PROTOZOOLOGY - ORIGINAL PAPER

Untargeted metagenomics shows a reliable performance for synchronous detection of parasites

Claudia Wylezich[1] · Simone M. Caccio[2] · Julia Walochnik[3] · Martin Beer[1] · Dirk Höper[1]

Abstract

Shotgun metagenomics with high-throughput sequencing (HTS) techniques is increasingly used for pathogen identification and characterization. While many studies apply targeted amplicon sequencing, here we used untargeted metagenomics to simultaneously identify protists and helminths in pre-diagnosed faecal and tissue samples. The approach starts from RNA and operates without an amplification step, therefore allowing the detection of all eukaryotes, including pathogens, since it circumvents the bias typically observed in amplicon-based HTS approaches. The generated metagenomics datasets were analysed using the RIEMS tool for initial taxonomic read assignment. Mapping analyses against ribosomal reference sequences were subsequently applied to extract 18S rRNA sequences abundantly present in the sequence datasets. The original diagnosis, which was based on microscopy and/or PCR, could be confirmed in nearly all cases using ribosomal RNA metagenomics. In addition to the pre-diagnosed taxa, we detected other intestinal eukaryotic parasites of uncertain pathogenicity (of the genera *Dientamoeba*, *Entamoeba*, *Endolimax*, *Hymenolepis*) that are often excluded from routine diagnostic protocols. The study clearly demonstrates the applicability of untargeted RNA metagenomics for the parallel detection of parasites.

Keywords Metagenomics · Parasite detection · Ribosomal RNA · Reference mapping · Sequence type · Proficiency test

Introduction

Microscopy remains the gold standard for diagnosis of many parasitic infections, but requires skilled personnel and is time consuming. Moreover, morphologic features are often insufficient to identify the organism at the species level. Furthermore, the relatively low prevalence of many parasitic

✉ Claudia Wylezich
Claudia.Wylezich@fli.de

[1] Institute of Diagnostic Virology, Friedrich-Loeffler-Institut, Federal Research Institute for Animal Health, Südufer 10, 17493 Greifswald-Insel Riems, Germany

[2] Department of Infectious Diseases, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome, Italy

[3] Molecular Parasitology, Institute for Specific Prophylaxis and Tropical Medicine, Medical University of Vienna, Kinderspitalgasse 15, 1090 Vienna, Austria

infections in Western industrialized countries has caused a decrease in awareness and a loss of expertise. Conventional molecular pathogen detection is also a challenge for emerging or neglected pathogens for which no sequence information and consequently no well-established test systems are available. The establishment of new specific tests is complicated by the lack of standardized reference material, especially in the case of parasites. For all these reasons, a fast and efficient one-serves-all approach is desirable for rapid and synchronous parasite detection.

Untargeted (PCR-free) metagenomics approaches are well suited for the detection of pathogens of different phylogenetic affiliations including viral and bacterial pathogens (e.g., Frank et al. 2011; Hanke et al. 2017; Rubbenstroth et al. 2019). In recent years, also infections caused by parasites (including asymptomatic intestinal infections caused by protists and helminths) have been detected using metagenomics (e.g., Kawai et al. 2012; Gao et al. 2016; Schneeberger et al. 2016). The detection of parasite signature sequences in metagenomics datasets can be improved using RNA instead of DNA, as shown for enteric protists in faecal samples (Wylezich et al. 2019). Indeed, the very large number of ribosomal reference sequences present in public databases allows the extraction

and taxonomic assignments of ribosomal sequences from metagenomics datasets. In contrast, fully sequenced and well-annotated reference genomes are still lacking for many parasites, limiting the use of DNA-based sequencing and reference mapping approaches (e.g., Marzano et al. 2017; Stensvold and van der Giezen 2018). Another benefit of the untargeted metagenomics approach is that no prior decision is necessary for which pathogen to screen, i.e., which specific test system needs to be applied, since sequences derived from all potential pathogens will be present in the metagenomics dataset. This saves time and money by avoiding application of many laborious diagnostic techniques successively until a suspicion can be confirmed.

