Measuring and mitigating inhibition during quantitative real time PCR analysis of viral nucleic acid extracts from large-volume environmental water samples

K.E. Gibson a,1, K.J. Schwaba, S.K. Spencerb, M.A. Borchardtb,*

aThe Johns Hopkins Bloomberg School of Public Health, Department of Environmental Health Sciences, Division of Environmental Health Engineering, and the Johns Hopkins Center for Water and Health, 615 North Wolfe Street, Baltimore, MD 21205, USA
bUSDA – Agricultural Research Service, Dairy Forage Research Center, Institute for Environmentally Integrated Dairy Management, 2615 Yellowstone Drive, Marshfield, WI 54449, USA

Abstract

Naturally-occurring inhibitory compounds are a major concern during qPCR and RT-qPCR analysis of environmental samples, particularly large volume water samples. Here, a standardized method for measuring and mitigating sample inhibition in environmental water concentrates is described. Specifically, the method 1) employs a commercially available standard RNA control; 2) defines inhibition by the change in the quantification cycle (C_q) of the standard RNA control when added to the sample concentrate; and 3) calculates a dilution factor using a mathematical formula applied to the change in C_q to indicate the specific volume of nuclease-free water necessary to dilute the effect of inhibitors. The standardized inhibition method was applied to 3,193 large-volume water (surface, groundwater, drinking water, agricultural runoff, sewage) concentrates of which 1,074 (34%) were inhibited. Inhibition level was not related to sample volume. Samples collected from the same locations over a one to two year period had widely variable inhibition levels. The proportion of samples that could have been reported as false negatives if inhibition had not been mitigated was between 0.3% and 71%, depending on water source. These findings emphasize the importance of measuring and mitigating inhibition when reporting qPCR results for viral pathogens in environmental waters to minimize the likelihood of reporting false negatives and under-quantifying virus concentration.

1. Introduction

The need for more accurate quantitative assessments of microbial water quality, in both drinking water and recreational water, is driven by the shift toward quantitative risk-based approaches used for making regulatory decisions on specific microbial contaminants. For example, the United States Environmental Protection Agency (USEPA) proposed revisions of the criteria for assessing recreational water quality (USEPA, 2010a) that call for quantitative risk assessments. These new criteria will likely depend on real-time, molecular detection and quantification of bacterial indicators (i.e. Enterococcus and Bacteroidales) to gauge health risks related to exposure to recreational waters (Wade et al., 2010).
Accurate quantification of waterborne pathogens is equally important for studies designed to determine the occurrence of pathogens as well as for establishing baseline levels of microorganisms in diverse water types within various environments (e.g. groundwater under the influence of surface water and agricultural runoff in to nearby surface waters). The results from these studies often contribute to decisions regarding water quality regulation (e.g. Safe Drinking Water Act, USEPA Contaminant Candidate List (CCL)), identify research needs, and encourage the development of improved sampling and detection methods.

Over the past 15 years, there has been increased focus on optimizing and applying methods for concentrating large volumes of water for quantifying waterborne viruses, primarily owing from the human enteric viruses (caliciviruses, adenovirus, and enterovirus) listed on all three CCLs published by the USEPA every 5 years since 1998 (USEPA, 2010b). The difficulties related to sampling and detecting enteric viruses in environmental water sources have been well established and stem primarily from the low, but significant, levels of human enteric viruses that may be present (Teunis et al., 2008), and the lack of a reproducible, quantitative culture assay for many enteric viruses of concern (e.g. norovirus, rotavirus) (Duizer et al., 2004). Thus, researchers have been pushed to improve detection methods in order to inform regulatory inclusion of enteric viruses in USEPA water quality standards (i.e., Safe Drinking Water Act). Some of the improved sampling methods for enteric viruses – and all classes of microorganisms – include tangential flow and dead-end ultrafiltration as well as electropositive filtration using NanoCeram® filter and glass wool filtration (Gibson and Schwab, 2011; Hill et al., 2005; Karim et al., 2009; Lambertini et al., 2008; Mull and Hill, 2009).

