), real-time PCR (Learmont et al., 2009), multiplexed-tandem PCR (Roeber et al., 2011; Bisset et al., 2014) and ddPCR (Elmahalawy et al., 2018) assays are used in research studies but have varying degrees of accuracy, in terms of relative and/or absolute quantitation of species abundance, and are limited to those parasite species that are anticipated *a priori* and have primer sets and assays already developed and validated. We previously developed a new approach to determining the relative abundance of different species of strongylid gastrointestinal species of cattle and bison which we termed nemabiome sequencing (Avramenko et al., 2015, 2017 and 2018). This involves deep sequencing, using the Illumina Miseq platform, of the ITS-2 rDNA locus in an approach analogous to bacterial 16S rDNA sequencing commonly used in microbiome and metabarcoding studies of invertebrate communities.

In this paper, we describe the application of the nemabiome sequencing approach to ovine gastrointestinal communities. We determine and validate correction factors needed to compensate for species-specific amplification biases specifically for the major gastrointestinal nematode species found in sheep. We also adapt and validate the technique to use directly on eggs, hatched L1 larvae or cultured L3 larvae and show there is a high degree of agreement in the results whichever of these different parasites stages are used. Finally, we show the power of the approach in undertaking surveillance on the prevalence and relative infection intensities of the major ovine gastrointestinal nematode species using a large set of samples previously archived from UK sheep flocks. The nemabiome sequencing approach provides a new powerful tool for surveillance and diagnostics of sheep gastrointestinal nematodes and opens new research opportunities to study the interactions between complex parasite communities, their host and the environment.

2. Materials and methods

2.1. Pure strains and mock communities

Pure strains of the six most important GI nematode species of sheep were used to create mock communities of known species composition. The strains used were *H. contortus* MHco16_4881 (Moredun Institute), *T. circumcincta* TcRW_1 (Ridgeway Research), *T. colubriformis* MRC-Tcl-0114 (Merial), *T. axei* MTax_4731 (Moredun Institute), *T. vitrinus* MTvi1_5557 (Moredun Institute) and *C. curticei* CC_PFZ_1 (Pfizer via Moredun Institute). Ten different mixtures containing equal proportions of each of the six species were created. Each mixture was created by individually counting a total of 240 L3s into a single tube of lysis buffer (50 mM KCl, 10 mM Tris (pH 8.3), 2.5 mM $MgCl_2$, 0.45% Nonidet P-40, 0.45% Tween 20, 0.01% (w/v) gelatin) and proteinase K at 200 µg/ml (Redman et al., 2015), ensuring the species composition was accurately known. DNA lysate preparations were made using previously described techniques (Redman et al., 2008) and 4 µl of a 1:10 dilution of neat lysate was used as a PCR template.

2.2. Extraction and preparation of egg, L1 and L3 populations from the same fecal samples

48 independent sets of fresh fecal samples were obtained from 21 sheep farms in western Canada between June and August 2017. Each set comprised of fecal samples collected from 20 ewes collected by the producer and shipped to the laboratory in Ziploc bags, sealed having extinguished as much of the air as possible. Within 8 hours of arrival at the laboratory, equal amounts of feces from each individual ewe were pooled together and thoroughly manually homogenized to create a pooled fecal sample for each farm sample. Each pooled sample was then split in two: half was used for coproculture to generate L3s and the other used to extract eggs and culture L1s. The coproculture technique used was as described in Avramenko et al. (2015). To extract eggs, the fecal sample was washed through a course sieve (pore size: 1 mm), to remove large debris before a 13% salt solution (specific gravity 1.06) was added to the filtrate and centrifuged at 3600 RPM to float the eggs. The eggs were then washed in water on a 20 µm sieve to retain the eggs which were harvested by washing into 30 ml of water. The harvested eggs were counted and, if there were > 500eggs, half were fixed directly in 70% ethanol and the remainder were incubated at 22 °C for 24–48 hours to allow hatching to produce L1s. Only samples with an egg hatch rate of greater than 90% were used. Where the number of eggs extracted was < 500 all the eggs were allowed to hatch to L1s. For each farm sample, separate 100 µl DNA lysates were prepared from 200 eggs, 200 L1s and 200 L3s (or just L1 and L3 for those few farm samples where < 500 eggs were obtained). 4 µl of a 1:10 dilution of lysate was used as PCR template for nemabiome sequencing.

2.3. Archived L1 larvae from UK sheep farms

A survey of 119 UK sheep farms in 2008 has previously been described in detail (Burgess et al., 2012) and some of this material (61 lamb samples and 93 ewe samples from a total of 99 different farms) was used to trial the ovine ITS-2 rDNA nemabiome sequencing assay on a large sample set. Briefly, fresh fecal material was collected from either twenty ewes at lambing or twenty lambs at weaning from each farm, pooled together and used to culture L1s as previously described (Burgess et al., 2012). DNA lysate preparations were made from approximately 1000 L1s and 4 µl of a 1:40 dilution of lysate was used as a PCR template for nemabiome sequencing. Dilutions of several aliquots of lysate buffer made in parallel, were included as negative controls for all PCR amplifications.

2.4. Species identification of individual worms using standard species-specific PCR

90 individual L1 larvae per farm were lysed in a 20 µl volume of lysis buffer (see Section 2.1 for lysis buffer composition) and proteinase K at 200 µg/ml (Redman et al., 2015). This lysate was diluted 1:20 in water to be used as PCR template. Dilutions of several aliquots of lysate buffer, made in parallel, were included as negative controls for all PCR amplifications. Eight individual pairs of species-specific primers designed to amplify the most prevalent GI parasites of sheep (Supplementary Table S1) were used in simplex PCR assays to individually determine the species identity of the 90 L1 per farm and generate accurate data on the nematode species proportions in each farm sample. Conditions for the rDNA ITS-2 species-specific PCR assays were 94 °C for 2 min followed by 35 cycles of 92 °C for 30 s (the exception being *H. contortus* with 40 cycles); annealing for 30 s and 72 °C for 30 s with a final extension cycle at 72 °C for 10 min. PCR primers and annealing temperatures are given in Supplementary Table S1.

