

9-23-2019

A Case for Using Genomics and a Bioinformatics Pipeline to Develop Sensitive and Species-Specific PCR-Based Diagnostics for Soil-Transmitted Helminths

Jessica R. Grant
Smith College

Nils Pilotte
Smith College

Steven A. Williams
Smith College, swilliam@smith.edu

Follow this and additional works at: https://scholarworks.smith.edu/bio_facpubs



Part of the [Biology Commons](#)

Recommended Citation

Grant, Jessica R.; Pilotte, Nils; and Williams, Steven A., "A Case for Using Genomics and a Bioinformatics Pipeline to Develop Sensitive and Species-Specific PCR-Based Diagnostics for Soil-Transmitted Helminths" (2019). Biological Sciences: Faculty Publications, Smith College, Northampton, MA. https://scholarworks.smith.edu/bio_facpubs/159

This Article has been accepted for inclusion in Biological Sciences: Faculty Publications by an authorized administrator of Smith ScholarWorks. For more information, please contact scholarworks@smith.edu



A Case for Using Genomics and a Bioinformatics Pipeline to Develop Sensitive and Species-Specific PCR-Based Diagnostics for Soil-Transmitted Helminths

Jessica R. Grant^{1*}, Nils Pilotte^{1,2} and Steven A. Williams^{1,2}

¹ Department of Biological Sciences, Smith College, Northampton, MA, United States, ² Molecular and Cellular Biology Program, University of Massachusetts, Amherst, MA, United States

OPEN ACCESS

Edited by:

Makedonka Mitreva,
Washington University
Medical Center, United States

Reviewed by:

Emmanuel Dias-Neto,
A.C. Camargo Cancer Center, Brazil
John Stuart Gilleard,
University of Calgary, Canada

*Correspondence:

Jessica R. Grant
jgrant@smith.edu

Specialty section:

This article was submitted to
Evolutionary and
Genomic Microbiology,
a section of the journal
Frontiers in Genetics

Received: 26 April 2019

Accepted: 21 August 2019

Published: 23 September 2019

Citation:

Grant JR, Pilotte N and Williams SA
(2019) A Case for Using Genomics
and a Bioinformatics Pipeline to
Develop Sensitive and Species-
Specific PCR-Based Diagnostics for
Soil-Transmitted Helminths.
Front. Genet. 10:883.
doi: 10.3389/fgene.2019.00883

The balance of expense and ease of use vs. specificity and sensitivity in diagnostic assays for helminth disease is an important consideration, with expense and ease often winning out in endemic areas where funds and sophisticated equipment may be scarce. In this review, we argue that molecular diagnostics, specifically new assays that have been developed with the aid of next-generation sequence data and robust bioinformatic tools, more than make up for their expense with the benefit of a clear and precise assessment of the situation on the ground. Elimination efforts associated with the London Declaration and the World Health Organization (WHO) 2020 Roadmap have resulted in areas of low disease incidence and reduced infection burdens. An accurate assessment of infection levels is critical for determining where and when the programs can be successfully ended. Thus, more sensitive assays are needed in locations where elimination efforts are approaching a successful conclusion. Although microscopy or more general PCR targets have a role to play, they can mislead and cause study results to be confounded. Hyper-specific qPCR assays enable a more definitive assessment of the situation in the field, as well as of shifting dynamics and emerging diseases.

Keywords: soil-transmitted helminth, molecular diagnostics, DNA diagnostics, polymerase chain reaction (PCR), quantitative PCR

INTRODUCTION

Parasitic worms impact the health and economic well-being of billions of people worldwide. Soil-transmitted helminths (STH) are a burden in the tropics and subtropics and contribute to an estimated 1.9 to 2.1 million disability-adjusted life years (DALYs) and US \$7.5 billion to US \$138.9 billion in loss of productivity (Bartsch et al., 2016; Kyu et al., 2018). Efforts are underway to eliminate STH, with the ambitious goal of controlling morbidity by the year 2020 (Becker et al., 2018; Uniting to Combat NTDs). Mass drug administrations (MDA) and water, sanitation and hygiene (WASH) programs across endemic countries are making headway (Hicks et al., 2015; Truscott et al., 2016; Weatherhead et al., 2017; Truscott et al., 2019), but with 2020 fast approaching, there are still many challenges to reaching this goal. An important concern is where to enact and when to cease MDA. This depends on accurately mapping the current burden in communities (2018 Action Group

Meeting). Sensitive, species-specific diagnostics are critical to properly evaluating the success of these programs, as well as addressing where to focus efforts and when interventions can be ended (Weatherhead et al., 2017).

Diagnostic techniques need to be inexpensive, practical, and give consistent results across technicians and laboratories. Importantly, they must be accurate, sensitive, and easily interpreted. Microscopy has long been relied on as the standard for diagnosis of intestinal parasites, including soil transmitted helminths (Beaver and Martin, 1968). Several copromicroscopic methods are in use, including FLOTAC (Cringoli et al., 2010), MINI-FLOTAC (Maurelli et al., 2014), several modifications of the McMaster technique (Mines, 1977), and Kato-Katz (Katz et al., 1972) (KK). Of these, KK is the most commonly used for STH diagnosis because it is relatively easy to perform in the field and is generally more sensitive than other microscopic methods (Moser et al., 2018). With any of these tests, even highly trained microscopists can misidentify species or give inconsistent results (Krauth et al., 2012), and they are notoriously insensitive in regions with low infection rates (Nikolay et al., 2014; Buonfrate et al., 2015; Speich et al., 2015; Acosta Soto et al., 2017).

Molecular diagnostics have been garnering more interest in the last few years, as their superior sensitivity has been proven and their acceptance by the research community has increased (Easton et al., 2016; Halfon et al., 2017; Holt et al., 2017). However, as with any advance, there are technical problems to overcome. For example, DNA extraction efficiency and preservation of samples prior to testing will affect diagnostic reliability (Andersen et al., 2013; Sarhan et al., 2015; Hidalgo et al., 2018; Papaïakovou et al., 2018). Notably, *Trichuris trichiura* eggs are notoriously difficult to break open, and this impacts the sensitivity of molecular assays, but techniques are being developed and improved to the point where consistently good results are achievable (Harmon et al., 2006; Nunes et al., 2006; Kaisar et al., 2017). Although molecular diagnostics are not inexpensive, microscopy techniques are also expensive and can be difficult to scale up, whereas the costs of qPCR have the potential to decrease, as studies show that multiple technical replicates may not be crucial and other cost-cutting measures, such as cheaper, more effective sample preservation and pooling are explored (Easton et al., 2017; Papaïakovou et al., 2018; Truscott et al., 2019). Until recently, most PCR-based assays have targeted well-characterized and conserved regions, such as ITS and 18S (Verweij and Stensvold, 2014; Hii et al., 2018), but increased availability of whole-genome sequence data is facilitating the discovery of more sensitive and species-specific genomic targets (Pilotte et al., 2016a; Papaïakovou et al., 2017).

Repetitive elements are essential parts of eukaryotic genomes that have structural and regulatory functions (Shapiro and Sternberg, 2005; López-Flores and Garrido-Ramos, 2012), and different types of repetitive DNA elements have been studied and classified (Charlesworth et al., 1994; Plohl et al., 2008; López-Flores and Garrido-Ramos, 2012; Biscotti et al., 2015). Ribosomal DNA is found in repeat arrays (Lafontaine and Tollervy, 2001; Pruesse et al., 2007). These have traditionally been used for primer design and can give sensitive results depending on the size of the array (Verweij and Stensvold, 2014). However,

the repeat is oftentimes conserved between species and even genera, and rDNA-based assays are often less specific than those designed from other repeat types (Pilotte et al., 2016; O'Connell et al., 2018).

Tandemly repeated DNA is classified by the size of the repeated monomer, resulting in microsatellites (< 9 bp), minisatellites (< 15 bp in arrays of 0.5–30 kb), and satellites (satDNA, up to ~200 bp per monomer, in megabase-sized arrays) that are generally enriched within the centromeric, pericentromeric, and subtelomeric regions of the chromosome (López-Flores and Garrido-Ramos, 2012). Copy number can be quite variable in mini- and micro-satellites but larger satDNA monomers are more consistent within species (López-Flores and Garrido-Ramos, 2012). Microsatellites and minisatellites are not useful for assay design, as the repeats tend to be too short to allow for primer/probe design. Larger satDNA monomers, on the other hand, offer the best options for assay design, in that the repeat monomers are an optimal size for qPCR, they are extremely abundant, and the copy number is relatively stable within species (López-Flores and Garrido-Ramos, 2012).