Molecular techniques such as real time qPCR and reverse transcription (RT) qPCR allow for the sensitive and specific detection and quantification of target nucleic acids. The application of these methods for detection of pathogens, however, can be hampered by the presence of molecular inhibitors. Inhibitory compounds can cause a shift, or however, can be hampered by the presence of molecular inhibitors present. In addition, their effectiveness during the RT quality cDNA for amplification during qPCR is not well known (Chandler et al., 1998; Staub et al., 1995). It is important to note that adding these compounds would require optimization of each qPCR assay, and the optimal concentration may still vary depending on the sample matrix and specific inhibitors present. In addition, their effectiveness during the RT phase where inhibitors can prevent the formation of high quality cDNA for amplification during qPCR is not well known (Chandler et al., 1998; Staub et al., 1995). Several tools are available for identifying inhibition during qPCR. These include the use of external controls and internal amplification controls – competitive or non-competitive – during qPCR (Hoorfar et al., 2004; Mackay et al., 2002). These inhibition standard controls may be 1) the same sequence as the DNA or RNA of the target viral genome (Hata et al., 2011); 2) plasmid DNA transcripts containing a recombinant sequence of the target viral genome (Scipioni et al., 2008); or 3) exogenous nucleic acid sequences that are not the same as the DNA or RNA of the target viral genome (Ninove et al., 2011; Stevenson et al., 2008). With respect to large volume water concentrates, researchers do not always report a method for identification of inhibition during molecular assays (Mull and Hill, 2009; Plutzer et al., 2010; Polaczyk et al., 2008; Shapiro et al., 2010). Studies reporting methods for identification of inhibition within qPCR of nucleic acid extracts from water samples primarily use external controls (Albinana-Gimenez et al., 2009; Gibson and Schwab, 2011; Hill et al., 2007, 2010; Knappett et al., 2010; Lambertini et al., 2008) or internal purification techniques (e.g., phenol-chloroform extraction, polyethylene glycol precipitation, nucleic acid extraction kits) to avoid co-concentrating inhibitors or to ablate isolated inhibitors.

Gel filtration resins like Sephadex or Sepharose or similar products are generally effective in removing inhibitors although the degree of inhibitor removal can vary widely among products (Borchardt et al., 2003; Hata et al., 2011; Miller, 2001; Wilson, 1997). Sometimes the most effective column is not the commercial purchase, but the one made by hand in the lab (Borchardt, unpublished data). Schriewer et al. (2011) recently applied Supelite™ DAX-8 non-ionic macroporous resin, which can irreversibly absorb humic acids, to 100 L surface water concentrates as well as prepared humic acid standards that were spiked with known amounts of DNA. The DAX-8 was found to reduce inhibition, but there also was loss of DNA during treatment with the resin.

PCR facilitators are used by some investigators to avoid the loss of nucleic acid that can result from chromatographic techniques or binding resins. Bovine serum albumin (BSA) added to PCR reaction mixtures has been shown to reduce inhibition by physically binding to the PCR inhibitory substances as well as serving as a target for substances (i.e. proteinases) that can degrade DNA polymerase (Oikarinen et al., 2009; Radstrom et al., 2004). Polyvinylpyrrolidone (PVP) has been reported to bind polyphenolics, compounds present in leaves and bark, that may be co-purified during nucleic acid extraction; however, the inclusion of PVP during nucleic acid extraction and/or within the reaction mix has shown conflicting results for alleviating inhibition (Jiang et al., 2005; Koonjul et al., 1999). The protein gp32, which binds single-stranded DNA, has been consistently reported to increase reaction efficiency for long PCR products and alleviate inhibition caused by humic acids (Kreader, 1996). It is important to note that adding these compounds would require optimization of each qPCR assay, and the optimal concentration may still vary depending on the sample matrix and specific inhibitors present. In addition, their effectiveness during the RT phase where inhibitors can prevent the formation of high quality cDNA for amplification during qPCR is not well known (Chandler et al., 1998; Staub et al., 1995).
amplification controls (Hata et al., 2011; Varela Villarreal et al., 2010).

Another approach for identifying inhibition is analyzing shifts in amplification efficiency for each qPCR sample. Examples include statistical and kinetic modeling approaches (e.g., sigmoidal curve fit model, linear regression of efficiency model, and modified standard curve) (Guescini et al., 2008; Rutledge, 2004; Rutledge and Stewart, 2008). Depending on the modeling approach chosen to analyze amplification efficiency, one may overestimate the actual qPCR efficiency while another may underestimate efficiency (Guescini et al., 2008). Very few studies involving molecular analysis of water sample concentrates have used shifts in amplification efficiency for identification and quantification of inhibition (Dubois et al., 2007; Rajal et al., 2007).

Overall, inhibition makes interpreting the public health risk of pathogens in environmental samples more problematic. When pathogens are present in these samples, the concentrations are typically low, and inhibition could result in underestimating exposure and consequently health risk. This may be particularly true for enteric viruses, which tend to have a low infectious dose (i.e. 10 to 100 virions) (Teunis et al., 2008). Viruses have also been shown by qPCR measurements to be significantly associated with an elevated risk of acute gastrointestinal illness at low concentrations in drinking water (i.e. 1 to 10 genomic copies per liter) (Borchardt et al., in press). Therefore, for recent methodological advances in water sampling and qPCR to contribute to a more accurate assessment of health risk from environmental exposures to pathogens, the issue of PCR inhibition cannot be ignored.