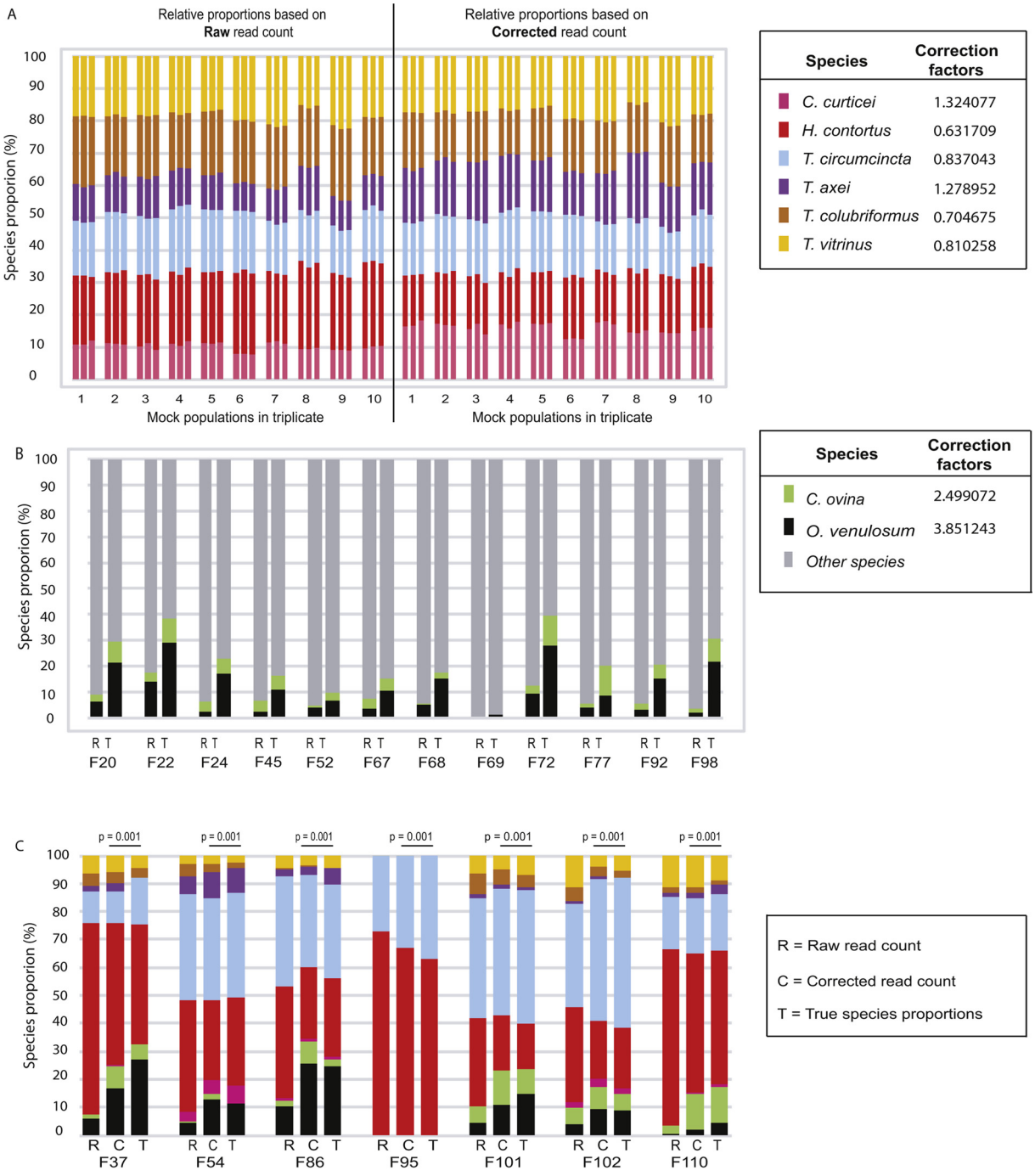


Fig. 1. Estimating and validating nemabiome representation bias correction factors for the major ovine gastrointestinal nematode species.

Panel A. Calculation of correction factors for the major abomasal and small intestinal nematode species: 10 mock populations each consisting of 240 L3 larvae (40 of each of six different species) were used to experimentally determine correction factors. Nemabiome ITS-2 rDNA sequencing was performed on each of the 10 mock populations in triplicate and the results of the species proportions, based on the raw uncorrected observed read counts is shown in the left-hand chart. For each mock population the observed species proportion was divided by the expected proportion (0.1667) and an average was generated from the ten mock populations to determine correction factors for each species (shown in the adjacent table). These species-specific correction factors were then applied to the observed read count for each species to produce "corrected" species proportions. These "corrected" results are shown in the right-hand chart.

Panel B. Calculation of correction factors for the two major large intestinal nematode species: The species of 90 individual L3 larvae from each of 12 field populations were individually identified using species-specific PCR to determine the true species proportions of *Chabertia ovina* and *Oesophagostomum venulosum*. The chart shows the species proportions in each of the 12 field populations calculated from the proportion of read counts of the nemabiome sequencing read counts (R = Raw read count data from nemabiome sequencing of 1000 larvae) compared to the true species proportions in the same population (T = true species proportions as determined by individual L3 PCR). The correction factors for the two large intestinal species calculated from this data is shown in the adjacent table.

Panel C. Validation of correction factors using field populations: The use of the calculated correction factors was validated using seven independent field populations. The chart compares the species proportions in the seven field populations; R = raw read counts from nemabiome sequencing of 1000 larvae, C = corrected Nemabiome read counts, T = true species composition determined from species-specific PCR of 90 individual larvae. No statistical difference between corrected and true species composition was observed based of Chi-square analysis ($p < 0.001$).