Other types of repetitive DNA include transposons and retrotransposons, which are dispersed throughout the genome. These include short interspersed nuclear elements (SINEs) which are 100 to 500 bases long, and long interspersed nuclear elements (LINEs) which are larger—6,000 or 7,000 bases long. These may also be useful for repeat-based assay design (Funakoshi et al., 2017).

The amount of repetitive DNA in any eukaryotic species is variable and can make up quite a large percentage of the genome. A recent study of parasitic worms revealed that both genome size and repeat content of the genomes range widely, with repeat elements forming up to 37% of the genomes in STH of interest (Table 1). Repetitive elements make up even greater percentages in other eukaryotes (Brindley et al., 2003; Wickstead et al., 2003; Shapiro and Sternberg, 2005; Coghlan et al., 2019), up to an astonishing 97% in some plants (Flavell et al., 1974; Piednoël et al., 2012). These repetitive elements, because of their abundance in the genome, provide targets for molecular assays of exquisite sensitivity. In addition, since many appear unbound by selective pressure, they can be highly species-specific. This combination of improved sensitivity and species-specificity makes repetitive elements a prime target for molecular diagnostics. However, as mentioned above, some forms of repetitive DNA, such as simple short nucleotide repeats, will be unsuitable for assay design. Another potential stumbling block is sequence variation in the repeat itself as polymorphism in the primer and probe sites will decrease the sensitivity of the assay. However, although the repeats we have targeted have not been specifically investigated, maintenance of homogeneity in repetitive elements by concerted evolution has been discussed in relation to other repeats in other species (Ganley and Kobayashi, 2007; Teruel et al., 2014). Concerted evolution of repeats conserves the sequence within a species while allowing significant heterogeneity between species. There can, of course, be some variation within repeats (Onyabe and Conn, 1999; Lindner and Banik, 2011) which could lessen the sensitivity of repeat-based molecular assays. Copy number variation of repeats between individuals may also impact the sensitivity of these assays in some populations. Studies have

TABLE 1 | Genome size of representative helminth species and the amount/percent of genome masked as repetitive. Low complexity repeats and simple (for example di- or tri-nucleotide repeats) are not targets for molecular assays.

Species	Assembly size (Mb)	Repeat-masked (Mb)	Repeat %	% of assembly that is low complexity/simple repeats
<i>Ancylostoma ceylanicum</i> (Hii et al., 2018)	349	128.6	36.8	0.5
<i>Ancylostoma duodenale</i> (Hii et al., 2018)	332.9	116.1	34.9	0.7
<i>Ascaris lumbricoides</i> germline (Okorie and de Souza, 2016)	334	n/a	16.8	n/a
<i>Ascaris lumbricoides</i> soma (Okorie and de Souza, 2016)	291	n/a	7	1.5
<i>Necator americanus</i> (Masny et al., 2016; Hii et al., 2018)	244.1	67.1	27.5	1.2
<i>Strongyloides stercoralis</i> (Pilotte et al., 2017; Hii et al., 2018)	42.7	4.4	10.3	4.2
<i>Trichuris trichiura</i> (Hii et al., 2018)	75.5	18.4	24.4	0.4

found population-level copy number variation in ribosomal repeats of different species (Bik et al., 2013; Schaap et al., 2013; Mascagni et al., 2018; Zhao and Gibbons, 2018), but there is a suggestion that there are both copy number variable-type repeats and constant-type repeats, whose copy number is consistent within species (Umemori et al., 2013). An understanding of copy number variation and its impact on assay sensitivity will likely need to be studied on an individual species-by-species basis.

The value of repeats as diagnostic tools has been understood for some time but these sequences were more difficult to find in the pre-genomics era (McReynolds et al., 1986; Meredith et al., 1989; Chanteau et al., 1994; Nekrutenko et al., 2000; Hamburger et al., 2001; Rao et al., 2002; Demas et al., 2011; Lodh et al., 2016). Now, abundant genomic data and robust bioinformatics tools are available to make these targets easier to identify and use in developing PCR-based assays. The pipeline leading from low coverage NGS data to hyper-sensitive and specific qPCR assay is not overly complicated or time-consuming, and the ability to repurpose low coverage NGS data from other studies makes this an attractive option for diagnostic development for helminthology and many other fields.

FINDINGS

Diagnostics give an estimate of the true prevalence of a disease, with the probability of correctly estimating the truth given by the sensitivity and species-specificity of the diagnostic. WHO guidelines for when to treat a community are informed by the prevalence of disease in that community, and are thus influenced by the sensitivity of the diagnostic used. A study modeling the probability of making the correct treatment decisions given WHO guidelines for treatment and varying the true prevalence and diagnostic sensitivity shows that there is a significant difference in outcome when more sensitive diagnostics are used (Medley et al., 2016). Medley et al. (Medley et al., 2016) found it especially true in areas of intermediate true prevalence (between 30% and 50%). More sensitive tests allowed the correct treatment decision to be made more often in intermediate cases.

The aforementioned study measured outcome by looking at DALYs and found that these were not as influenced by diagnostic sensitivity in low or high prevalence areas. However, there are other reasons to prefer more sensitive tests in low prevalence areas.

As the goals of the WHO elimination programs are reached, there will be pressure to reallocate the funds spent on MDA to other programs. A well-defined threshold under which recrudescence will not occur is critical to preventing the reoccurrence of disease after the completion of MDA. Restarting such programs would be extremely difficult and expensive once they have ended. Modeling has shown that the threshold must be based on true prevalence (Truscott et al., 2017; Ásbjörnsdóttir et al., 2017), which can only be accurately estimated with highly sensitive and species-specific diagnostics. Improved diagnostics are crucial to meet this need (Kongs et al., 2001; Andersen et al., 2013; Clarke et al., 2018; Hidalgo et al., 2018). Post-MDA surveillance is also necessary. With highly sensitive diagnostics, the reappearance of disease can be recognized and addressed before infection levels rise, increasing the probability of controlling the recrudescence (Farrell and Anderson, 2018). In addition to testing human populations for infection, vectors that transmit helminths (or other parasites) or intermediate hosts can be screened to track disease prevalence in the community without the need for taking human samples (Sanpool et al., 2012; Pilotte et al., 2016b; Ramírez et al., 2018; Zaky et al., 2018). Pooling of samples is a common way to decrease cost but diagnostic tests used for the screening of pooled samples need to be highly sensitive (Masny et al., 2016; Okorie and de Souza, 2016; Pilotte et al., 2017), especially when the infection level is low, so as not to miss positive results in dilute samples.

In recent years, efforts to sequence nematode genomes by groups, such as the International Helminth Genomes Consortium, have made great strides in increasing the availability of helminth genomic sequence data (Howe et al., 2016). Although many of the genomes are in draft form, this is sufficient for probing the genome for species-specific repeats. Our method for recovering highly repetitive sequence from low depth, raw, short-read genome sequence data uses the Galaxy-based tool RepeatExplorer (Novák et al., 2010; Novák et al., 2013; Pilotte et al., 2016a). Originally used to investigate repeat sequences in plant genomes, RepeatExplorer takes as input short-read next-generation sequence data and creates graph-based clusters based on the similarity of the sequences. In these graphs (see **Figure 1**), each read is represented by a node, and each sequence pair (by default defined as $\geq 90\%$ identity over at least 55% of the read lengths) is represented by an edge. The density of the graph represents the number and similarity of reads in the cluster. In