The research reported here describes a simple and effective method for measuring and mitigating sample inhibition during qPCR analysis of viral nucleic acid extracts from over 3,000 large volume water samples (surface water, groundwater, drinking water, agricultural runoff, and sewage). Sample inhibition is: 1) identified by using a commercially available standard RNA control; 2) measured by a specific change in quantification cycle \( C_q \) of the RNA standard when seeded into a sample; and 3) mitigated by using a mathematical formula applied to the change in \( C_q \) to indicate the specific volume of nuclease-free water necessary to dilute the effect of inhibitors. Moreover, for the 3,193 samples we report inhibition frequency, the effect of sample volume on inhibition, the change in inhibition at the same sample site over time, and the magnitude of false negatives if inhibition had not been mitigated.

2. Materials and methods

2.1. Study laboratories

For the methods and data presented herein, two separate laboratories combined data for their samples to demonstrate the results of applying the standardized method for measuring inhibition to a large number of samples. Methods for each lab are designated as either “Laboratory A” or “Laboratory B”. The major difference in methods between the two laboratories lies within the filtration and/or concentration of large-volume water samples. Lesser differences are described for the preparation of inhibition control RNA. Both laboratories relied on the same shift in the control \( C_q \) for defining inhibition level.

2.2. Measurement of sample inhibition

Real time qPCR inhibition was evaluated for every water sample by spiking a known amount of hepatitis G virus (HGV; also known as GB virus C) Armored RNA® (Asuragen, Austin, TX) into the PCR reaction mixture prior to the reverse transcription (RT) step. It was appropriate to evaluate inhibition encompassing both the RT and PCR steps because the majority of enteric viruses quantified in these samples are RNA viruses. When measuring inhibition in studies involving only DNA targets, Laboratory B spikes a known amount of HGV cDNA into PCR reaction mixture prior to amplification. The quantity of HGV measured in the unknown sample was compared to the quantity measured in corresponding HGV positive controls. Methods for the preparation and quantification of internal standard HGV RNA were performed by two laboratories and are described below. Each batch of samples assayed for inhibition included a negative control of HGV master mix containing diethylpyrocarbonate (DEPC)-treated water substituted for HGV RNA and at least 3 positive control reactions containing only HGV RNA and no sample.

2.3. Dilution factor formula

A sample was deemed inhibited if the quantification cycle \( C_q \) of the seeded HGV was higher by one cycle or more than the mean of the expected \( C_q \) obtained from the HGV positive controls. We considered a one cycle deviation indicative of inhibition given the precision of the \( C_q \) measurements of the HGV positive controls. For example, Laboratory A) mean \( C_q = 26.9, \) standard deviation = 0.3, range 26.2–27.3, \( n = 10 \); Laboratory B) mean \( C_q = 28.5, \) standard deviation = 0.3, range 28.2–29.0, \( n = 14 \). \( n \) refers to the number of \( C_q \) measurements from a single preparation of the HGV positive control; different HGV positive control preparations were used between the laboratories. \( C_q \) measurements of other HGV positive control preparations had similar levels of precision. The \( C_q \) of the seeded HGV and HGV positive control were determined by the same method, either fix point (Laboratory A) or second derivative maximum (Laboratory B). If the sample was inhibited (>1 cycle higher for the seeded HGV compared to the HGV positive controls), the difference in \( C_q \) values was used to calculate an appropriate dilution as follows:

\[
\text{Dilution factor} = 10^x,
\]

where \( x = \frac{\text{expected HGV } C_q - \text{measured HGV } C_q}{\text{standard curve slope}} \)

For example, if \( x = 0.845 \) the dilution factor = 7 and the nucleic acid from the unknown sample should be diluted 1:7 with nuclease-free water prior to analysis.

For both laboratories A and B the target HGV \( C_q \) was 28. HGV Armored RNA was quantified by the manufacturer and we made the appropriate dilution that when HGV was added to the one-step or two-step RT-qPCR mixture it would result in a \( C_q \) of 28 (i.e. approximately 800 copies HGV in the PCR
reaction volume). Establishing a Cq of 28 enabled an inhibition effect of a 7-cycle increase up to a Cq of 35 (i.e. approximately a 100-fold dilution) to be reliably determined.

Inhibition levels are reported here as dilution factors. For an ideal standard curve slope of $-3.3$, a one cycle shift upward in the measured HGV Cq translates to a dilution factor of 2. Therefore, samples with dilution factors $< 2$ are considered uninhibited while samples with a dilution factor $\geq 2$ are considered inhibited.