2.5. Deep amplicon sequencing of the rDNA ITS-2

The ITS-2 rDNA fragment was amplified and sequenced by deep amplicon sequencing as previously described (Avramenko et al., 2015). The library was sequenced using the MiSeq Illumina platform with the 2 × 250 v2 Reagent Kit (Illumina). An average read depth of 53,258 reads (range: 5,259–120,068) was obtained for each sample used in the analysis. A conservative read-depth threshold of 2000 reads per sample is routinely used for the Nemabiome assay and any samples with less than 2000 reads is considered a sequencing failure, discarded from further analysis and attempts are made to re-amplify and re-sequence the sample. All samples discussed here had a read-depth of > 2000 reads on the first sequencing attempt.

2.6. Bioinformatic analysis

Short-read sequence data was analysed using the Mothur bioinformatics tool version 1.36.1 (Schloss et al., 2009) as previously described (Avramenko et al., 2017, www.nemabiome.ca). Briefly, raw paired-end reads were assembled into single contigs and then filtered to remove contigs that were < 200bp or > 450bp. Contigs with ambiguities between the overlapping paired-end reads were also filtered out. An in-house curated database containing all available ITS-2 rDNA sequences from the relevant nematodes, created from a combination of our own reference sequences and those from public databases, was used in this analysis (Supplementary File S1). Contigs were aligned to the ITS-2 rDNA curated database and discarded if they did not align to at least 10% of any ITS-2 rDNA amplicon in the database with at least 90% sequence similarity. The remaining sequences were classified to reference sequences in the database using the k-nearest-neighbor method (k = 3) which is an option for the classify.seqs command (e.g. method = knn, numwanted = 3) in Mothur 1.36.1 (Schloss et al., 2009). When the three nearest matches did not match to a single species, the sequence was classified at the next highest taxonomic level. Further detailed information regarding the analysis pipeline can be viewed at www.nemabiome.ca.

2.7. Statistical analysis

The majority of statistics were calculated in IBM SPSS Statistics version 24. Chi-square analysis was used to compare the species proportions generated from the amplicon sequencing with the actual species proportions expected in a given sample and two-sample *t*-tests (2-tailed) were used to compare mean percentages of a species between groups of samples (i.e. ewe versus lamb samples). Lin's Concordance Correlation Coefficient was calculated in SPSS with the syntax provided by https://gjyp.nl/marta/Lin.sps. r_c values > 0.8 are classified as a near perfect agreement (r_c = 1, would represent perfect agreement).

3. Results

3.1. Assessment of species representation bias and the determination of correction factors

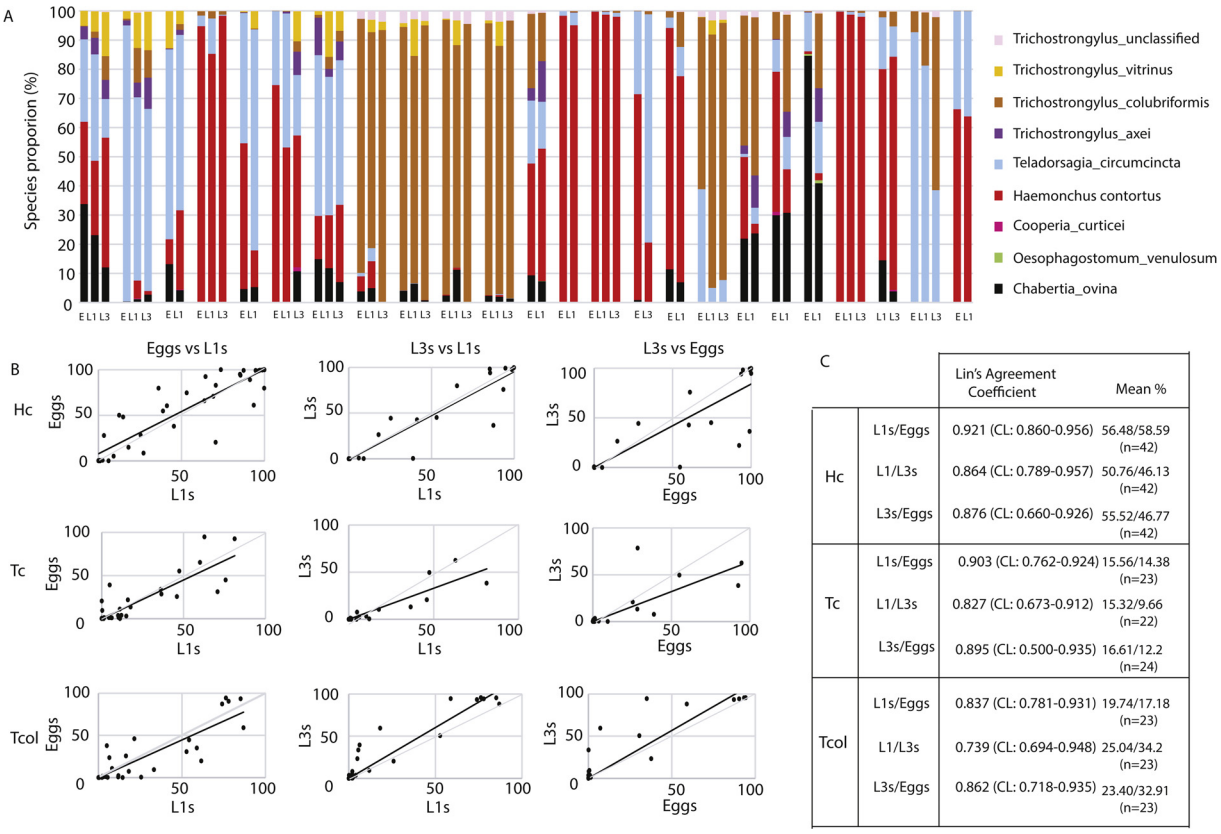
Mock communities containing exact known proportions of each of the six of the most important gastrointestinal nematode species of sheep were created by the counting of 40 individual larvae of each species into mixtures prior to DNA lysate preparation. This produced mock populations of 240 L3s comprising 40 L3s of *H. contortus, T. circumcincta, T. colubriformis, T. axei, T. vitrinus* and *C. curticei* (equal proportions of 16.67% each). Ten separate replicate mock communities were created, and the ITS-2 rDNA was independently amplified from each mock community three times to produce ten replicate communities each with three technical replicates. The accuracy of the ITS-2 amplicon sequencing in determining the species proportions was examined by comparing the proportion of sequence reads classified for each species with the expected equal proportion of 16.67% for each species. Overall, the species proportions determined by nemabiome sequencing were found to be similar to the true proportions in each mock community with very little variance between the ten replicate samples or the three PCR replicates for each sample (Fig. 1A). The range of differences between percentage proportions for each species being < 2.2% between the replicate samples and < 7.8% between the three PCR replicates. Nevertheless, some species were over-represented (*H. contortus, T. cicumcincta, T. colubriformis and T. vitrinus*) and some under-represented (*C. curticei and T. axei*) relative to their true proportions (Fig. 1A). Mean correction factors for each of the six species were estimated from the nemabiome mock community data by dividing the true proportions by the observed proportions of the compiled data (Fig. 1A).