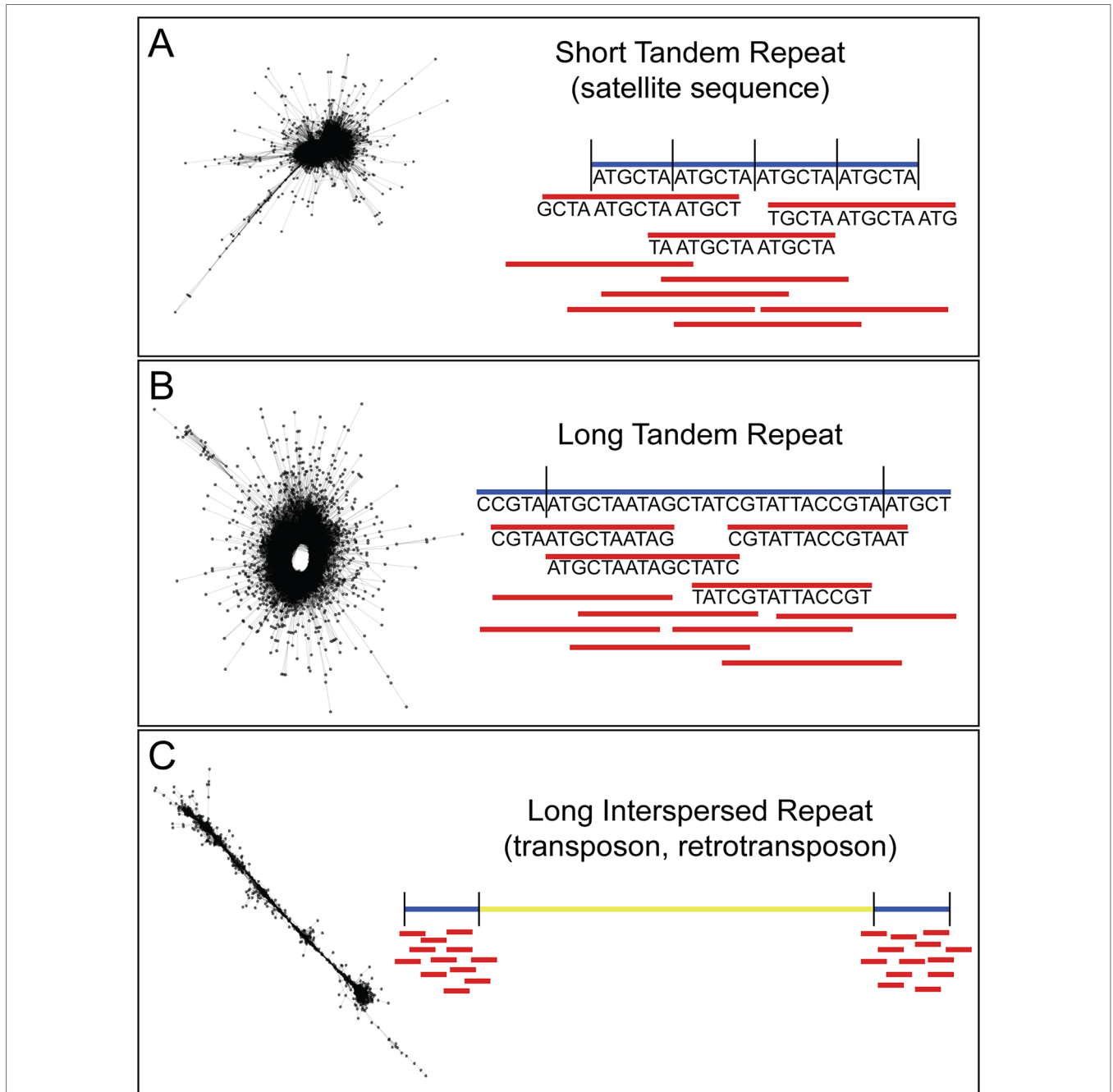


FIGURE 1 | Description of RepeatExplorer cluster output. **(A)** Short tandem repeats, including many satellite sequences, form characteristic “star burst”-type clusters. Because they are tandem, and such repeats are of similar size or shorter than the length of an individual sequence read, a very high percentage of reads within the cluster meet the RepeatExplorer-defined criteria for pair formation. This results in each read successfully pairing with a very high percentage of the other reads assigned to the cluster. Because nearly all of the reads within the cluster are paired with nearly all of the other reads within this same cluster, a compact network of very short edges forms between reads. This in turn generates a very dense cluster with a core of paired reads possessing nearly identical sequences. If of sufficient length for assay design, the consensus sequences for these clusters make ideal diagnostic targets, as they contain the greatest number of repeats per read. **(B)** Long tandem repeats characteristically result in “doughnut”-like clusters. In such clusters, neighboring reads within the underlying scaffold meet the criteria for pair formation. However, because the length of the repeat monomers generating these clusters is longer, reads may be significantly shorter than the monomer itself. This results in many reads within the cluster that do not meet the criteria for pair formation as they map to different regions of the same monomer unit. Yet because they are tandemly arranged, reads spanning a repeat-repeat junction will meet the pair formation criteria, closing the sequence “loop” and resulting in a “doughnut”-shaped cluster. **(C)** Long interspersed repeats, such as transposable elements form characteristic “line”-type clusters. While reads neighboring each other in the underlying scaffold meet the criteria for pair formation, the extended length of a repeat monomer means that distant reads within a single monomer will not meet this threshold. This results in similar pairings to those seen in clusters generated from long tandem repeats. However, because these elements are interspersed, reads do not span repeat-repeat junctions, so a “loop” is not formed, and clusters attain a linear appearance.

low-depth sequence data sets, low copy number sequences will not be well represented and will, therefore, graph as individual nodes or small clusters, whereas high copy number repetitive sequences will be found in dense clusters. The number of reads in a given cluster, combined with the structure and density of that cluster, can be used as a proxy for the representation of the number of repeats of that sequence in the genome. The copy number of the repeat will impact the sensitivity of the assay, since each copy in the genome will be an additional target for the assay. In addition, RepeatExplorer and its sister software TAREAN (Novák et al., 2017) provide information on the count of individual nucleotides in the repeat contigs. These counts can be used to design assays to the most conserved regions, limiting the problems associated with intragenomic variation. We have developed, and made available here, custom Python scripts to parse the output from RepeatExplorer and return highly repetitive sequences (<https://github.com/JessicaGrant/RepeatTargetScripts>). Primer/probe qPCR assays targeting sequences discovered using this technique have been shown to amplify as little as 20 ag of genomic DNA, or less than the amount of DNA found in a single egg (Pilotte et al., 2016a). Care is needed in choosing which repeat to select for use as a diagnostic target, as some may be found in closely related species (Williams et al., 2000; Rao et al., 2006); however, similar to what has been reported in the literature (Subirana and Messegue, 2013; Subirana et al., 2015), we have found that many repetitive sequences are species-specific. There may be times when a more general assay—one that will amplify several species of the same genus, for example—may be desired (Rao et al., 2006). A careful search of RepeatExplorer output can often reveal both species- and genera-specific targets.

Diagnostics based on targets discovered using this technique have proven useful in both past and ongoing field tests in Kenya (Pickering et al., 2019), Bangladesh (Benjamin-Chung et al., 2019), Ethiopia, Uganda, Timor Leste (Papaiakovou et al., 2017), Thailand (O'Connell et al., 2018), Liberia (Fischer et al., 2018), Japan, Benin, Malawi, India, and the Southern US, and have been adopted for use by large operational research efforts, such as the DeWorm3 cluster randomized trials (Ásbjörnsdóttir et al., 2018). However, testing biological samples, whether for diagnosing individuals or getting an overview of the epidemiological environment of a region, involves a myriad of factors, such as unexpected or emerging parasites, zoonotic infections, and misleading material in the samples. Although most of the criticism of the KK technique has been on the lack of sensitivity and potential for missed infection, there is also the risk of false positives, for example, mistaking other material in stool as eggs (Speich et al., 2015). Some fecal elements may resemble parasite ova, depending on environmental or dietary factors. Confounding elements may include pollen grains, fungal spores, diatoms, or any number of items. An entire chapter in Ash and Orihel's "Atlas of Human Parasitology" is dedicated to artifacts in fecal samples that can mislead copromicroscopic diagnostics (Cushion et al., 1990). Thus, this problem is more frequent than many researchers realize or acknowledge; what follows are some examples demonstrating the importance of this issue.

A field study comparing KK with repeat-based qPCR in Bangladesh (Benjamin-Chung et al., 2019) found that hookworm

species and *Trichuris trichiura* prevalence, as measured by qPCR, was significantly higher than was measured by KK. This was expected, given the greater sensitivity of the qPCR assays and the results of many previous studies comparing KK and PCR (Pontes et al., 2003; Stensvold et al., 2006; Knopp et al., 2014; Pilotte et al., 2016a; Easton et al., 2017; Ng-Nguyen et al., 2017). For *Ascaris*, however, prevalence as measured by KK was significantly higher than by qPCR. This surprising result was investigated further, both by qPCR targeting a different part of the *Ascaris* genome, and also by amplicon sequencing that targeted the 18S gene of all eukaryotes in several of the KK positive/qPCR negative samples. All of the samples that were positive by KK but negative by the initial qPCR assay were also negative using the second qPCR target. Additionally, the 18S amplicon sequencing revealed no *Ascaris* in these samples, but did find it in the control samples that were positive by both KK and qPCR. Not one organism was found in the amplicon sequencing that could explain all of the false-positive results. What material in the samples had confounded the microscopists is still unknown, but there was no evidence by any of the molecular assays that *Ascaris* was present in the samples. Had the study relied on copromicroscopic results alone, the conclusion would have been that MDA or WASH interventions were less effective as an *Ascaris* intervention than they likely were, since the true prevalence of the parasite was in fact much lower than was measured by KK.