2.4. Nucleic acid extractions

The two laboratories used different methods for extracting HGV RNA: Laboratory A) RNA was extracted from 140 μL HGV Armored RNA® using QIAGEN® Viral RNA Mini Spin kit (Qiagen, Valencia, CA) following the manufacturer’s Spin Protocol; HGV RNA was eluted from the Qiangen spin column by performing a double elution using 2 $\times$ 40 μL of DEPC-treated water supplemented with 0.01% 500 U/μL RNase Inhibitor (Applied Biosystems, Carlsbad, CA). A similar method of nucleic acid extraction was also applied to concentrated water samples using a QIAGEN® MinElute® Virus Spin kit (Qiagen) to extract total nucleic acid (i.e. RNA and DNA) as opposed to just RNA. The water sample nucleic acid extraction was added to the RT-PCR master mix to which was also added a known amount of extracted HGV (Section 2.5.1). Laboratory B) HGV Armored RNA was seeded into DEPC-treated water for the positive control or into the extracted nucleic acid suspension from a concentrated water sample. The latter was obtained using the QIAamp DNA blood mini kit and buffer AVL (Qiagen). After seeding, HGV RNA was released following the manufacturer’s directions by heating at 99 °C for 4 min; thermal conditions for reverse transcription were begun immediately following HGV RNA release (Section 2.5.2).

2.5. Real time RT-qPCR assay for HGV

2.5.1. Laboratory A

HGV RNA was amplified by real time, one-step RT-qPCR using an ABI Prism 7300 Sequence Detection System (Applied Biosystems). Primers and probe for the HGV assay are as follows: HepG-F (5’ CGG CCA AAA GGT GGT GGA TG 3’), HepG-R (5’ CGA GCC TGA GTG CCG G 3’), HepG probe (FAM 5’ AGG TTC CTC TGG GCC TGT TGG CCA G 3’ BHQ-1) (Schlueter et al., 1996). Amplification was performed as described previously in Gibson and Schwab (2011). Briefly, each 25 μL reaction mixture contained 12.5 μL of 2 $\times$ master mix [QuantiTect Probe RT-PCR Kit (Qiagen)], 5 μL of prepared sample, 2 μL of a known amount of HGV RNA, and DEPC-treated water for the remaining volume. Real time RT-qPCR amplification for HGV was performed under the following conditions: reverse transcription for 30 min at 50 °C, then denaturation for 15 min at 95 °C, followed by 40 cycles of denaturation at 95 °C for 15 s and annealing/extension at 60 °C for 60 s.

2.5.2. Laboratory B

HGV RNA was amplified by two-step RT-qPCR where 8.6 μL of the extracted water sample RNA was combined with 8.6 μL Armored HGV RNA and 0.7 μL (0.007 μg/μL) random hexamers (Promega, Madison, WI) and heating for 4 min at 99 °C. Added to this mixture was 32.1 μL RT master mix containing the following components reported as final concentrations in the 50 μL reaction volume: 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 3 mM MgCl₂, 10 mM dithiothreitol, 70 μM of each deoxy-nucleoside triphosphate (Promega), 30 U of RNasin® Ribonuclease Inhibitor (Promega) and 100 U of SuperScript II reverse transcriptase (Invitrogen Life Technologies, Rockville, MD). Reaction incubation was at 25 °C for 15 min, 42 °C for 60 min, and 99 °C for 5 min and then held at 4 °C. Following RT, real-time qPCR was performed with the LightCycler 480 (Roche Diagnostics, Mannheim, Germany) using the master mix provided by the manufacturer (LightCycler 480 Probes Master kit, Roche Diagnostics). Six microliters of cDNA was added to 14 μL master mix containing HGV primers and dual-labeled TaqMan probe (TIB Molbiol, Berlin, Germany) (Schlueter et al., 1996) at final concentrations of 500 nm and 100 nm, respectively. Laboratory B used the same sequence for the primers and probe as described for Laboratory A. Thermocycling conditions were a hot start polymerase activation for 10 min at 95 °C followed by 45 cycles of 15 s at 94 °C and 1 min at 60 °C.