For the large-intestinal worm species *Oesophagostomum venulosum* and *Chabertia ovina,* we were unable to source pure strains of L3s to include in the mock communities described above. Consequently, we took a different approach to determine the correction factors for these species. We determined the relative proportions of these two large intestinal nematode species in twelve field sample populations by determining the species identity of 90 single L3s using species-specific ITS-2 rDNA PCR (Supplementary Table S1). Nemabiome sequencing was then performed on DNA lysates prepared from pools of ~1000 L3s from each of these twenty populations. The proportions of *Oesophagostomum venulosum* and *Chabertia ovina* determined from this nemabiome sequence data was compared with the "true" proportions determined from single larva species-specific PCRs (Fig. 1B). These species were considerably under-represented in the uncorrected nemabiome data relative their true proportions. Correction factors of 2.499072 for *C. ovina* and 3.851243 for *O. venulosum* were calculated by dividing the true proportions (individual larvae PCR data) by the observed proportions from the nemabiome (Fig. 1B).

The effectiveness of these correction factors to compensate for species biases of the uncorrected nemabiome sequencing data were tested on a sub-set of farm samples (n = 7). The true species proportions for these farm samples was determined using species-specific ITS-2 rDNA PCR to determine the species identity of 90 individual larvae per population. Nemabiome ITS-2 rDNA sequencing data, generated from lysates prepared from ~ 1000 L3s from the same populations, was used to determine the species proportions with and without the application of the calculated correction factors. Comparison of the raw and corrected nemabiome data with the "true" species proportions, determined by single L3 species specific-PCR, revealed the corrected nemabiome sequence data was more accurate than the uncorrected data in each case (Fig. 1C). There was no statistically significant difference between the species proportions determined by the corrected nemabiome sequence data and the "true" species proportions determined by species-specific PCR of 90 larvae, based on Chi-square analysis ($p < 0.01$, Fig. 1C).

3.2. Comparison of Nemabiome sequencing applied to eggs, L1 and L3 larvae

The ITS-2 rDNA nemabiome amplicon sequencing assay was applied to DNA lysates made from eggs, L1s and L3s harvested from a set of 48 independent pooled fecal samples (pooled from 20 ewes per set) from 21 sheep farms in Western Canada (Fig. 2A). Pairwise comparisons between eggs versus L1s, L1s versus L3s and eggs versus L3s nemabiome sequences (following use of correction factors) was conducted for the predominant species across the samples; *H. contortus* (Hc), *T. circumcincta* (Tc) and *T. colubriformis* (Tcol). There was a high level of agreement between the species proportions determined from nemabiome sequencing data derived from eggs, L1s and L3s. Lin's agreement analysis indicated that there was very little difference in the species proportions calculated between eggs and L1s (Hc = 0.921, Tc = 0.903, Tcol = 0.837), between L3s and L1s (Hc = 0.864, Tc = 0.827, Tcol =



C

		Lin's Agreement Coefficient	Mean %
Hc	L1s/Eggs	0.921 (CL: 0.860-0.956)	56.48/58.59 (n=42)
	L1/L3s	0.864 (CL: 0.789-0.957)	50.76/46.13 (n=42)
	L3s/Eggs	0.876 (CL: 0.660-0.926)	55.52/46.77 (n=42)
Tc	L1s/Eggs	0.903 (CL: 0.762-0.924)	15.56/14.38 (n=23)
	L1/L3s	0.827 (CL: 0.673-0.912)	15.32/9.66 (n=22)
	L3s/Eggs	0.895 (CL: 0.500-0.935)	16.61/12.2 (n=24)
Tcol	L1s/Eggs	0.837 (CL: 0.781-0.931)	19.74/17.18 (n=23)
	L1/L3s	0.739 (CL: 0.694-0.948)	25.04/34.2 (n=23)
	L3s/Eggs	0.862 (CL: 0.718-0.935)	23.40/32.91 (n=23)

Fig. 2. Nemabiome results from eggs, L1s and L3s cultured from the same ovine fecal sample.
Panel A. Stacked bar-charts representing the species proportions calculated from corrected nemabiome sequencing data for 200 eggs (E), 200 L1s (L1) and 200 L3s (L3) harvested and cultured from the same fecal samples from each farm sample. A sub-set of 24 representative farm samples (out of a total number of 48 farm samples) are plotted.
Panel B. Scatter-plots of the species proportions of *H. contortus*, *T. circumcincta* and *T. colubriformis* resulting from the pairwise comparisons of eggs versus L1s, L3s versus L1s and L3s versus eggs respectively.
Panel C. Statistical analysis of the pairwise comparisons of the corrected nemabiome sequencing data from eggs, L1s and L3s for *H. contortus* (Hc), *T. circumcincta* (Tc) and *T. colubriformis* (Tcol). The left column shows Lin's agreement analysis and the right column shows the mean percentage abundance of each species in the two sample types being compared.