In a similar case, higher than expected rates of hookworm were noted in a survey of children in rural Rwanda. Further investigation suggested these results may have been confounded by *Caenorhabditis elegans* eggs (Irisarri-Gutiérrez et al., 2016). Additional examples of misidentification of hookworm ova as other eggs (Ralph et al., 2006; Werneck et al., 2007; Yong et al., 2007) show that such confusion may be a more common problem than previously thought. Thus, relying solely on microscopy may be misleading in some instances.

Discrepancies can occur between molecular assays as well, since some PCR targets are less species-specific than others. In developing our pipeline for repeat-based primer discovery, a previously published qPCR assay targeting the internal transcribed spacer region was compared against our newly developed assay targeting an *Ancylostoma duodenale* species-specific repeat (Llewellyn et al., 2016; Pilotte et al., 2016a). Surprisingly, the repeat-based assay failed to detect any of the samples that were determined to be positive for *A. duodenale* by the ITS-based assay. A previously published semi-nested PCR assay (George et al., 2015; Chidambaram et al., 2017) and Sanger sequencing later determined that the discordant results were due to all of the infections being the zoonotic species *Ancylostoma ceylanicum*. The ITS of these two species is highly conserved in the region targeted by the original qPCR assay and so the ITS-based assay did not distinguish between these closely related species. Our repeat-based *A. duodenale* assay, on the other hand, only detects *A. duodenale*, and so all of the *A. ceylanicum*-containing samples were negative. We have since used our pipeline to develop a species-specific qPCR assay for *A. ceylanicum* (Papaiakovou et al., 2017), which is more sensitive and specific than the ITS-based assay and easier to use than semi-nested PCR.

In a similar case, the same ITS-based primer set (Llewellyn et al., 2016) was used to investigate *Ancylostoma duodenale* in a field study of a refugee population in Thailand (O'Connell et al., 2018). Since the most common human *Ancylostoma* parasite is *A. duodenale*, the results were initially believed to indicate *A. duodenale* infection. Again, however, a corroboratory qPCR targeting the highly specific *A. duodenale* repeat failed to detect any *A. duodenale*. Use of the *A. ceylanicum*-specific assay (Papaiakovou et al., 2017), as well as confirmation with semi-nested PCR and Sanger sequencing, revealed that all of these infections were, in fact, caused by *A. ceylanicum* and not *A. duodenale*. In this case, although the more general ITS-based assay misdiagnosed the species causing the infection, the more specific assay for the expected parasite (*A. duodenale*) would have missed the infection. This highlights a risk of using extremely specific qPCR assays in the field if the precise parasite community is unknown. It also highlights that *A. ceylanicum* may be a much more common human pathogen than previously supposed. Here, the repeat-based, species-specific assays can be used to identify the true prevalence of various related parasite species.

The specificity of *Trichuris trichiura* (whipworm) detection by microscopy is assumed to be high, given the relatively distinct morphology of *Trichuris* eggs. However, there are several species of *Trichuris*, including some that infect companion or farm animals. Distinction between species of *Trichuris* relies on size differentiation, but there are overlaps in some species making misdiagnosis by microscopy possible. The most common species in humans is *T. trichiura*, but *Trichuris suis*, which commonly infects pigs, and *Trichuris vulpis*, usually found in dogs, have also been found infecting humans (Areekul et al., 2010; Mohd-Shaharuddin et al., 2019). Microscopy and genera-specific qPCR assays may be confounded by these zoonotic species. Discordance between the ITS-based *Trichuris* qPCR assay and the repeat-based qPCR assay has been noted in one study where all but one of the discordant results were later shown to be *Trichuris ovis*, a species found in sheep and goats that has not been known to infect humans (Pilotte et al., 2016a). Whether this is evidence of human infection or merely false positives due to close contact with infected animals or ingesting food contaminated with animal feces is an open question. In these cases, qPCR targeting highly species-specific repetitive targets alone could easily miss the prevalence of zoonotic infection, leading to a misunderstanding of the health of the population. On the other hand, the use of the repeat-based, highly specific assays gives a true picture of the prevalence of various species.

Despite the distinct morphology of *Trichuris*, the potential of misidentifying unexpected infection with zoonotic species is not the only risk when relying solely on microscopy for investigating whipworm. A study of STH in two regions in Liberia that used both microscopy and qPCR found discordance between the tests for *T. trichiura* in one of the regions (Fischer et al., 2018), with 25 of 27 putatively positive samples for *T. trichiura*, as determined by KK, being negative by qPCR. In the second region, the agreement between the two tests was high. In this case, the discrepancies were investigated further, and the eggs were determined to be a species of *Capillaria*, a human parasite that is associated with

eating raw fish and was not expected to be found in this region. Reinvestigation by microscopy in this case elucidated subtle differences in the eggs found in the KK-positive, qPCR-negative samples. A microscopist, highly trained and expecting to have to differentiate between two extremely similar egg morphologies, could have noticed the difference and provided a correct result. However, since *Capillaria* had not previously been reported in Liberia, the eggs that looked like *Trichuris* were reported as such. Without the complementary qPCR, this mistake would not have been discovered. This study also found discrepancies between the microscopy and qPCR measuring the prevalence of *Ascaris lumbricoides* in the same region where *Capillaria* was discovered. Surprisingly, these samples were also determined to be confounded by the presence of *Capillaria*, which can appear rounder and have more subtle polar plugs, leading to its misidentification as the eggs of *A. lumbricoides*. These examples highlight the tendency for microscopists to sometimes see what they are looking for. In this case, the unexpected presence of *Capillaria* eggs misled the microscopists and would have resulted in significant misinterpretation of the distribution of STH species in the study region.

CONCLUSION

The elimination of STH is a worthy goal, given the distress and disability they cause to a large portion of the global population. The goal is attainable, but will not be easily reached. Monitoring and evaluation of progress is critical and depends on highly accurate reporting. Species-specific repeat-based target discovery and qPCR assays deliver this accuracy. Microscopy and genera-specific molecular assays have a place in this effort, especially in surveys where full mapping of parasite diversity has not occurred. These tools, used in conjunction with highly sensitive and specific molecular assays targeting repetitive elements, can give a clear and accurate assessment where one tool alone could yield misleading results.

New target development is fairly easy, since the web-based bioinformatics tool RepeatExplorer provides output in a manner that makes finding repeat-based qPCR targets straightforward. Only a skim of the genome is necessary, so the data needed to develop new assays is already available for many species and is fairly inexpensive to produce for a species whose genome has not yet been sequenced. We have used this technique to explore new diagnostics for other nematode and protist parasites, and we think that the pipeline for assay development has potential for improving diagnostic sensitivity for many other classes of infectious agents.

Correct identification of species is of interest if one wants to understand evolution, biogeography, and emerging disease. The treatment for infection with one species of helminth is often the same as for another; however, we would argue that lumping all species of a genus together is careless. It is not clear that different species respond the same way under the same drug treatment, and misidentification of species in the past might be confounding the results of studies of resistance or drug response based on microscopy alone. Different species within a genus may also vary

in their capacity for animal infection, resulting in some parasites having animal reservoirs while others remain obligate human pathogens. A more detailed understanding of the underlying community structure will offer crucial insight into subjects, such as antihelminthic resistance, emerging or zoonotic diseases, and optimal threshold levels for elimination of disease.

BIOINFORMATIC RESOURCES

RepeatExplorer (Novák et al., 2010; Novák et al., 2013) is made freely available here:

<http://repeatexplorer.org/>

We make our scripts available for others to adapt and use here:

<https://github.com/JessicaGrant/RepeatTargetScripts>

AUTHOR CONTRIBUTIONS

Conceptualization: JG, NP, and SW. Funding acquisition: SW. Writing—original draft: JG. Writing—review and editing: JG, NP, SW.