2.6. Concentration of water samples

Water samples included in the present study were concentrated using two separate methods described previously (Gibson and Schwab, 2011; Lamberti et al., 2008). Briefly, large-volume water samples from California, Indiana, Iowa, Maryland, Michigan, Ohio, Pennsylvania, Virginia and Washington as well as water samples around Accra, the capital city of Ghana, Africa were concentrated utilizing a tangential flow, hollow fiber ultrafiltration (UF) method. Sample volumes processed by UF ranged from 40 L to 100 L ($n = 75$). Water samples from the remaining studies including Water and Health Trial for Enteric Risks (WAHTER) Study sites, Madison, WI, Milwaukee, WI and United States Department of Agriculture (USDA) field research site located near Marshfield, WI were concentrated using the sodocalcic glass wool filtration method (Lamberti et al., 2008). Sample volumes processed by the glass wool filtration method ranged from 2 L to 110,886 L ($n = 3,118$) with mean and median volumes of 1,953 L and 753 L, respectively. Both laboratories performed secondary concentration procedures on the filtration concentrates by centrifugal filtration devices (Laboratory A) and polyethylene glycol precipitation (Laboratory B) as described previously in Gibson and Schwab (2011) and Lamberti et al. (2008), respectively.

2.7. Description of water samples

Water sample data pooled for inhibition data analysis were from five water source types (Table 1): 1) Groundwater samples were from household taps and 36 municipal wells drilled in sand/sandstone aquifers located in 14 communities in Wisconsin, USA that do not treat or disinfect their municipal water (i.e. WAHTER Study); 10 municipal wells in a deep confined sandstone aquifer supplying Madison, WI; four wells in Ghana and 10 wells in Washington; 2) Surface water
samples were from lakes Mendota, Monona, and Wingra in Madison, WI; the rivers Milwaukee, Menomonie, Cedar, Underwood, Kinnickinnic near Milwaukee, WI; rivers and lakes in Southern Ghana near Accra; rivers and irrigation canals in Yakima Valley, Washington; rivers from USGS study sites across the U.S. (Indiana, Iowa, Maryland, Pennsylvania, Ohio, Michigan, Virginia); a creek in Harford County, MD; and surface water used as source water for drinking water treatment plants in Southern California and Baltimore, Maryland; 3) Sewage samples were raw influent entering wastewater treatment plants in Madison and Milwaukee, Wisconsin, and four communities in Ohio, USA; 4) Agricultural runoff samples were from a single study involving four 1.6 ha fields (Withee silt loam) located near Marshfield, WI; 5) Drinking water samples (i.e., tap water derived from treated surface water) were from a research laboratory at Johns Hopkins Bloomberg School of Public Health in Baltimore, MD; two drinking water treatment plants, one each in California and Maryland; and community water kiosks located in rural villages near Accra, Ghana.

### 3. Results and Discussion

#### 3.1. Measuring qPCR inhibition and mitigation strategies

The number of samples inhibited and the level of inhibition as expressed by the calculated dilution factor (DF) are shown in Fig. 1 for each source water type. A similar pattern of inhibition is seen for each source water type — a high frequency of uninhibited samples (DF < 2), a low frequency of samples with partial inhibition, and a high frequency of samples with extreme inhibition resulting in a DF ≥ 100 in the qPCR assay. Extremely inhibited samples are grouped in the category DF ≥ 100; even though the calculated dilution factor may have been greater than 100, we kept the actual dilution at 1:100 to minimize the possibility of diluting the virus target below the limit of detection. Of the 3,193 samples analyzed, 1,074 (34%) were considered inhibited (DF ≥ 2), which could have led to reporting false negatives or when viral nucleic acid was detected, the measured concentrations could have been under-quantified. Comparisons of inhibition levels by source type cannot be made from these data because of differences in inter- and intra-laboratory methods applied to the various water types.

HGV Armored RNA® was selected as an internal standard for control of inhibition in environmental samples for several reasons. First, HGV should not be present in any environmental water samples as it is an enveloped, blood-borne, single-stranded RNA virus of the Flaviviridae family causing liver inflammation in immunocompromised populations (Lindenbach et al., 2007). Second, by using HGV instead of a positive control comprised of the target viruses, we eliminate the risk of sample contamination (HGV is not used as a surrogate or control for any other work in our laboratories thus minimizing the likelihood of sample contamination with HGV). Third, by selecting a non-competitive internal control, we minimize potential competition between control and target amplification, which may decrease assay sensitivity (Stevenson et al., 2008). Last, HGV armored RNA is commercially available, allowing for this inhibition method to be easily transferable to other research laboratories. A potential shortcoming of using HGV is that PCR inhibition of HGV RNA may not be identical to the level that would have been measured for virus-specific PCR reactions. However, we believe this is a negligible concern when compared to the consequences of undetected laboratory contamination with target virus amplicon, leading to the reporting of false positives. An additional consideration is cost; depending on sample volume the inhibition measure adds approximately US$ 3.00 to 5.00 per sample to the analysis, not accounting for labor.