In a recent proof-of-concept study, we demonstrated the applicability of RNA metagenomics for parasite detection using faecal samples that were not pre-diagnosed for any enteric eukaryote (Wylezich et al. 2019). In the present follow-up study, we investigated pre-diagnosed samples with the same approach to confirm the original diagnosis made with conventional methods (microscopy, PCR) and to provide a more comprehensive picture of the non-host eukaryotes in the samples.

Materials and methods

Investigated sample materials

Diagnostic stool samples from different archive collections were included in this study: (1) Ethanol-fixed parasite-containing stool samples were obtained from the Medical University of Vienna, Austria (MUV samples). The samples were pre-diagnosed by light microscopy after sodium acetate-acetic acid-ethanol concentration. *Entamoeba*-positive samples were further differentiated at the species level by real-time PCR (Blessmann et al. 2002). The Friedrich Loeffler Institut (FLI) was provided with the samples including the results of the pre-diagnoses. (2) Stool samples were pre-diagnosed for protists and helminths using light microscopy and Lugol staining, and further screened by specific PCR assays for detection of *Ascaris, Giardia, Blastocystis* and *Entamoeba* (compare Table 1) at the Istituto Superiore di Sanità, Roma, Italy (ISS). To simulate an proficiency test for the metagenomics workflow, the ISS samples were sent frozen to the FLI without providing the results of the pre-diagnoses. Beside faecal material, (3) frozen samples of wild boar tissue infected with *Trichinella, Fasciola* and the lungworm *Dictyocaulus viviparus* were obtained from the official veterinary service of Western Pomerania, Anklam, Germany. Those veterinary-diagnostic (VD) samples were microscopically checked before transfer to the FLI and the diagnoses were provided with the material. Upon arrival at FLI, all

samples were stored deeply frozen (−80 °C) until further processing.

Metagenomics sequencing

Ethanol-fixed samples were centrifuged (12,000 rpm, 5 min, 4 °C) to remove ethanol, and the pellets were washed three times using 1 ml of 1 × TE buffer by mixing and subsequent centrifugation. The washed and re-suspended pellets (for MUV samples), the thawed ISS samples and pieces of the VD samples were treated with the cryoPREP (CP) instrument (Covaris) for disintegration as described (Wylezich et al. 2018). For two samples (ISS-2 and ISS-6), we additionally extracted RNA from subsamples without prior CP treatment and with combined treatment of CP plus Covaris M220 Focused-ultrasonicator (75 W, 1.5 min 7 °C). RNA was extracted in all cases, typically with the RNeasy Mini kit (Qiagen, Hilden, Germany). For two samples (ISS-C and ISS-D), disintegrated subsamples were additionally extracted using the Agencourt RNAdvance Tissue Kit (Beckman Coulter), following the manufacturer's instructions (Table 1). Libraries were prepared from the RNA as described by Wylezich et al. (2018), and sequenced on the Ion Torrent S5XL platform.

Detection of parasite signatures in metagenomics datasets

Generated metagenomics datasets were analysed using a combination of the RIEMS tool (Scheuch et al. 2015) and reference mapping against 18S rRNA gene sequences of suspected candidates using the Genome Sequencer software suite (versions 2.6; Roche) as described by Wylezich et al. (2019). Accession numbers of the sequences used as references are included in the **Supplementary Table S1**. Different identity thresholds (95–99%) and a minimum overlap length of reads of 95% were applied for reference mapping. The obtained 18S rRNA sequences were verified by Blast analysis (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome). Whereas for the proof-of-concept study (Wylezich et al. 2019) only nearly complete 18S rRNA sequences extracted from metagenomics datasets were rated as positive, in the present study, no minimum fragment length of the obtained contigs was presupposed for detection but instead each single read was considered positive if it was a specific hit.