0.739) or between L3s and eggs (Hc = 0.876, Tc = 0.895, Tcol = 0.862) for all three species (Fig. 2B and C).

3.3. Application of ITS-2 rDNA amplicon sequencing to assess ovine gastrointestinal nematode species prevalence and abundance across UK sheep farms

In order to investigate the use of the ovine ITS-2 rDNA amplicon sequencing assay as a surveillance tool, we applied it to archived field samples (93 ewe samples and 61 lamb samples) from 99 different UK sheep farms collected in 2008 (Burgess et al., 2012). We compared the originally published prevalence data, based on conventional species-specific PCR, with our nemabiome sequencing data (Fig. 3). Dealing with each of the major species in turn: the PCR data from the original study showed that *T. circumcincta* had 100% prevalence on the 119 farms sampled in the UK (Burgess et al., 2012 and Fig. 3A). The nemabiome sequencing data confirmed the presence of *T. circumcincta* on all farms but also revealed additional information on the high degree of variance in relative abundance between farms (range 7.24%–98.89% of flock parasite population) (Fig. 3A). The nemabiome sequencing data also revealed that *T. circumcincta* comprised a higher mean percentage of parasite populations in ewes than in lambs (T-test, 2-tail: t(147) = 5.4, p = $3.0e^{-07}$, Fig. 3A). The original PCR data also reported a very high prevalence of *T. vitrinus* across UK sheep flocks (93% overall) (Fig. 3B). As with *T. circumcincta*, the nemabiome sequencing revealed a high degree of variance in the relative abundance of *T. vitrinus* between farms (range 0.02%–62.22% of flock parasite population) (Fig. 3B). In addition, the nemabiome sequencing data showed that *T. vitrinus* comprised a higher mean percentage of the parasite populations in lambs than in than in ewes (*t*-test, 2-tail: t(147) = 4.8, p = $4.3e^{-06}$); the opposite trend to that seen with *T. circumcincta* (Fig. 3B). In the case of *H. contortus*, the nemabiome sequencing data revealed a much more skewed distribution than for *T. circumcincta* or *T. vitrinus* (Fig. 3C). The relative abundance of *H. contortus* was very low on the majority of farms (< 10% on 82% of the farms) but was high on a small number of farms (> 50% on 3.3% of the farms) with a striking regional difference; a greater number of farms in England, than in Wales or Scotland, had a high relative abundance of *H. contortus*. A skewed distribution of relative abundance was also seen with *T. colubriformis* again with England having a higher number of high relative abundance farms than Scotland which had none (Fig. 3D). Although the original PCR data revealed a relatively high prevalence of *T. axei* and *C. curticei* on UK farms, the nemabiome sequencing data revealed that the relative abundance of these two species was consistently low in the flocks where they were present (mean abundances; *T. axei* at 3.6% in England, 2.4% in Wales and 3.4% in Scotland and *C. curticei* at 2% in England, 1.1% in wales and 1.6% in Scotland, Fig. 3E and F). The original study did not examine the two major ovine large intestinal parasitic nematode species, *O. venulosum* and *C. ovina* but the nemabiome sequencing revealed the overall mean relative abundances were *O. venulosum* at 5.9% in England, 6.4% in Wales and 8.1% in Scotland and *C. ovina* at 7.6% in England, 6.8% in Wales and 9.2% in Scotland