We used dilution of the nucleic acid extraction eluate with nuclease-free water as the primary method for alleviating the effects of inhibitors. By using qPCR to measure the degree of inhibition and using the difference in measured and expected HGV Cq values to calculate an appropriate dilution factor, the trade-off between diluting the inhibitors and diluting the amplification target to the point of signal loss is minimized. This dilution trade-off is often cited as the reason for using

---

**Table 1 — Number of water samples by source type and study site or location.**

<table>
<thead>
<tr>
<th>Source water type</th>
<th>Groundwater</th>
<th>Surface water</th>
<th>Sewage</th>
<th>Agricultural run-off</th>
<th>Finished drinking water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ghana n = 4</td>
<td>California n = 3</td>
<td></td>
<td></td>
<td></td>
<td>California n = 3</td>
</tr>
<tr>
<td>Madison n = 162</td>
<td>Ghana n = 10</td>
<td></td>
<td></td>
<td></td>
<td>Ghana n = 5</td>
</tr>
<tr>
<td>WAHTER® n = 2051</td>
<td>Indiana n = 5</td>
<td></td>
<td></td>
<td></td>
<td>Maryland n = 8</td>
</tr>
<tr>
<td>Washington n = 10</td>
<td>Iowa n = 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Madison n = 48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maryland n = 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Michigan n = 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Milwaukee n = 404</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ohio n = 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pennsylvania n = 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Virginia n = 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Washington n = 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>n = 2227</td>
<td>n = 497</td>
<td>n = 76</td>
<td>n = 377</td>
<td>n = 16</td>
</tr>
</tbody>
</table>

a Wisconsin Water And Health Trial for Enteric Risks (WAHTER Study).
b United States Department of Agriculture.
other methods (Funes-Huacca et al., 2011), a justifiable concern if the only option is standard 10-fold dilutions. Using our described method, however, it is possible to determine the minimum dilution necessary to reduce inhibition.

We usually measure the inhibition level in a set of samples, make the necessary dilutions to those samples that are inhibited, and run the qPCR assays for the targets (e.g., human enteric viruses). To reduce costs and time we do not measure inhibition again after the dilution step. In a large study of hundreds to thousands of samples it becomes impractical to measure inhibition more than once. Laboratory B did check the effectiveness of the dilution on a subset of inhibited surface water samples. Dilution was very effective in reducing inhibition, bringing the $C_q$ of the seeded HGV within ± one cycle of the $C_q$ of the HGV positive control (Table 2). It is important to recognize we do not always make the exact dilution called for by the dilution factor formula. For example, if the formula specifies 1:72 we may make a 1:50 dilution. The actual dilution is made in reference to the calculated dilution as well as in the context of our experience working with the sample matrix and the expected concentration of target in the sample. For example, for wastewater influent with high target concentrations we will make the dilution calculated, but for treated drinking water with low target concentrations we might dilute less than calculated to avoid a false negative.

Several published studies and reviews have cited humic substances (humic acids and fulvic acids) as the primary inhibitors associated with application of PCR to environmental samples such as soil, water, sediments, and sewage (Radstrom et al., 2008; Tsai and Olson, 1992; Watson and Blackwell, 2000; Wilson, 1997). Identifying the type and quantity of inhibitory compounds present in each sample could allow a more targeted mitigation strategy to be selected. However, whether this approach is cost-effective and results in improved qPCR measurements has yet to be determined.

Fig. 1 – Frequency distribution of sample inhibition by water source type. Inhibition is defined as a dilution factor ≥2. Dilution factor ≥100 includes all samples with inhibition levels requiring a dilution of 1:100 or greater. The histogram intervals for the main plots and insets are 1.99. Insets magnify the y-axis scale in the 0–3% range to better visualize the frequency distribution of inhibition for those samples with dilution factors between approximately 4 and 99.
3.2. **Inhibition levels did not vary with sample volume**

Filtering or concentrating large sample volumes is often suspected to cause an increase in inhibition but our data do not show this association (Fig. 2). We observed both small (<10 L) and large (>1,000 L) volumes of water to have low levels of inhibition as well as elevated levels of inhibition. The range in groundwater sample volumes spanned more than four orders of magnitude, and yet inhibition did not exhibit any clear pattern (Fig. 2B), suggesting other factors such as site location or season have a greater effect than sample volume on inhibition level. In Fig. 2, the data are stratified by water source type (groundwater, surface water, and agricultural runoff) to remove the effect of this variable on inhibition level. Sample collection method also could be a confounding variable but we believe this effect on the analysis is negligible because the far majority of samples were collected by glass wool filtration. Ultrafiltration was used to collect only zero water samples and sewage samples are not reported in Fig. 2. Drinking water samples and sewage samples are not reported in Fig. 2 because these were collected at two fixed volumes, 100 L and 4 L, respectively.