The 18S rRNA sequences generated within this study were submitted to GenBank when the fragments were longer than 1000 base pairs (see **Supplementary Table S1**). They are accessible under the accession numbers MN914072-MN914086.

Table 1 Samples investigated in the present study, pre-diagnosis and results obtained using metagenomics sequencing. MUV and ISS samples are stool samples; VD samples represent tissue samples. Taxa detected in the pre-diagnosis and taxa confirmed using metagenomics are shown in green, respectively. Taxa that were additionally found using metagenomics although not pre-diagnosed are shown in blue

Sample	Taxa detected via pre-diagnosis	Metagenomics		
		Library ID *disintegration step extraction kit size [reads] (unclassified)*	**Suspected by RIEMS**	**Confirmed by reference mapping** (18S rRNA)
MUV-1	*Giardia duodenalis* – by microscopy	**L3195** *cryoPREP RNAdvance 3,420,917 (0.8%)*	*Giardia intestinalis* **Blastocystis sp.** *Entamoeba* spp.	*Giardia duodenalis* **Blastocystis ST1***
MUV-2	*Giardia* – by microscopy *Blastocystis* – by microscopy	**L2178** *cryoPREP RNeasy 232,188 (0.3%)*	*Giardia intestinalis* **Blastocystis sp.**	*Giardia intestinalis* **Blastocystis ST3***
MUV-4	*Entamoeba* – by microscopy and PCR *Blastocystis* – by microscopy	**L1963** *cryoPREP RNeasy 2,114*	Not applied	[pre-diagnosed taxa could not be confirmed]
MUV-5	*Enterobius vermicularis* – by microscopy	**L3196** *cryoPREP RNAdvance 2,549,786 (1.2%)*	*Blastocystis hominis* **Enterobius vermicularis**	*Enterobius vermicularis**
ISS-2	*Ascaris* – by microscopy *Blastocystis* – by PCR	**L2852** *none RNeasy 579,241 (14.0%)*	*Ascaris* spp. *Hymenolepis* spp. **Blastocystis sp.** *Entamoeba* spp. *Cryptosporidium* spp.	*Ascaris / Baylisascaris* **Blastocystis ST1***
		L2854 *cryoPREP RNeasy 1,587,391 (0.3%)*	*Ascaris* spp. **Blastocystis sp.** *Entamoeba* spp.	*Ascaris / Baylisascaris* **Blastocystis ST1***
		L2856 *cryoPREP + M220 RNeasy 2,006,185 (0.4%)*	*Ascaris* spp. **Blastocystis sp.** *Entamoeba histolytica*	*Ascaris / Baylisascaris* **Blastocystis ST1***
ISS-4	*Ascaris* – by microscopy *Blastocystis* – by PCR	**L2922** *cryoPREP + M220 RNeasy 3,174,771 (0.04%)*	*Ascaris* spp. **Blastocystis sp.** **Entamoeba spp.**	*Ascaris sp./ A. suum** **Blastocystis ST1*** **Entamoeba spp.***
ISS-5	*Ascaris* – by microscopy *Giardia* assemblage B – by microscopy and PCR *Blastocystis* – by PCR	**L2919** *cryoPREP RNeasy 2,610,224 (0.8%)*	*Ascaris* spp. *Giardia intestinalis* **Blastocystis sp.** *Entamoeba* spp. **Endolimax nana** *Cryptosporidium parvum*	Ascarididae *Giardia intestinalis* **Blastocystis ST1/ST3** **Endolimax nana**
ISS-6	*Giardia* assemblage A – by microscopy and PCR *Blastocystis* – by PCR	**L2853** *none RNeasy 2,924,441 (0.1%)*	*Ascaris lumbricoides* *Giardia intestinalis* **Blastocystis sp.** *Entamoeba dispar*	*Giardia intestinalis* **Blastocystis ST3**
		L2855 *cryoPREP RNeasy 1,747,497 (0.06%)*	*Giardia intestinalis* **Blastocystis sp.**	*Giardia intestinalis* **Blastocystis ST3***
		L2857 *cryoPREP + M220 RNeasy 1,693,733 (0.05%)*	*Giardia intestinalis* **Blastocystis sp.**	*Giardia intestinalis* **Blastocystis ST3***
ISS-7	*Giardia* assemblage A – by microscopy and PCR	**L2920** *cryoPREP RNeasy 3,067,228 (0.1%)*	*Giardia* spp. **Hymenolepis spp.** **Dientamoeba fragilis**	*Giardia intestinalis* **Hymenolepis nana** **Dientamoeba fragilis**
ISS-8	*Giardia* assemblage A – by microscopy and PCR	**L2921** *cryoPREP*	*Ascaris lumbricoides* *Giardia intestinalis*	*Giardia intestinalis* **Blastocystis ST3***