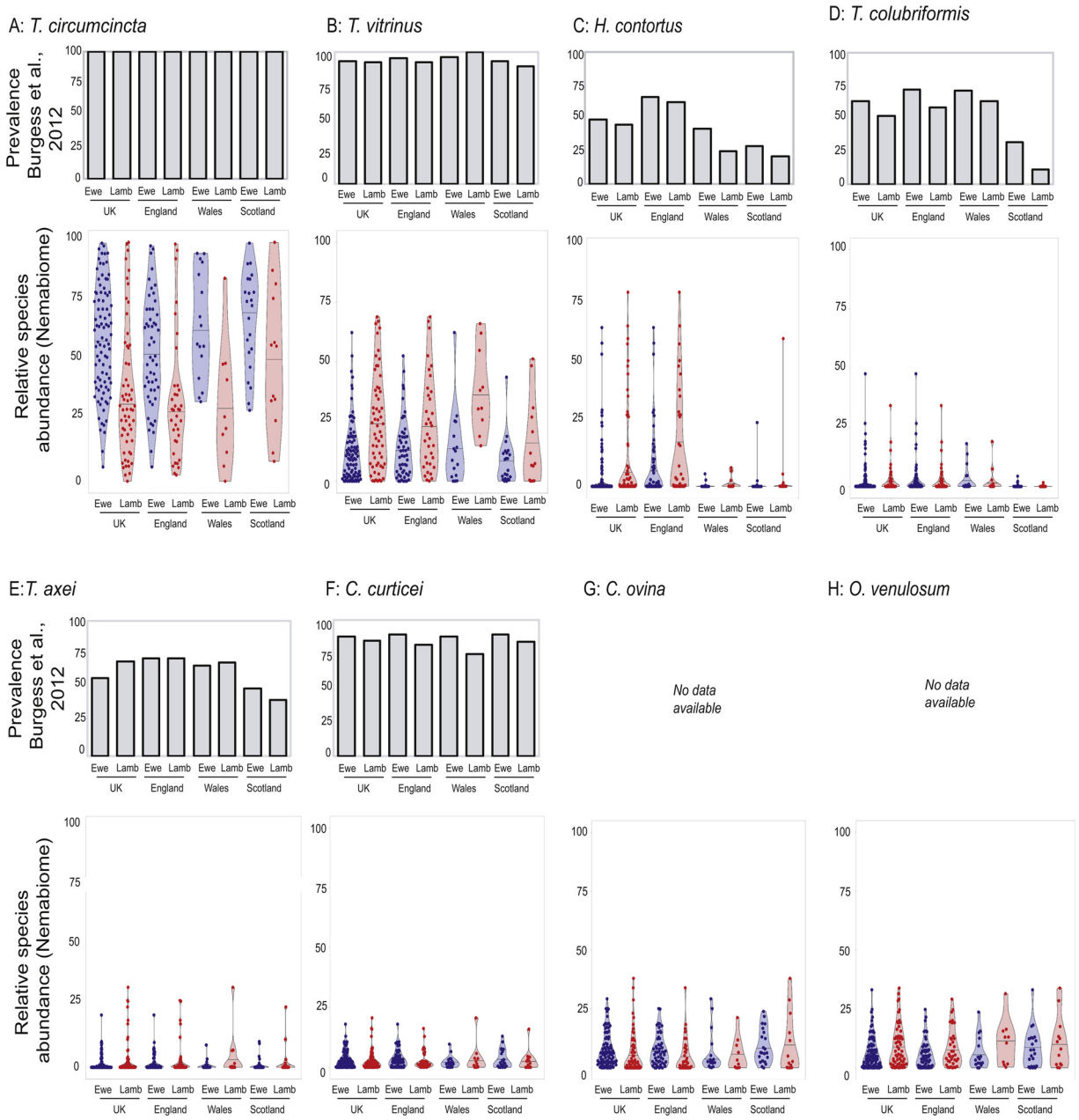


Fig. 3. Combination of species prevalence data with abundance data generated from the Nemabiome of archived sheep samples. Prevalence of six major GI parasite species of sheep estimated from species-specific assays (data taken from Burgess et al., 2012). Species proportion data generated by the nemabiome of GI parasite species of sheep (pools of approximately 1000 L1s) for the whole of the UK and for England, Wales and Scotland separately. Larvae collected early in the season were from ewes and later in the season were from lambs. Species proportions categorized by region/country and represented as violin plots showing datapoints and mean proportions. Data separated by species: (A) *T. circumcincta*, (B) *T. vitrinis*, (C) *H. contortus*, (D) *T. colubriformis*, (E) *T. axei*, (F) *C. curticei*, (G) *O. venulosum* and (H) *C. ovina.*

(Fig. 3G and H).

4. Discussion

We previously developed and validated ITS-2 rDNA nemabiome sequencing for determining the relative abundance of different gastrointestinal nematode species of cattle (Avramenko et al., 2015). In this paper, we validate the approach for ovine gastrointestinal nematode communities, investigate its use on eggs and L1s, as well as cultured L3s, and demonstrate its value as a surveillance tool to provide statistically robust data on relative GI nematode species abundance using a large set of samples from UK sheep farms.

As with bacterial 16S rDNA sequencing and integrated metabarcoding studies, there are a variety of potential sources of species representation bias in ITS-2 rDNA nemabiome sequencing data (Avramenko et al., 2015). These include factors related to the biology of the target organisms such as species-specific differences in cell number and thus DNA content, aswell as copy number variation of the ITS-2 rDNA target. They also include technical issues such as differences in the efficiency of genomic DNA extraction between species and sequence-dependant PCR amplification biases. It is very difficult to measure and control for all these differences individually but the net effect of the different factors contributing to bias can be experimentally determined and empirically compensated for, at least partially, using correction factors (Avramenko et al., 2015). For the major ovine gastrointestinal species found in the abomasum and small intestine, we had

access to pure populations of single species and so could use a similar approach that we had previously used for the cattle nemabiome work (Avramenko et al., 2015). That is, to make mock communities of known species composition and use them to determine the overall species representation bias of the ITS-2 rDNA nemabiome sequencing data. For the six nematode species from the abomasum and small intestine, we found that the biases were relatively small and so applying small correction factors was sufficient to provide a relatively accurate measure of relative species abundance (Fig. 1A). For the two most common nematode species of the ovine large intestine - *O. venulosum* and *C. ovina* – pure populations were not available. Consequently, we undertook a different approach. We determined the actual relative abundance of these two species, using species-specific PCR of 90 individual larvae, in each of 12 field populations (Fig. 1B). We then applied the nemabiome sequencing assay to the same populations and compared the nemabiome sequencing results to the actual species proportions. The representation biases for these two species was larger than for the abomasal and small intestinal species, with the relative abundance of both species being consistently underestimated and so larger correction factors were required (Fig. 1B). The use of these correction factors, for all eight species, were then tested on seven independent field populations, where the true relative abundance for each of the target species had been determined by species-specific PCR on 90 individual larvae. In all cases, the application of correction factors reduced biases and improved the accuracy of the results (Fig. 1C).