To our knowledge, Hata et al. (2011) is the only published study to investigate water sample volume as related to qPCR inhibition. In contrast to our findings, Hata et al. (2011) reported large sample volumes (8–200 L) to be more qPCR inhibited than small sample volumes (0.05–2 L). However, in that report, only 24 samples were analyzed, and more importantly, the methods of primary and secondary concentration were specific to the sample volume collected; that is, one method was applied to small volumes while a different method was applied to large volume samples. This correlation of concentration method with sample volume makes it difficult to distinguish whether levels of inhibition were attributable solely to sample volume or if the concentration methods were also contributing factors. Our analysis is also limited in that the sample volume comparisons are not side-by-side. Ideally, inhibition levels should be compared between small and large sample volumes collected at the same site at the same time with the same method.

3.3. **Every sample must be analyzed for inhibition**

Among the sample locations pooled for this research, several sites were sampled repeatedly over one to two year periods, presenting an opportunity to examine how inhibition levels may vary in samples collected at the same site over time. For example, a groundwater well was sampled twenty-six times over a two year period with inhibition levels remaining mostly steady with only a single spike in inhibition (Fig. 3). In contrast, the inhibition levels in samples collected at two surface water locations varied considerably over time (Fig. 3). As it is not possible to assess the inhibition potential of a sampling site from a single sample, we recommend every sample be analyzed for inhibition.

3.4. **Number of false negatives if inhibited sample is not diluted**

The ability to identify inhibition and avoid reporting false negatives is important for protecting public health. Fig. 4 demonstrates the potential number of false negatives that could have been reported had inhibition not been measured. These are samples that were positive for a human enteric virus (either adenovirus, enterovirus, norovirus, hepatitis A virus, or rotavirus) after making the dilution according to the calculated DF, and the C<sub>q</sub> value for the virus detected was >37. We selected this C<sub>q</sub> value as it is near the limit of detection for most qPCR assays and low level inhibition before dilution could have moved the C<sub>q</sub> below the limit of detection and resulted in the sample being virus-negative. For the purpose of illustration, we use two definitions of potential false negatives: 1) The sample was virus-positive with a C<sub>q</sub> ≥ 37 and the measured inhibition level required a dilution factor of 2 or greater; and 2) The sample was virus-positive with a C<sub>q</sub> ≥ 37 and the measured inhibition level required a dilution factor of 10 or greater. In other words, if the sample had not been diluted at least 1:2 or 1:10 the inhibitors present would have most likely resulted in the target virus not being detected. This potential to be a false positive is based on the one and 3.3 cycle shifts that translate from calculated dilution factors of 2 and 10, respectfully. A sample with a C<sub>q</sub> ≥ 37 would have reported a C<sub>q</sub> ≥ 38 and C<sub>q</sub> ≥ 40, both potentially "non detects", if the sample had not been diluted 1:2 and 1:10, respectively. Among the various water types, the fraction of samples with the potential to be false negatives ranged from 0.3% to 71% (Fig. 4).

The opposing concern, diluting out the sample too much that it brings the viral target below the limit of detection, is also illustrated by Fig. 4. Because the two definitions of a potential false negative are not mutually exclusive, the difference in the samples numbers between the two definitions (i.e., gray bar minus black bar in Fig. 4) indicates the number of samples that could have been virus-negative had the nucleic acid extraction eluate been automatically diluted.

---

**Table 2 – Effect of dilution of the sample nucleic acid extract on reducing qPCR inhibition.**

<table>
<thead>
<tr>
<th>Inhibition group&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Number of samples</th>
<th>Mean (Expected HGV C&lt;sub&gt;q&lt;/sub&gt; – Measured HGV C&lt;sub&gt;q&lt;/sub&gt;)&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Before dilution</th>
<th>After dilution&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete inhibition</td>
<td>12</td>
<td>&gt;12</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>Partial inhibition</td>
<td>10</td>
<td>5.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Complete inhibition refers to those samples inhibited to the extent that the seeded HGV could not be detected (i.e., C<sub>q</sub> > 40). Partial inhibition refers to those samples in which seeded HGV could be detected and the inhibition level quantified.

<sup>b</sup> The target C<sub>q</sub> of the HGV positive control (i.e., expected) was approximately 28.

<sup>c</sup> Completely inhibited samples were diluted no more than 1:100.
1:10 without measuring inhibition. An automatic 1:10 dilution could have diluted the virus target beyond the limit of detection in 40% (44 of 109), 93% (91 of 98), 37% (20 of 54), and 26% (9 of 34) of the inhibited surface water, groundwater, sewage, and agricultural runoff samples, respectively. These numbers demonstrate the value of using the dilution factor formula instead of guessing for making the correct inhibitor dilution.