Table 1 (Continued)

	Blastocystis – by PCR	RNeasy 5,429,451 (0.1%)	*Blastocystis* sp. Entamoeba histolytica *Dientamoeba fragilis* Cryptosporidium parvum	*Dientamoeba fragilis*
ISS-C	*Cryptosporidium parvum* – by microscopy and PCR	L3197 cryoPREP RNeasy 2,828,010 (2.3%)	Blastocystis hominis *Cryptosporidium* spp.	*Cryptosporidium parvum*
		L3193 cryoPREP RNAdvance 1,862,001 (3.4%)	Blastocystis hominis *Cryptosporidium* spp.	*Cryptosporidium parvum*
ISS-D	*Giardia* – by microscopy (IFA) *Cryptosporidium parvum* - by microscopy and PCR	L3198 cryoPREP RNeasy 430,212 (28.1%)	Giardia intestinalis *Cryptosporidium* spp.	*Giardia intestinalis* *Cryptosporidium parvum*
		L3194 cryoPREP RNAdvance 2,187,213 (25.6%)	Blastocystis sp. Giardia intestinalis *Cryptosporidium* spp.	*Giardia intestinalis* *Cryptosporidium parvum*
VD-1	**Trichinella** - by microscopy	L1806 cryoPREP RNeasy 1,627,079 (0.4%)	Blastocystis sp. *Trichinella* spp. Cryptosporidium	*Trichinella spiralis*
VD-2	*Fasciola hepatica* - by microscopy	L1949 cryoPREP RNeasy 568,673 (0.7%)	Blastocystis sp. *Fasciola* spp. Entamoeba polecki Hymenolepis spp.	[pre-diagnosed taxon could not be confirmed in terms of assembled contigs]
		L1950 cryoPREP RNeasy 491,774 (0.2%)	Blastocystis sp. *Fasciola* spp.	[pre-diagnosed taxon could not be confirmed in terms of assembled contigs]
		L1951 cryoPREP RNeasy 555,425 (0.3%)	Blastocystis sp. *Fasciola* spp. Cryptosporidium parvum	[pre-diagnosed taxon could not be confirmed in terms of assembled contigs]
VD-3	*Dictyocaulus viviparus* - by microscopy	L1804 cryoPREP RNeasy (AL buffer) 1,200,725 (0.1%)	Blastocystis sp.	[pre-diagnosed taxon could not be confirmed]
		L1805 cryoPREP RNeasy (RLT buffer) 1,271,891 (0.1%)	Blastocystis sp.	[pre-diagnosed taxon could not be confirmed]