REFERENCES

- Acosta Soto, L., Santísima-Trinidad, A. B., Bornay-Llinares, F. J., Martín González, M., Pascual Valero, J. A., and Ros Muñoz, M. (2017). Quantitative PCR and Digital PCR for Detection of *Ascaris lumbricoides* Eggs in Reclaimed Water. *BioMed Res. Int.* 2017, 1–9. doi: 10.1155/2017/7515409
- Andersen, L. O., Röser, D., Nejsun, P., Nielsen, H. V., and Stensvold, C. R. (2013). Is Supplementary bead beating for DNA extraction from nematode eggs by use of the NucliSENS easyMag protocol necessary? *J. Clin. Microbiol.* 51, 1345–1347. doi: 10.1128/JCM.03353-12
- Areekul, P., Putaporntip, C., Pattanawong, U., Sitthicharoenchai, P., and Jongwutiwes, S. (2010). *Trichuris vulpis* and *T. trichiura* infections among school children of a rural community in northwestern Thailand: the possible role of dogs in disease transmission. *Asian Biomed.* 4, 49–60. doi: 10.2478/abm-2010-0006
- Ásbjörnsdóttir, K. H., Means, A. R., Werkman, M., and Walson, J. L. (2017). Prospects for elimination of soil-transmitted helminths. *Curr. Opin. Infect. Dis.* 30, 482. doi: 10.1097/QCO.0000000000000395
- Ásbjörnsdóttir, K., Ajjampur, SSR, Anderson, R., Bailey, R., Gardiner, I., Halliday, K., et al. (2018). Assessing the feasibility of interrupting the transmission of soil-transmitted helminths through mass drug administration: The DeWorm3 cluster randomized trial protocol. *PLoS Negl. Trop. Dis.* 12(1):e0006166. doi: 10.1371/journal.pntd.0006166. eCollection
- Bartsch, S. M., Hotez, P. J., Asti, L., Zapf, K. M., Bottazzi, M. E., Diemert, D. J., et al. (2016). The Global Economic and Health Burden of Human Hookworm Infection. *PLoS Negl. Trop. Dis.* 10, e0004922. doi: 10.1371/journal.pntd.0004922
- Beaver, P. C., and Martin, L. K. (1968). Evaluation of Kato thick-smear technique for quantitative diagnosis of helminth infections *. *Am. J. Trop. Med. Hyg.* 17, 382–391. doi: 10.4269/ajtmh.1968.17.382
- Becker, S. L., et al. (2018). Toward the 2020 goal of soil-transmitted helminthiasis control and elimination. *PLoS Negl. Trop. Dis.* 12, e0006606. doi: 10.1371/journal.pntd.0006606
- Benjamin-Chung, J., Pilote, N., Ercumen, A., Grant, J. R., Maasch, J. R. M. A., Gonzalez, A. M., et al. (2019). Comparison of multi-parallel qPCR and Kato-Katz for detection of soil-transmitted helminth infection among children in rural Bangladesh. *bioRxiv* 629501. doi: 10.1101/629501
- Bik, H. M., Fournier, D., Sung, W., Bergeron, R. D., and Thomas, W. K. (2013). Intra-genomic variation in the ribosomal repeats of nematodes. *PLOS ONE* 8, e78230. doi: 10.1371/journal.pone.0078230

FUNDING

This work received financial support from the Coalition for Operational Research on Neglected Tropical Diseases (CORNTD), which is funded at The Task Force for Global Health primarily by the Bill & Melinda Gates Foundation, by the UK aid from the British government, and by the United States Agency for International Development through its Neglected Tropical Diseases Program. In addition, this work was partially funded through a grant to the DeWorm3 Project, which is funded by a grant to the Natural History Museum from the Bill and Melinda Gates Foundation.

ACKNOWLEDGMENTS

The authors would like to thank Marina Papaikovou for helpful conversations, as well as Dr. Judd Walson and Dr. D. Timothy J. Littlewood (Ásbjörnsdóttir et al., 2018). We also thank Dr. Eric Ottesen and Dr. Patrick Lammie (Task Force for Global Health, Decatur, GA, USA) for their ongoing support and advice.

- Biscotti, M. A., Olmo, E., and Heslop-Harrison, J. S. (Pat). (2015). Repetitive DNA in eukaryotic genomes. *Chromosome Res.* 23, 415–420. doi: 10.1007/s10577-015-9499-z
- Brindley, P. J., Laha, T., McManus, D. P., and Loukas, A. (2003). Mobile genetic elements colonizing the genomes of metazoan parasites. *Trends Parasitol.* 19, 79–87. doi: 10.1016/S1471-4922(02)00061-2
- Buonfrate, D., Formenti, F., Perandin, F., and Bisoffi, Z. (2015). Novel approaches to the diagnosis of Strongyloides stercoralis infection. *Clin. Microbiol. Infect.* 21, 543–552. doi: 10.1016/j.cmi.2015.04.001
- Chanteau, S., Luquiaud, P., Failloux, A. B., and Williams, S. A. (1994). Detection of *Wuchereria bancrofti* larvae in pools of mosquitoes by the polymerase chain reaction. *Trans. R. Soc. Trop. Med. Hyg.* 88, 665–666. doi: 10.1016/0035-9203(94)90219-4
- Charlesworth, B., Sniegowski, P., and Stephan, W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371, 215. doi: 10.1038/371215a0
- Chidambaram, M., Parija, S. C., Toi, P. C., Mandal, J., Sankaramoorthy, D., George, S., et al. (2017). Evaluation of the utility of conventional polymerase chain reaction for detection and species differentiation in human hookworm infections. *Trop. Parasitol.* 7, 111. doi: 10.4103/tp.TP_26_17
- Children Without Worms. (2018). STH Coalition Action Group Meeting 2018 Report. Retrieved April 2019 from http://www.childrenwithoutworms.org/sites/default/files/2018%20Action%20Group%20Meeting%20Report_FINAL%2007.16.18.pdf
- Clarke, N. E., Llewellyn, S., Traub, R. J., McCarthy, J., Richardson, A., Nery, S. V. (2018). Quantitative polymerase chain reaction for diagnosis of soil-transmitted helminth infections: a comparison with a flotation-based technique and an investigation of variability in DNA detection. *Am. J. Trop. Med. Hyg.* 99, 1033–1040. doi: 10.4269/ajtmh.18-0356
- Coghlan, A., Tyagi, R., Cotton, J. A., Holroyd, N., Rosa, B. A., Tsai, I. J., et al. (2019). Comparative genomics of the major parasitic worms. *Nat. Genet.* 51, 163–174. doi: 10.1038/s41588-018-0262-1
- Cringoli, G., Rinaldi, L., Maurelli, M. P., and Utzinger, J. (2010). FLOTAC: new multivalent techniques for qualitative and quantitative copromicroscopic diagnosis of parasites in animals and humans. *Nat. Protoc.* 5, 503–515. doi: 10.1038/nprot.2009.235
- Cushion, M. T., Brubaker, R., Ash, L. R., and Orihel, T. C. (1990). *Atlas of Human Parasitology*. 3rd ed. Chicago: American Society of Clinical Pathologists. ISBN 0-89189-292-3. doi: 10.1111/j.1550-7408.1992.tb04458.x