Our previous ITS-2 rDNA nemabiome sequencing of cattle gastrointestinal nematode communities used populations of L3 larvae harvested from fecal samples following culture for 21 days at 20–21 °C (Avramenko et al., 2015, 2017, 2018). This was largely for practical reasons as, in the case of cattle, harvesting L3s from fecal cultures is less labour intensive than extracting sufficient numbers of eggs from fecal samples due to the typically low egg counts. However, there are a number of potential problems with using cultured L3s for ITS-2 rDNA nemabiome sequencing. Firstly, the fecal environment is complex and so fecal culture conditions are difficult, indeed impossible, to standardize. Fungal, bacterial and protozoal communities will vary between different fecal samples, as will the moisture content and general composition related to environmental, dietary and health factors of the host. These complex, and poorly defined differences, could potentially have a major differential impact on the growth and survival of different nematode species depending on the nature of the samples and incubation conditions. Consequently, the species proportions determined from L3s could differ significantly from those determined from parasite eggs in fresh fecal samples. Secondly, from a practical perspective, the delay of 10–21 days to allow development to L3 is a potential problem if results are required more immediately such as in a diagnostic, or some experimental, situations. Consequently, we compared the ITS-2 rDNA nemabiome sequencing results generated from approximately 200 eggs, 200 freshly hatched L1s and 200 L3s harvested from fecal cultures from the same 48 samples. There was relatively little difference between the results generated from each of the three stages for each sample. Although there were some differences in the proportions of the minor species present, the nemabiome profiles were very similar overall and there was a high level of agreement for those parasite species present at sufficiently high proportions for meaningful statistical analysis (Fig. 2). This is an encouraging result as it suggests that the different parasite harvesting methodologies and parasite stages can be used with minimal impact on the results. Based on these results, we have adopted the use of freshly hatched L1s as our standard methodology for the ovine ITS-2 rDNA nemabiome sequencing assay. This is because preparative egg flotations can be undertaken reasonably efficiently in sheep due to the higher egg counts (hundreds or thousands of eggs per gram) and visual checking of hatching rates allows a visible check on sample quality and the viability of the harvested parasites prior to nemabiome sequencing. Further, the L1 stage represents a fixed point in development which is easier to standardize than eggs which are potentially at different development stages depending on the ages and storage condition of the fecal sample.

We investigated the use of the validated ITS-2 rDNA nemabiome sequencing assay as a surveillance tool using a set 93 ewe samples and 61 lamb samples, previously harvested and archived from 99 different UK sheep flocks in 2008 (Burgess et al., 2012). These samples had been previously characterized by conventional species-specific PCR assays to assess trichostrongylid nematode species prevalence in ewes and lambs across the UK (Burgess et al., 2012). The nemabiome sequencing data provided a much more comprehensive picture of the relative abundance and distribution of the major gastrointestinal nematode species in UK sheep farms than the original species-specific PCR data. The latter simply showed the most highly prevalent ovine gastrointestinal species to be *T. circumcincta* (overall prevalence of 100%) followed very closely by *T. vitrinus* (overall prevalence 93%). However, the nemabiome sequence data revealed much more detail on the relative abundance and distribution of these two parasite species. First, it revealed substantial variance in the relative abundance of both these species between farms (Fig. 3A and B). It also showed that the abundance of *T. circumcincta* was much higher than *T. vitrinus* overall in all three regions (England, Scotland and Wales). Further, it revealed significant differences in the relative abundance in these two species between ewes and lambs with *T. circumcinta* being more predominant in ewes than lambs and the opposite being the case for *T. vitrinus*. This provides a comprehensive, statistically supported, dataset to inform what is otherwise a combination of anecdotal and "expert" knowledge on the biology of these parasites. Similarly, the species-specific PCR data, showing a prevalence of 49% and 45% of *H. contortus* for ewes and lambs respectively, provides a somewhat misleading view of the situation for this parasite when taken in isolation. The nemabiome sequencing data reveals that the relative abundance of this parasite is very low on most of these farms but also shows it to be very high in a few cases. Hence, the nemabiome data reveals the sporadic distribution of this parasite in the UK very well. Finally, the detection, and determination of relative abundance, of the large intestinal nematodes, *C. ovina* and *O. venulosum,* illustrates the ability of nemabiome sequencing to detect additional species for which prior specific assays were not available. This demonstrates how the approach can provide a more comprehensive view of parasitic nematode community composition than is achieved when limited by the availability of assays for each individual species and/or applied on the basis of *a priori* assumptions of the likely species present.

5. Conclusion

This work demonstrates that ITS-2 rDNA nemabiome sequencing provides a more powerful and comprehensive approach to define ovine gastrointestinal nematode communities than the currently available molecular assays. The information presented in this paper on the relative abundance and distribution of eight different parasitic nematode species across 99 sheep flocks was achieved using just a single PCR on each sample and, following library preparation and pooling, a single sequencing run on the Illumina Miseq platform. This serves to illustrate the power of the nemabiome sequencing approach in large scale surveillance and epidemiological studies.

Funding

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC, grant number RGPIN-2015-03976), the University of Calgary, Canada, NSERC-CREATE Host Parasite Interactions (HPI) graduate training program and Alberta Agriculture and Forestry, (AAF, grant number 2015R029R).

CRediT authorship contribution statement

Elizabeth Redman: Formal analysis, Investigation, Methodology, Resources, Validation, Writing - original draft. **Camila Queiroz:** Methodology, Investigation. **David J. Bartley:** Resources, Writing - review & editing. **Michel Levy:** Resources, Writing - review & editing. **Russell W. Avramenko:** . **John Stuart Gilleard:** Conceptualization, Funding acquisition, Supervision, Writing - review & editing.

Declaration of Competing Interest

None.

Acknowledgement

The UK samples were originally collected as part of a CEDFAS BBRSC grant, (2008-2012).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.vetpar.2019.108933.

References

Agneessens, J., Dorny, P., Hollanders, W., Claerebout, E., Vercruysse, J., 1997. Epidemiological observations on gastrointestinal nematode infections in grazing cow-calf pairs in Belgium. Vet. Parasitol. 69, 65–75.

Avramenko, R., Redman, E.M., Lewis, R., Yazwinski, T.A., Wasmuth, J., Gilleard, J.S., 2015. Exploring the gastrointestinal "Nemabiome": deep amplicon sequencing to quantify the species composition of parasitic nematode communities. PLoS One 10 (12), e0143559. https://doi.org/10.1371/journal.pone.0143559.

Avramenko, R.A., Redman, E.M., Lewis, R., Bichuette, M.A., Palmeira, B.M., Yazwinski, T.A., Gilleard, J.S., 2017. The use of nemabiome metabarcoding to explore gastrointestinal nematode species diversity and anthelmintic treatment effectiveness in beef calves. Int. J. Parasitol. 47, 893–902.