*18S rRNA fragment >1,000 bp length was obtained using reference mapping

Results and discussion

The 18S rRNA gene and other sequences of the ribosomal RNA cluster have been a preferential target for diagnostics of parasites (Kounosu et al. 2019, and references therein). However, primers used for diagnosis might miss certain parasite taxa or lead to cross-amplification of the (also eukaryotic) host or sample background, or generate short fragments not sufficient for a reliable phylogenetic assignment. Kounosu et al. (2019) evaluated several primer sets for 18S and 28S rRNA amplification and highlighted difficulties in identifying primers that enable synchronous detection of all eukaryotes, including parasites, in a single PCR. In particular, unicellular parasites were often not detected and received short amplicons were unsuited for an accurate taxonomic resolution (Kounosu et al. 2019). In the present study, we exploited the abundance of ribosomal sequences in metagenomics datasets obtained using RNA as starting material. The use of shotgun sequencing of total RNA, instead of DNA, should also positively shift the detection limit of PCR-diagnostics, due to the higher abundance of 18S rRNA transcripts compared to the corresponding coding genes (DNA), as recently demonstrated for the protistan parasite *Babesia* (Hanron et al. 2017).

General performance of the untargeted metagenomics workflow

An overview of samples and associated results is given in Table 1. A detailed overview of used conditions (e.g. identity thresholds), resulting contigs and their identity to sequenced taxa from databases is provided as **Supplementary Table S1**. Overall, pre-diagnosed taxa were confirmed in all cases with the exception of samples MUV-4, VD-2 and VD-3. Concerning sample MUV-4, a very poor library could be constructed from the low amount of extracted RNA only delivering a minute sequence dataset (2114 reads). The fixation of the MUV-4 sample might have had a negative impact on the recovery of RNA. The resulting dataset was therefore too small for a successful detection of the taxa. Since helminths can be distributed very unevenly in the host tissue and their eggs in stool samples, respectively, the reason for failure with sample VD-3 could be the use of a disadvantageous subsample not containing any specimen of the pre-diagnosed taxa. This might also be true for VD-2. The dataset contained only a few reads of the liver fluke (*Fasciola hepatica* was found via RIEMS analysis, Table 1), which could not be assembled into contigs. For all other samples, pre-diagnosed parasites as well as additional protists were detected via untargeted metagenomics with yields between 0.00004% (for *Blastocystis* ST1 in sample ISS-5) and 15.7% (for *Cryptosporidium parvum* in sample ISS-C; **Supplementary Table S1**).

For two samples (ISS-2 and ISS-6), sample processing with disintegration (CP and CP + M220 Focused-ultrasonicator) and without disintegration was compared. For sample ISS-2, the detected number of reads from *Blastocystis* and *Ascaris* was higher without disintegration than with any disintegration. In these datasets, the number of parasite reads is especially influenced by shifts of the detected bacterial read numbers. As already shown, the application of a disintegration step may result in a higher representation of nucleic acids from specific certain bacteria (Wylezich et al. 2018). Regarding sample ISS-6, no difference in the number of parasite reads was found, yet inclusion of a disintegration step (CP and CP + M220 Focused-ultrasonicator) resulted in longer contigs compared to those obtained without prior disintegration (for *Blastocystis*). This is mainly caused by the amount of unclassified reads (Table 1) that was higher in datasets generated without prior disintegration compared with datasets that were generated with any disintegration (datasets of samples ISS-2 and ISS-6; Table 1). This negative impact in terms of unclassified reads in datasets generated without disintegration step has already been documented (Wylezich et al. 2018). While in the present cases, the disintegration resulted not always in an obviously better performance in terms of the parasite read amount, we recommend including this step to obtain high quality datasets and especially for detecting protistan cysts and/or helminth eggs.

With respect to the RNA extraction kits used, we could detect higher parasite read numbers for both samples ISS-C and ISS-D (*Cryptosporidium* spp., *Giardia* spp.; **Supplementary Table S1**) extracted with the RNeasy Mini Kit compared to the datasets of the same samples extracted with the Agencourt RNAdvance Tissue Kit. Although the results for the two samples rather support the use of the RNeasy Mini Kit, the selection of the optimal kit needs to be investigated in a systematic validation and might depend on the sample type.