- Demas, A., Oberstaller, J., DeBarry, J., Lucchi, N. W., Srinivasamoorthy, G., Sumari, et al. (2011). Applied genomics: data mining reveals species-specific malaria diagnostic targets more sensitive than 18S rRNA. *J. Clin. Microbiol.* 49, 2411–2418. doi: 10.1128/JCM.02603-10
- Easton, A. V., Oliveira, R. G., O'Connell, E.M., Kepha, S., Mwandawiro, C. S., Njenga, S. M., et al. (2016). Multi-parallel qPCR provides increased sensitivity and diagnostic breadth for gastrointestinal parasites of humans: field-based inferences on the impact of mass deworming. *Parasit. Vectors* 9, 38. doi: 10.1186/s13071-016-1314-y
- Easton, A. V., Oliveira, R. G., Walker, M., O'Connell, E. M., Njenga, S. M., Mwandawiro, C. S., et al. (2017). Sources of variability in the measurement of *Ascaris lumbricoides* infection intensity by Kato-Katz and qPCR. *Parasit. Vectors* 10, 256. doi: 10.1186/s13071-017-2164-y
- Farrell, S. H., and Anderson, R. M. (2018). Helminth lifespan interacts with non-compliance in reducing the effectiveness of anthelmintic treatment. *Parasit. Vectors* 11, 66. doi: 10.1186/s13071-018-2670-6
- Fischer, K., Gankpala, A., Gankpala, L., Bolay, F.K., Curtis, K.C., Weil, G.J., et al. (2018). Capillaria Ova and Diagnosis of *Trichuris trichiura* Infection in Humans by Kato-Katz Smear, Liberia. *Emerg. Infect. Dis.* 24, 1551. doi: 10.3201/eid2408.180184
- Flavell, R. B., Bennett, M. D., Smith, J. B., and Smith, D. B. (1974). Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochem. Genet.* 12, 257–269. doi: 10.1007/BF00485947
- Funakoshi, K., Bagheri, M., Zhou, M., Suzuki, R., Abe, H., Akashi, H. (2017). Highly sensitive and specific Alu-based quantification of human cells among rodent cells. *Sci. Rep.* 7, 13202. doi: 10.1038/s41598-017-13402-3
- Ganley, A. R. D., and Kobayashi, T. (2007). Highly efficient concerted evolution in the ribosomal DNA repeats: total rDNA repeat variation revealed by whole-genome shotgun sequence data. *Genome Res.* 17, 184–191. doi: 10.1101/gr.5457707
- George, S., Kaliappan, S. P., Kattula, D., Roy, S., Geldhof, P., Kang, G., et al. (2015). Identification of *Ancylostoma ceylanicum* in children from a tribal community in Tamil Nadu, India using a semi-nested PCR-RFLP tool. *Trans. R. Soc. Trop. Med. Hyg.* 109, 283–285. doi: 10.1093/trstmh/trv001
- Halfon, P., Sylla, K., Ndiaye, J. L., Delaunay, P., Gaye, O., Faye, B., et al. (2017). Performance of Real-Time Polymerase Chain Reaction Assays for the Detection of 20 Gastrointestinal Parasites in Clinical Samples from Senegal. *Am. J. Trop. Med. Hyg.* 97, 173–182. https://doi.org/10.4269/ajtmh.16-0781
- Hamburger, J., He-Na, Abbasi, I., Ramzy, R. M., Jourdane, J., Ruppel, A. (2001). Polymerase chain reaction assay based on a highly repeated sequence of *Schistosoma haematobium*: a potential tool for monitoring schistosome-infested water. *Am. J. Trop. Med. Hyg.* 65, 907–911.
- Harmon, A. F., Zarlenga, D. S., and Hildreth, M. B. (2006). Improved methods for isolating DNA from *Ostertagia ostertagi* eggs in cattle feces. *Vet. Parasitol.* 135, 297–302. doi: 10.1016/j.vetpar.2005.10.014
- Hicks, J. H., Kremer, M., and Miguel, E. (2015). The Case for Mass Treatment of Intestinal Helminths in Endemic Areas. *PLoS Negl. Trop. Dis.* 9, e0004214. doi: 10.1371/journal.pntd.0004214
- Hidalgo, A., Melo, A., Romero, F., Hidalgo, V., Villanueva, J., Fonseca-Salamanca, F. (2018). DNA extraction in *Echinococcus granulosus* and *Taenia* spp. eggs in dogs stool samples applying thermal shock. *Exp. Parasitol.* 186, 10–16. doi: 10.1016/j.exppara.2018.01.016
- Hii, S.E., Senevirathna, D., Llewellyn, S., Inpankaew, T., Odermatt, P., Khieu, V., et al. (2018). Development and Evaluation of a Multiplex Quantitative Real-Time Polymerase Chain Reaction for Hookworm Species in Human Stool. *Am. J. Trop. Med. Hyg.* 99, 1186–1193. doi: 10.4269/ajtmh.18-0276
- Holt, D., Shield, J., Harris, T., Mounsey, K., Aland, K., McCarthy, J., et al. (2017). Soil-Transmitted Helminths in Children in a Remote Aboriginal Community in the Northern Territory: Hookworm is Rare but Strongyloides stercoralis and *Trichuris trichiura* Persist. *Trop. Med. Infect. Dis.* 2, 51. doi: 10.3390/tropicalmed2040051
- Howe, K. L., Bolt, B. J., Cain, S., Chan, J., Chen, W. J., Davis, P., et al. (2016). WormBase 2016: expanding to enable helminth genomic research. *Nucleic Acids Res.* 44, D774–D780. doi: 10.1093/nar/gkv1217
- Irisarri-Gutiérrez, M. J., Muñoz-Antolí, C., Acosta, L., Parker, L.A., Toledo, R., Bornay-Llinares, F.J., et al. (2016). Hookworm-like eggs in children's faecal samples from a rural area of Rwanda. *Afr. Health Sci.* 16, 83–88. https://doi.org/10.4314/ahs.v16i1.11
- Kaiser, M.M.M., Brienen, E. a. T., Djuardi, Y., Sartono, E., Yazdanbakhsh, M., Verweij, J. J., et al. (2017). Improved diagnosis of *Trichuris trichiura* by using a bead-beating procedure on ethanol preserved stool samples prior to DNA isolation and the performance of multiplex real-time PCR for intestinal parasites. *Parasitology* 144, 965–974. doi: 10.1017/S0031182017000129
- Katz, N., Chaves, A., and Pellegrino, J. (1972). A simple device for quantitative stool thick-smear technique in *Schistosomiasis mansoni*. *Rev. Inst. Med. Trop. Sao Paulo* 14, 397–400.
- Knopp, S., Salim, N., Schindler, T., Voules, D. A. K., Rothen, J., Lweno, O., et al. (2014). Diagnostic accuracy of Kato-Katz, FLOTAC, Baermann, and PCR methods for the detection of light-intensity hookworm and Strongyloides stercoralis infections in Tanzania. *Am. J. Trop. Med. Hyg.* 90, 535–545. doi: 10.4269/ajtmh.13-0268
- Kongs, A., Marks, G., Verlé, P., and Stuyf, P. V. der (2001). The unreliability of the Kato-Katz technique limits its usefulness for evaluating *S. mansoni* infections. *Trop. Med. Int. Health* 6, 163–169. doi: 10.1046/j.1365-3156.2001.00687.x
- Krauth, S. J., Coulibaly, J. T., Knopp, S., Traoré, M., N'Goran, E. K., Utzinger, J. (2012). An In-Depth Analysis of a Piece of Shit: Distribution of *Schistosoma mansoni* and Hookworm Eggs in Human Stool. *PLoS Negl. Trop. Dis.* 6, e1969. doi: 10.1371/journal.pntd.0001969
- Kyu, H. H., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., et al. (2018). Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 392, 1859–1922. doi: 10.1016/S0140-6736(18)32335-3
- Lafontaine, D. L. J., and Tollervey, D. (2001). The function and synthesis of ribosomes. *Nat. Rev. Mol. Cell Biol.* 2, 514. doi: 10.1038/35080045
- Lim, M. D., Brooker, S. J., Belizario, V. Y., Gay-Andrieu, F., Gilleard, J., Levecke, B., et al., Annecy STH diagnostic experts group (2018). Diagnostic tools for soil-transmitted helminths control and elimination programs: A pathway for diagnostic product development. *PLoS Negl. Trop. Dis.* 12, e0006213. doi: 10.1371/journal.pntd.0006213
- Lindner, D. L., and Banik, M. T. (2011). Intragenomic variation in the ITS rDNA region obscures phylogenetic relationships and inflates estimates of operational taxonomic units in genus *Laetiporus*. *Mycologia* 103, 731–740. doi: 10.3852/10-331
- Llewellyn, S., Inpankaew, T., Nery, S. V., Gray, D. J., Verweij, J. J., Clements, A. C. A., et al. (2016). Application of a Multiplex Quantitative PCR to Assess Prevalence and Intensity Of Intestinal Parasite Infections in a Controlled Clinical Trial. *PLoS Negl. Trop. Dis.* 10. doi: 10.1371/journal.pntd.0004380
- Lodh, N., Caro, R., Sofer, S., Scott, A., Krolewiecki, A., Shiff, C. (2016). Diagnosis of Strongyloides stercoralis : Detection of parasite-derived DNA in urine. *Acta Trop.* 163, 9–13. doi: 10.1016/j.actatropica.2016.07.014
- López-Flores, I., and Garrido-Ramos, M. A. (2012). “The repetitive DNA content of eukaryotic genomes,” in *Genome Dynamics*, vol. 7. Eds. Garrido-Ramos and M. A. (S. KARGER AG), 1–28. doi: 10.1159/000337118
- Mascagni, F., Vangelisti, A., Giordani, T., Cavallini, A., and Natali, L. (2018). Specific LTR-retrotransposons show copy number variations between wild and cultivated sunflowers. *Genes* 9(9), 433. doi: 10.3390/genes9090433
- Masny, A., Salamatin, R., Rozej-Bielicka, W., and Golab, E. (2016). Is molecular xenomonitoring of mosquitoes for *Dirofilaria repens* suitable for dirofilariasis surveillance in endemic regions? *Parasitol. Res.* 115, 511–525. doi: 10.1007/s00436-015-4767-6
- Maurelli, M. P., Rinaldi, L., Alfano, S., Pepe, P., Coles, G. C., Cringoli, G. (2014). Mini-FLOTAC, a new tool for copromicroscopic diagnosis of common intestinal nematodes in dogs. *Parasit. Vectors* 7, 356. doi: 10.1186/1756-3305-7-356
- McReynolds, L. A., DeSimone, S. M., and Williams, S. A. (1986). Cloning and comparison of repeated DNA sequences from the human filarial parasite *Brugia malayi* and the animal parasite *Brugia pahangi*. *Proc. Natl. Acad. Sci. U.S.A.* 83, 797–801. doi: 10.1073/pnas.83.3.797
- Medley, G. F., Turner, H. C., Baggaley, R. F., Holland, C., and Hollingsworth, T. D. (2016). “Chapter Six—The Role of More Sensitive Helminth Diagnostics in Mass Drug Administration Campaigns: Elimination and Health Impacts,” in *Advances in Parasitology*, vol. 94. Eds. Basáñez, M. G. , Anderson, and R. M. (Academic Press), 343–392. doi: 10.1016/bs.apar.2016.08.005
- Meredith, S. E., Unnasch, T. R., Karam, M., Piessens, W. F., and Wirth, D. F. (1989). Cloning and characterization of an *Onchocerca volvulus* specific DNA sequence. *Mol. Biochem. Parasitol.* 36, 1–10. doi: 10.1016/0166-6851(89)90194-1