Avramenko, R.A., Braas, A., Redman, E.M., Woodbury, M., Wagner, B., Shury, T., Liccioli, S., Windeyer, C., Gilleard, J.S., 2018. High species diversity of trichostrongyle parasite communities within and between Western Canadian commercial and conservation bison herds revealed by nemabiome metabarcoding. Parasit. Vectors 11, 299.

Bisset, S.A., Knight, J.S., Bouchet, C.L., 2014. A multiplex PCR-based method to identify strongylid parasite larvae recovered from ovine faecal cultures and/or pasture samples. Vet. Parasitol. 200, 117–127.

Besier, R.B., Kahn, L., Sargison, N.D., Van Wyk, J.A., 2016. The pathophysiology, ecology and epidemiology of *Haemonchus contortus* infection in small ruminants. In: In: Gasser, R., Samson-Himmelstjerna, G.V. (Eds.), *Haemonchus contortus* and Haemonchosis Past, Present and Future Trends, vol. 9. pp. 95–144.

Blaxter, M., Koutsovoulos, G., 2015. The evolution of parasitism in Nematoda. Parasitol. 142 (Suppl. 1), S26–S39. https://doi.org/10.1017/S0031182014000791.

Burgess, C.G., Bartley, Y., Redman, E., Skuce, P.J., Nath, M., Whitelaw, F., Tait, A., Gilleard, J.S., Jackson, F., 2012. A survey of the trichostrongylid nematode species present on UK sheep farms and associated anthelmintic control practices. Vet. Parasitol. 189, 299–307.

Elmahalawy, S., Halvarsson, P., Skarin, M., Höglund, J., 2018. Droplet digital polymerase chain reaction (ddPCR) as a novel method for absolute quantification of major gastrointestinal nematodes in sheep. Vet. Parasitol. 261, 1–8.

Gasser, R.B., Chilton, N.B., Hoste, H., Stevenson, L.A., 1994. Species identification of trichostrongyle nematodes by pcr-linked rflp. Int. J. Parasitol. 24, 291–293.

Gasser, R.B., Bott, N.J., Chilton, N.B., Hunt, P., Beveridge, I., 2008. Toward practical, DNA-based diagnostic methods for parasitic nematodes of livestock – bionomic and biotechnological implications. Biotechnol. Adv. 26, 325–334.

Giudici, C., Aumont, G., Mahieu, M., Saulai, M., Cabaret, J., 1999. Changes in gastrointestinal helminth species diversity in lambs under mixed grazing on irrigated pastures in the tropics (French West Indies). Vet. Res. 30, 573 -58.

Learmont, J., Conyers, C., Hird, H., Morgan, C., Craig, B.H., von-Samson-Himmelstjerna, G., Taylor, M., 2009. Development and validation of real-time PCR methods for diagnosis of *Teladorsagia circumcincta* and *Haemonchus contortus* in sheep. Vet. Parasitol. 166, 268–274.

McMurtry, L.W., Donaghy, M.J., Vlassoff, A., Douch, P.G.C., 2000. Distinguishing morphological features of the third larval stage of ovine Trichostrongylus spp. Vet. Parasitol. 90, 73–81.

O'Connor, L.J., Walkden-Brown, S.W., Kahn, L.P., 2006. Ecology of the free-living stages of major trichostrongylid parasites of sheep. Vet. Parasitol. 142, 1–15.

Redman, E., Packard, E., Grillo, V., Smith, J., Jackson, F., Gilleard, J.S., 2008. Microsatellite analysis reveals marked genetic differentiation between *Haemonchus contortus* laboratory isolates and provides a rapid system of genetic fingerprinting. Int. J. Parasitol. 38, 111–122.

Redman, E., Whitelaw, F., Tait, A., Burgess, C., Bartley, Y., Skuce, P.J., Jackson, F., Gilleard, J.S., 2015. The emergence of resistance to the benzimidazole anthelmintics in parasitic nematodes of livestock is characterised by multiple independent hard and soft selective sweeps. PLoS Negl. Trop. Dis. 9 (2), e0003494. https://doi.org/10.1371/journal.pntd.0003494.

Roeber, F., Jex, A.R., Campbell, A.J.D., Campbell, B.E., Anderson, G.A., Gasser, R.B., 2011. Evaluation and application of a molecular method to assess the composition of strongylid nematode populations in sheep with naturally acquired infections. Infect. Genet. Evol. 11, 849–854.

Sangster, N.C., Kelly, J.D., Whitlock, H.V., Gunawanf, M., Porter, C.J., 1980. *Trichostrongylus colubriformis* and *Ostertagia* sp resistant to levamisole, morantel tartrate and thiabendazole: infectivity, pathogenicity and drug efficacy in two breeds of sheep. Res. Vet. Sci. 29, 26–30.

Sargison, N., Wilson, D., Bartley, D.J., Penny, C., Jackson, J., 2007. Haemonchosis and teladorsagiosis in a Scottish sheep flock putatively associated with the overwintering of hypobiotic fourth stage larvae. Vet. Parasitol. 147, 326–333.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. (December), 7537–7541 2009.

Stromberg, B.E., Gasbarre, L.C., Ballweber, L.R., Dargatz, D.A., Rodriguez, J.M., Kopral, C.A., Zarlenga, D.S., 2015. Prevalence of internal parasites in beef cows in the United States: results of the national animal health monitoring systems (NAHMS) beef study, 2007–2008. Can. J. Vet. Res. 79, 290–295.

Van Wyk, J.A., Cabaret, J., Michael, L.M., 2004. Morphological identification of nematode larvae of small ruminants and cattle simplified. Vet. Parasiol. 119, 277–306.

Vlassoff, A., 1976. Seasonal incidence of infective trichostrongyle larvae on pasture: the contribution of the ewe and the role of the residual pasture infestation as sources of infection to the lamb. New Zealand J. Exp. Agric. 4, 281–284.