Difficulties of the detection of *Giardia*

Giardia is the most commonly detected gut flagellate, worldwide, and can cause infections in humans and animals. The infective dose is known to be particularly low, already a few cysts can result in infection. In the present study, only one or two reads were found in microscopically *Giardia*-positive samples when using the 18S rRNA (M54878, AF006676, HQ179642) as reference sequence. When using a full genome sequence (UZAE01000001) for reference mapping, we found that most of the detected reads match with the rRNA tandem repeat unit but not necessarily with the 18S rRNA gene. Therefore, we re-run the reference mapping using the ribosomal tandem repeat (X52949) as reference, which resulted in a better mapping outcome though the read amount was still very low (<0.0009%) and reads mostly belonged to the 28S

rRNA gene that is longer than the 18S rRNA gene. The ribosomal sequences of *Giardia* exhibit a very high GC content (about 74%, Wang et al. 2011) in contrast to the *Giardia* genome (about 46%, Smith et al. 1998). However, since genome sequences of *Giardia* were not overrepresented in the datasets over ribosomal sequences, we exclude a GC-shift of sequences that can be caused by some library preparation kits (Grützke et al. 2019). A possible explanation for the low recovery of *Giardia* reads can be a low number of *Giardia* cysts in the samples. Unfortunately, we have no quantitative data for the samples to verify this possibility. Also the storage conditions might have been suboptimal for these samples.

Storage of samples

Basically, metagenomics sequencing enables to describe the sample complexity in terms of a sequence dataset, provided optimal storage of the sample to preserve its original state. Basically, fresh or deeply frozen material is well suited for metagenomics, especially when working with RNA. In the present study, samples were from the archives of diagnostic laboratories and not originally intended for metagenomics sequencing. Therefore, storage conditions might have been suboptimal, e.g., the cold chain was not fully maintained, and this may have caused loss of phylogenetic signals in terms of molecules for certain parasites. Ethanol fixation might be an alternative as shown here with the MUV samples when a freezer is not or not immediately available. However, one of the ethanol-fixed samples (MUV-4) resulted in a very poor library delivering only a small dataset.

Additional findings besides the pre-diagnosed taxa

Blastocystis was confirmed in all pre-diagnosed cases and additionally found in a sample that had been pre-screened only by light microscopy but not by PCR indicating a low *Blastocystis* density in the respective sample (MUV-1). We detected additional species using the metagenomics approach. Since these taxa cause no severe diseases, according to the current state of knowledge, they are typically not included in routine diagnostics. For example, *Endolimax nana* belongs to the family Entamoebidae and is often observed in faecal samples but, due to its non-pathogenic nature, is typically not reported. Likewise, enteric infections with *Hymenolepis nana* and *Dientamoeba fragilis* are typically considered benign, although controversy persists regarding their pathogenic potential. For example, *H. nana* was identified as causative agent in a case of invasive cestodiasis in an immunocompromised patient with a spread of the infection to an ectopic location (Olson et al. 2003). In terms of diagnostics, *D. fragilis* requires specific stainings or at least concentration methods and the analysis of fresh samples to observe the relevant morphological features. PCR diagnostics is sometimes applied for this

parasite but bears the risk of false-positive results (Intra et al. 2019 and references therein). In the present study, the species was readily detected in metagenomics datasets.

Some implausible findings obtained with the RIEMS tool (e.g. *Blastocystis* in the VD tissue samples) are false-positive hits based on misassignments due to not correctly curated genomes in public databases or the lack of suitable reference genomes as already discussed in detail (Wylezich et al. 2019). For example, recently published sequences attributed to *Blastocystis* sp. (accession numbers MN339606, MK782501, MK782521) occurred in some Blast analyses as closest hits to the query sequence. However, the mentioned sequences are closely related to Dipodascaceae (Ascomycota) and do not show any close relationship to *Blastocystis* species (e.g. ST1, KY610205).