Similar to other methods for identifying PCR inhibition, there are potential limitations associated with the method we describe here. First, PCR inhibition of HGV RNA may not be identical to the inhibition level measured using virus-specific PCR reactions. As mentioned previously, we choose not to use target viruses as inhibition controls to minimize the chance of false positives resulting from contaminating samples with target amplicon. Second, because this method requires
a separate PCR reaction it could be more time consuming and less cost effective than multiplex methods using internal amplification controls. Third, the method is not a panacea; false negatives are still possible when inhibition is very high. For example, if the calculated dilution factor is 1,000, making this dilution to counter inhibition would almost certainly dilute any target viral nucleic acid present in the sample to below the limit of detection. When inhibition is this high, Laboratory B passes the nucleic acid extract through a Sephadex column to remove inhibitors and bring the required dilution to ≤1:100.

It is important to emphasize the risk of obtaining false negatives is not the only reason to control and correct for sample inhibition. Researchers reporting these data also need to be concerned about under-quantifying pathogen concentration as inhibitors can cause extremely inaccurate quantitative results and thus possibly underestimating health risk (Guescini et al., 2008). This is especially important as quantitative microbial risk assessment (QMRA) tools are used by researchers and regulators to examine and calculate the risks individual pathogenic microorganisms pose through contaminated water sources including drinking water and recreational water (Rose and Grimes, 2001; Schoen et al., 2011). Similar to the conventional, or chemical, risk assessment framework, QMRA includes hazard identification, dose–response analysis, exposure assessment, and risk characterization. Traditionally, exposure assessments utilized for microorganisms in ambient water were based on index microorganisms for each group of pathogens since there are a large number of microorganisms associated with waterborne disease (Boehm et al., 2009; Reynolds et al., 2008). Information gathered from the index organism was then used to extrapolate risk for a specific pathogen. However, with the introduction of new detection techniques, including qPCR, and water concentration methods, researchers are increasingly able to assess exposure to the actual pathogen of interest. While new molecular-based technologies are important, researchers should report quantitative data with caution and adhere to MIQE guidelines (Bustin et al., 2009) as well as include appropriate controls and mitigation steps for sample inhibition.

4. Conclusions

Important conclusions from these data are:

- Inhibition of qPCR and RT-qPCR assays is a significant concern during the analysis of viral nucleic acid extracts from environmental water samples; thus, the identification and quantification of inhibition in each sample is critical for proper quality assurance and control during analysis and reporting.
- The dilution factor formula allows a more accurate calculation of the dilution needed to mitigate inhibition, reducing the likelihood of over-diluting the sample to the point where the target signal is below the limit of detection.
- Increasing the sample volume filtered or concentrated is often suspected to cause an increase in inhibition; however, the data presented here do not support this assumption.
- Because it is important to analyze every sample for inhibition, even if a sample set are all collected from the same location, there is a need for a simple and cost-effective method such as the one we describe, for measuring inhibition.
- The issue of not measuring and correcting for sample inhibition is greater than just the potential for false negatives; inhibition can also result in under-quantifying pathogen concentration and thus possibly underestimating infection risk.
**Acknowledgments**

We thank our collaborators for permitting us to report the inhibition levels we measured in samples from their various studies: Ken Bradbury and Madeline Gotkowitz at the Wisconsin Natural History and Geological Survey for Madison groundwater and lake samples; Steve Corsi at the US Geological Survey in Middleton, WI for river samples from the Milwaukee area; Donna Francy at US Geological Survey in Columbus, OH for sewage influent samples from plants in Ohio; Bill Jokela at the USDA-Agricultural Research Service in Marshfield, WI for agricultural runoff samples; and Sandra McLeLLan at the Water Institute, University of Wisconsin—Milwaukee for samples from the Milwaukee harbor and Lake Michigan. We also thank James Schissler for his laboratory assistance and support. Dr. Tim Julian, Johns Hopkins Bloomberg School of Public Health, reviewed an earlier manuscript version and we thank him for his helpful comments. Analyses for water samples obtained for the Wisconsin Water And Health Trial for Enteric Risks (WAHTER Study) were funded by US EPA STAR Grant R831630. Studies using the hollow fiber, ultrafiltration method were partially supported by an US EPA STAR Grant (R833002) the Johns Hopkins Global Water Program, the Osprey Foundation of Maryland as well as the Johns Hopkins Center for a Livable Future Pre-doctoral Fellowship – K.E. Gibson was a Predoctoral Fellow. The views expressed herein have not been subjected to US EPA review and therefore do not necessarily reflect the views of the agency, and no official endorsement should be inferred.

**References**