- Mines, J. (1977). Modifications of McMaster worm egg counting method. *Aust. Vet. J.* 53, 342–343. doi: 10.1111/j.1751-0813.1977.tb00247.x
- Mohd-Shaharuddin, N., Lim, Y. A. L., Hassan, N.-A., Nathan, S., and Ngui, R. (2019). Molecular characterization of *Trichuris* species isolated from humans, dogs and cats in a rural community in Peninsular Malaysia. *Acta Trop.* 190, 269–272. doi: 10.1016/j.actatropica.2018.11.026
- Moser, W., Bärenbold, O., Mirams, G.J., Cools, P., Vlamincx, J., Ali, S.M., et al. (2018). Diagnostic comparison between FECPAKG2 and the Kato-Katz method for analyzing soil-transmitted helminth eggs in stool. *PLoS Negl. Trop. Dis.* 12, e0006562. doi: 10.1371/journal.pntd.0006562
- Nekrutenko, A., Makova, K. D., and Baker, R. J. (2000). Isolation of binary species-specific PCR-based markers and their value for diagnostic applications. *Gene* 249, 47–51. doi: 10.1016/S0378-1119(00)00168-2
- Nikolay, B., Brooker, S. J., and Pullan, R. L. (2014). Sensitivity of diagnostic tests for human soil-transmitted helminth infections: a meta-analysis in the absence of a true gold standard. *Int. J. Parasitol.* 44, 765–774. doi: 10.1016/j.ijpara.2014.05.009
- Ng-Nguyen, D., Stevenson, M. A., Dorny, P., Gabriël, S., Vo, T. V., Nguyen, V.-A.T., et al. (2017). Comparison of a new multiplex real-time PCR with the Kato Katz thick smear and copro-antigen ELISA for the detection and differentiation of *Taenia* spp. in human stools. *PLoS Negl. Trop. Dis.* 11, e0005743. doi:10.1371/journal.pntd.0005743
- Novák, P., Neumann, P., and Macas, J. (2010). Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC Bioinformatics* 11, 378. doi: 10.1186/1471-2105-11-378
- Novák, P., Neumann, P., Pech, J., Steinhaisl, J., and Macas, J. (2013). RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* 29, 792–793. doi: 10.1093/bioinformatics/btt054
- Novák, P., Ávila Robledillo, L., Koblížková, A., Vrbová, I., Neumann, P., Macas, J. (2017). TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Res.* 45, e111. doi: 10.1093/nar/gkx257
- Nunes, C.M., Lima, L.G.F., Manoel, C.S., Pereira, R.N., Nakano, M.M., and Garcia, J.F. (2006). Fecal specimens preparation methods for PCR diagnosis of human taeniosis. *Rev. Inst. Med. Trop. de São Paulo* 48(1), 45–47. doi: 10.1590/S0036-46652006000100010
- O'Connell, E. M., et al. (2018). *Ancylostoma ceylanicum* Hookworm in Myanmar Refugees, Thailand, 2012–2015. *Emerg. Infect. Dis.* 24. doi: 10.3201/eid2408.180280
- Okorie, P. N., and de Souza, D. K. (2016). Prospects, drawbacks and future needs of xenomonitoring for the endpoint evaluation of lymphatic filariasis elimination programs in Africa. *Trans. R. Soc. Trop. Med. Hyg.* 110, 90–97. doi: 10.1093/trstmh/trv104
- Onyabe, D. Y., and Conn, J. E. (1999). Intragenomic heterogeneity of a ribosomal DNA spacer (ITS2) varies regionally in the neotropical malaria vector *Anopheles numeztovari* (Diptera: Culicidae). *Insect Mol. Biol.* 8, 435–442. doi: 10.1046/j.1365-2583.1999.00134.x
- Papaiakovou, M., Pilotte, N., Grant, J.R., Traub, R.J., Llewellyn, S., McCarthy, et al. (2017). A novel, species-specific, real-time PCR assay for the detection of the emerging zoonotic parasite *Ancylostoma ceylanicum* in human stool. *PLoS Negl. Trop. Dis.* 11, e0005734. doi: 10.1371/journal.pntd.0005734
- Papaiakovou, M., Pilotte, N., Baumer, B., Grant, J., Asbjornsdottir, K., Schaer, E., et al. (2018). A comparative analysis of preservation techniques for the optimal molecular detection of hookworm DNA in a human fecal specimen. *PLoS Negl. Trop. Dis.* 12, e0006130. doi: 10.1371/journal.pntd.0006130
- Pickering, A.J., Njenga, S.M., Steinbaum, L., Swarthout, J., Lin, A., Arnold, B.F., et al. (2019). Effects of single and integrated water, sanitation, handwashing, and nutrition interventions on child soil-transmitted helminth and *Giardia* infections: A cluster-randomized controlled trial in rural Kenya. *PLoS Med.* 16, e1002841. doi: 10.1371/journal.pmed.1002841
- Piednoël, M., Aberer, A.J., Schneeweiss, G.M., Macas, J., Novak, P., Gundlach, H., et al. (2012). Next-generation sequencing reveals the impact of repetitive DNA across phylogenetically closely related genomes of orobanchaceae. *Mol. Biol. Evol.* 29, 3601–3611. doi: 10.1093/molbev/mss168
- Pilotte, N., Papaiakovou, M., Grant, J.R., Bierwert, L.A., Llewellyn, S., McCarthy, J.S., et al. (2016a). Improved PCR-based detection of soil transmitted helminth infections using a next-generation sequencing approach to assay design. *PLoS Negl. Trop. Dis.* 10, e0004578. doi: 10.1371/journal.pntd.0004578
- Pilotte, N., Unnasch, T. R., and Williams, S. A. (2017). The current status of molecular xenomonitoring for lymphatic filariasis and onchocerciasis. *Trends Parasitol.* 33, 788–798. doi: 10.1016/j.pt.2017.06.008
- Pilotte, N., Zaky, W. I., Abrams, B. P., Chadee, D. D., and Williams, S. A. (2016b). A novel xenomonitoring technique using mosquito excreta/feces for the detection of filarial parasites and malaria. *PLoS Negl. Trop. Dis.* 10, e0004641. doi: 10.1371/journal.pntd.0004641
- Plohl, M., Luchetti, A., Mestrovic, N., and Mantovani, B. (2008). Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* 409, 72–82. doi: 10.1016/j.gene.2007.11.013
- Pontes, L. A., Oliveira, M. C., Katz, N., Dias-Neto, E., and Rabello, A. (2003). Comparison of a polymerase chain reaction and the Kato-Katz technique for diagnosing infection with *Schistosoma mansoni*. *Am. J. Trop. Med. Hyg.* 68, 652–656. doi: 10.4269/ajtmh.2003.68.652
- Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., et al. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 35, 7188–7196. doi: 10.1093/nar/gkm864
- Ralph, A., O'Sullivan, M. V.N., Sangster, N. C., and Walker, J. C. (2006). Abdominal pain and eosinophilia in suburban goat keepers—trichostrongylosis. *Med. J. Aust.* 184, 467–469.
- Ramirez, A.L., Hall-Mendelin, S., Doggett, S.L., Hewitson, G.R., McMahon, J.L., Ritchie, S.A., et al. (2018). Mosquito excreta: a sample type with many potential applications for the investigation of Ross River virus and West Nile virus ecology. *PLoS Negl. Trop. Dis.* 12, e0006771. doi: 10.1371/journal.pntd.0006771
- Rao, R. U., Weil, G. J., Fischer, K., Supali, T., and Fischer, P. (2006). Detection of *Brugia* parasite DNA in human blood by real-time PCR. *J. Clin. Microbiol.* 44, 3887–3893. doi: 10.1128/JCM.00969-06
- Rao, U. R., Williams, S. A., and Klei, T. R. (2002). Quantification of PCR amplification products of *Brugia* HhaI repeat DNA using a semiautomated Q-PCR system. *Mol. Cell. Probes* 16, 13–23. doi: 10.1006/mcpr.2001.0393
- Sanpool, O., Intapan, P.M., Thanchomnang, T., Sri-Aroon, P., Lulitanond, V., Sadaow, L., et al. (2012). Development of a real-time PCR assay with fluorophore-labelled hybridization probes for detection of *Schistosoma mekongi* in infected snails and rat feces. *Parasitology* 139, 1266–1272. doi: 10.1017/S0031182012000649
- Sarhan, R. M., Kamel, H. H., Saad, G. A., and Ahmed, O. A. (2015). Evaluation of three extraction methods for molecular detection of *Schistosoma mansoni* infection in human urine and serum samples. *J. Parasit. Dis. Off. Organ Indian Soc. Parasitol.* 39, 499–507. doi: 10.1007/s12639-013-0385-3
- Schaap, M., Lemmers, R.J., Maassen, R., van der Vliet, P.J., Hoogerheide, L.F., van Dijk, H.K., et al. (2013). Genome-wide analysis of macrosatellite repeat copy number variation in worldwide populations: evidence for differences and commonalities in size distributions and size restrictions. *BMC Genomics* 14, 143. doi: 10.1186/1471-2164-14-143
- Shapiro, J. A., and Sternberg, R. (2005). Why repetitive DNA is essential to genome function. *Biol. Rev.* 80, 227–250. doi: 10.1017/S1464793104006657
- Speich, B., Ali, S.M., Ame, S.M., Albonico, M., Utzinger, J., Keiser, J. (2015). Quality control in the diagnosis of *Trichuris trichiura* and *Ascaris lumbricoides* using the Kato-Katz technique: experience from three randomised controlled trials. *Parasit. Vectors* 8, 82. doi: 10.1186/s13071-015-0702-z
- Stensvold, C.R., Saijuntha, W., Sithithaworn, P., Wongratanacheewin, S., Strandgaard, H., Ørnberg, N., et al. (2006). Evaluation of PCR based coprodiagnosis of human opisthorchiasis. *Acta Trop.* 97, 26–30. doi: 10.1016/j.actatropica.2005.08.008
- Subirana, J. A., and Messegue, X. (2013). A satellite explosion in the genome of holocentric nematodes. *PLoS One* 8, e62221. doi: 10.1371/journal.pone.0062221
- Subirana, J. A., Albà, M. M., and Messegue, X. (2015). High evolutionary turnover of satellite families in *Caenorhabditis*. *BMC Evol. Biol.* 15, 218. doi: 10.1186/s12862-015-0495-x
- Teruel, M., Ruiz-Ruano, F.J., Marchal, J.A., Sánchez, A., Cabrero, J., Camacho, J.P., et al. (2014). Disparate molecular evolution of two types of repetitive DNAs in the genome of the grasshopper *Eyprepocnemis plorans*. *Heredity* 112, 531–542. doi: 10.1038/hdy.2013.135
- Truscott, J. E., Turner, H. C., Farrell, S. H., and Anderson, R. M. (2016). Soil-transmitted helminths: mathematical models of transmission, the impact of mass drug administration and transmission elimination criteria. *Adv. Parasitol.* 94, 133–198. doi: 10.1016/bs.apar.2016.08.002