The aforementioned taxa might be often overlooked, at least when only present in low densities, and do therefore not contribute to the overall picture of a sample and are not included in the interpretation of the symptoms. For unresolved cases of intestinal disorders, however, it might be advantageous to get a more comprehensive picture of the individual gut microbiome allowing the interpretation of all comprised opportunistic agents and their mutual effects. In relation to this, a further benefit of the present method should be mentioned; it allows the parallel detection of prokaryotic taxa based on 16S and 23S rRNA without any primer bias. Another important and growing field of application for metagenomics-based diagnostics is the screening of organs and donor blood before transplantation or transfusion, respectively, as also asymptomatic, opportunistic pathogens may have serious consequences for the immunocompromised recipients (Hanron et al. 2017).

Concluding remarks

Altogether, the application of untargeted RNA metagenomics sequencing for diagnostic samples was promising in the present study, especially since it can be used for the parallel identification of different unrelated taxa. In addition, it provides a more complete picture on protists that are often not included in routine diagnostics (e.g., *Dientamoeba fragilis*, "non-pathogenic" intestinal amoebae). Untargeted metagenomics sequencing is therefore applicable as one-serves-all pathogen detection tool swiftly providing comprehensive information of the sample. We recommend using it as a snapshot approach especially in cases with a difficult diagnosis because of contradictory symptoms or for infections in immunocompromised individuals. However, reasonable care has to be taken with the analysis of the obtained results using current public databases.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All samples included in this study were retrieved from the archives of diagnostic institutions. Samples were fully anonymised, the sole information linked to the samples was being positive or negative for the respective parasite in the respective diagnostic test.

References

Blessmann J, Buss H, Nu PA, Dinh BT, Ngo QT, Van AL, Alla MD, Jackson TF, Ravdin JI, Tannich E (2002) Real-time PCR for detection and differentiation of *Entamoeba histolytica* and *Entamoeba dispar* in fecal samples. J Clin Microbiol 40:4413–4417. https://doi.org/10.1128/JCM.40.12.4413-4417.2002

Frank C, Werber D, Cramer JP, Askar M, Faber M, an der Heiden M, Bernard H, Fruth A, Prager R, Spode A, Wadl M, Zoufaly A, Jordan S, Kemper MJ, Follin P, Müller L, King LA, Rosner B, Buchholz U, Stark K, Krause G (2011) Epidemic profile of Shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany. N Engl J Med 365:1771. DOI: https://doi.org/10.1056/NEJMoa1106483, 1780

Gao D, Yu Q, Wang G, Wang G, Xiong F (2016) Diagnosis of a malayan filariasis case using a shotgun diagnostic metagenomics assay. Parasites Vectors 9:86. https://doi.org/10.1186/s13071-016-1363-2

Grützke J, Malorny B, Hammerl JA, Busch A, Tausch SH, Tomaso H, Deneke C (2019) Fishing in the soup – pathogen detection in food safety using metabarcoding and metagenomic sequencing. Front Microbiol 10:1805. https://doi.org/10.3389/fmicb.2019.01805

Hanke D, Pohlmann A, Sauter-Louis C, Höper D, Stadler J, Ritzmann M, Steinrigl A, Schwarz BA, Akimkin V, Blome S, Beer M (2017) Porcine epidemic diarrhea in Europe: in-detail analyses of disease dynamics and virus evolution. Viruses 9:177. https://doi.org/10.3390/v9070177

Hanron AE, Billman ZP, Seilie AM, Chang M, Murphy SC (2017) Detection of *Babesia microti* parasites by highly sensitive 18S rRNA reverse transcription PCR. Diagnost Microbiol Infect Dis 87:226–228. https://doi.org/10.1016/j.diagmicrobio.2016.11.021

Intra J, Sarto C, Besana S, Tiberti N, Brambilla P (2019) The importance of considering the neglected intestinal protozoan parasite *Dientamoeba fragilis*. J Med Microbiol 68:890–892. https://doi.org/10.1099/jmm.0.000983

Kawai T, Sekizuka T, Yahata Y, Kuroda M, Kumeda Y, Iijima Y, Kamata Y, Sugita-Konishi Y, Ohnishi T (2012) Identification of *Kudoa septempunctata* as the causative agent of novel food poisoning outbreaks in Japan by consumption of *Paralichthys olivaceus* in raw fish. Clin Infect Dis 54:1046–1052. https://doi.org/10.1093/cid/cir1040

Kounosu A, Murase K, Yoshida A, Maruyama H, Kikuchi T (2019) Improved 18S and 28S rDNA primer sets for NGS-based parasite detection. Sci Rep 9:15789. https://doi.org/10.1038/s41598-019-52422-z

Marzano V, Mancinelli L, Bracaglia G, Del Chierico F, Vernocchi P, Di Girolamo F, Garrone S, Kuekou HT, D'Argenio P, Dallapiccola B, Urbani A, Putignani L (2017) Omic" investigations of protozoa and worms for a deeper understanding of the human gut "parasitome. PLoS Negl Trop Dis 11:e0005916. https://doi.org/10.1371/journal.pntd.0005916

Olson PD, Yoder K, Fajardo LFL-G, Marty AM, van de Pas S, Olivier C, Relman DA (2003) Lethal invasive cestodiasis in immunosuppressed patients. J Infect Dis 187:1962–1966. https://doi.org/10.1086/375357

Rubbenstroth D, Peus E, Schramm E, Kottmann D, Bartels H, McCowan C, Schulze C, Akimkin V, Fischer N, Wylezich C, Hlinak A, Spadinger A, Grossmann E, Petersen H, Grundhoff A, Rautenschlein S, Teske L (2019) Identification of a novel clade of group a rotaviruses in fatally diseased domestic pigeons in Europe. Transbound Emerg Dis 66:552–561. https://doi.org/10.1111/tbed.13065

Scheuch M, Höper D, Beer M (2015) RIEMS: a software pipeline for sensitive and comprehensive taxonomic classification of reads from metagenomics datasets. BMC Bioinf 16:69. https://doi.org/10.1186/s12859-015-0503-6

Schneeberger PHH, Becker SL, Pothier JF, Duffy B, N'Goran EK, Beuret C, Frey JE, Utzinger J (2016) Metagenomic diagnostics for the simultaneous detection of multiple pathogens in human stool specimens from Côte d'Ivoire: a proof-of-concept study. Infect Genet Evol 40:389–397. https://doi.org/10.1016/j.meegid.2015.08.044

Smith MW, Aley SB, Sogin M, Gillin FD, Evans GA (1998) Sequence survey of the *Giardia lamblia* genome. Mol Biochem Parasitol 95: 267–280. https://doi.org/10.1016/s0166-6851(98)00113-3

Stensvold CR, van der Giezen M (2018) Associations between gut microbiota and common luminal intestinal parasites. Trends Parasitol 34:369–377. https://doi.org/10.1016/j.pt.2018.02.004

Wang W, Wie Z, Lam TW, Wang J (2011) Next generation sequencing has lower sequence coverage and poorer SNP-detection capability in the regulatory regions. Sci Rep 1:55. https://doi.org/10.1038/srep00055

Wylezich C, Papa A, Beer M, Höper D (2018) A versatile sample processing workflow for diagnostic metagenomics. Sci Rep 8:13108. https://doi.org/10.1038/s41598-018-31496-1

Wylezich C, Belka A, Hanke D, Beer M, Blome S, Höper D (2019) Metagenomics for broad and improved parasite detection: a proof-of-concept study using swine faecal samples. Int J Parasitol 49:769–777. https://doi.org/10.1016/j.ijpara.2019.04.007