- Truscott, J. E., Werkman, M., Wright, J.E., Farrell, S.H., Sarkar, R., Ásbjörnsdóttir, K., et al. (2017). Identifying optimal threshold statistics for elimination of hookworm using a stochastic simulation model. *Parasit. Vectors* 10, 321. doi: 10.1186/s13071-017-2256-8
- Truscott, J. E., Dunn, J. C., Papaikovou, M., Schaer, F., Werkman, M., Littlewood, D. T. J., et al. (2019). Calculating the prevalence of soil-transmitted helminth infection through pooling of stool samples: choosing and optimizing the pooling strategy. *PLoS Negl. Trop. Dis.* 13, e0007196. doi: 10.1371/journal.pntd.0007196
- Umemori, J., Mori, A., Ichiyanagi, K., Uno, T., and Koide, T. (2013). Identification of both copy number variation-type and constant-type core elements in a large segmental duplication region of the mouse genome. *BMC Genomics* 14, 455. doi: 10.1186/1471-2164-14-455
- Uniting to Combat NTDs. (2012) London declaration on neglected tropical diseases. Available at: <https://unitingtocombatntds.org/london-declaration-neglected-tropical-diseases/> (Accessed April 9, 2019).
- Verweij, J. J., and Stensvold, C. R. (2014). Molecular testing for clinical diagnosis and epidemiological investigations of intestinal parasitic infections. *Clin. Microbiol. Rev.* 27, 371–418. doi: 10.1128/CMR.00122-13
- Weatherhead, J. E., Hotez, P. J., and Mejia, R. (2017). The global state of helminth control and elimination in children. *Pediatr. Clin.* 64, 867–877. doi: 10.1016/j.pcl.2017.03.005
- Werneck, J. S., Carniato, T., Gabriel, A., Tufik, S., and Andrade, S. S. (2007). Mites in clinical stool specimens: potential misidentification as helminth eggs. *Trans. R. Soc. Trop. Med. Hyg.* 101, 1154–1156. doi: 10.1016/j.trstmh.2007.07.006
- Wickstead, B., Ersfeld, K., and Gull, K. (2003). Repetitive elements in genomes of parasitic protozoa. *Microbiol. Mol. Biol. Rev.* 67, 360–375. doi: 10.1128/MMBR.67.3.360-375.2003
- Williams, S. A., Lizotte-Waniewski, M. R., Foster, J., Guiliano, D., Daub, J., Scott, A.L., et al. (2000). The filarial genome project: analysis of the nuclear, mitochondrial and endosymbiont genomes of *Brugia malayi*. *Int. J. Parasitol.* 30, 411–419. doi: 10.1016/S0020-7519(00)00014-X
- Yong, T.-S., Lee, J.-H., Sim, S., Lee, J., Min, D.-Y., Chai, J.-L., et al. (2007). Differential diagnosis of *Trichostrongylus* and hookworm eggs via PCR using ITS-1 sequence. *Korean J. Parasitol.* 45, 69–74. doi: 10.3347/kjp.2007.45.1.69
- Zaky, W. I., Tomaino, F. R., Pilotte, N., Laney, S. J., and Williams, S. A. (2018). Backpack PCR: a point-of-collection diagnostic platform for the rapid detection of *Brugia* parasites in mosquitoes. *PLoS Negl. Trop. Dis.* 12, e0006962. doi: 10.1371/journal.pntd.0006962
- Zhao, S., and Gibbons, J. G. (2018). A population genomic characterization of copy number variation in the opportunistic fungal pathogen *Aspergillus fumigatus*. *PLoS One* 13, e0201611 doi: 10.1371/journal.pone.0201611.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Grant, Pilotte and Williams. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.