

# Whole genome amplification (WGA) for archiving and genotyping of clinical isolates of *Cryptosporidium* species

MAHA BOUZID<sup>1</sup>, DARREN HEAVENS<sup>2</sup>, KRISTIN ELWIN<sup>3</sup>, RACHEL M. CHALMERS<sup>3</sup>,  
STEPHEN J. HADFIELD<sup>3</sup>, PAUL R. HUNTER<sup>1</sup> and KEVIN M. TYLER<sup>1\*</sup>

<sup>1</sup> Biomedical Research Centre, School of Medicine, Health Policy and Practice, University of East Anglia, Norwich NR4 7TJ, UK

<sup>2</sup> Genome Lab, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK

<sup>3</sup> UK *Cryptosporidium* Reference Unit, NPHS Microbiology Swansea, Singleton Hospital, Swansea SA2 8QA, UK

(Received 3 April 2009; revised 10 June 2009; accepted 30 June 2009; first published online 21 September 2009)

## SUMMARY

Clinical and environmental isolates of pathogens are often unique and may be unculturable, yielding a very limited amount of DNA for genetic studies. *Cryptosporidium* in particular are difficult to propagate. Whole genome amplification (WGA) is a valuable technique for amplifying genomic material. In this study, we tested 5 WGA commercial kits using *Cryptosporidium* clinical isolates. DNA of 5 *C. hominis* and 5 *C. parvum* clinical isolates and *C. parvum* IOWA reference strain were used. The majority of the samples were amplified by all of the kits tested. The integrity and fidelity of the amplified genomic DNA were assessed by sequence analysis of several PCR products of varying length. We found evidence that one kit in particular may be more error prone while another seemed the more suitable kit for *Cryptosporidium* clinical samples, generating high molecular weight DNA from all the samples with high fidelity. Thus WGA was found to be a useful technique for producing amplified DNA suitable for downstream genotyping techniques and archiving of *Cryptosporidium* clinical isolates.

Key words: whole genome amplification, multi-displacement amplification, *Cryptosporidium* DNA, genotyping.

## INTRODUCTION

The availability of adequate amounts of high quality genomic DNA is essential for genetic studies such as diagnosis and genotyping. Genotyping and subtyping of clinical and environmental isolates is desirable as it allows source tracking and better understanding of molecular epidemiology and population structure (Anderson *et al.* 2000; Han *et al.* 2000; Mallon *et al.* 2003; Burgos *et al.* 2004; Smith *et al.* 2007). Unfortunately, such isolates are often unculturable, yielding a very limited amount of DNA for study. *Cryptosporidium*, an emergent protozoan parasite causing mainly diarrhoeal illness in humans and animals, is particularly difficult to propagate. Although *Cryptosporidium* oocysts are excreted in high numbers in the faeces during acute clinical episodes (Goodgame *et al.* 1993), the purification methods used to obtain clean DNA, suitable for downstream molecular methods, usually result in losses. This presents difficulties for the study of isolates present in lower numbers, such as those in environmental

samples or during subclinical infection. These issues have limited biological studies to *C. parvum* calf propagated strains, particularly the IOWA reference strain (Cama *et al.* 2006).

Several subtyping techniques have been applied to *Cryptosporidium* species using different markers: glycoprotein GP60 (Strong *et al.* 2000; Leav *et al.* 2002), double-stranded RNA element (Leoni *et al.* 2003) and mini- and microsatellite repeats (Cacciò *et al.* 2000; Mallon *et al.* 2003). Ideally, each new isolate should be tested using a panel of markers. However, this is usually limited by the amount of DNA available. Whole genome amplification (WGA) can be used to increase the amount of nucleic acid available from clinical and environmental samples of waterborne pathogens (reviewed by Bouzid *et al.* 2008) and application of this technique to *Cryptosporidium* isolates should address the perceived need for multilocus typing (Smith *et al.* 2006).

The first described WGA methods were degenerate oligonucleotide-primed PCR (DOP-PCR) (Telenius *et al.* 1992; Cheung and Nelson, 1996) and primer extension preamplification (PEP) (Zhang *et al.* 1992). However, these PCR-based techniques produced short products (<3 kb) and were limited by substantial amplification bias and incomplete coverage of genetic markers (Paunio *et al.* 1996; Dean *et al.* 2002; Hawkins *et al.* 2002; Park *et al.* 2005). New

\* Corresponding author: Biomedical Research Centre, School of Medicine, Health Policy and Practice, University of East Anglia, Norwich NR4 7TJ, UK. Tel: +44 (0) 1603 591225. Fax: +44 (0) 1603 591750. E-mail: k.tyler@uea.ac.uk

Table 1. Origin and epidemiological data of clinical isolates of *Cryptosporidium hominis* and *C. parvum* used for this study

(The concentration of each DNA suspension was evaluated by nanodrop.)

Sample reference	Age of case	Gender of case	Origin of case	RFLP of the COWP gene (speciation)	DNA concentration (nanodrop ng/ $\mu$ l)
W15504	43	F	Scotland	<i>C. hominis</i>	12
W15507	9	F	Wales	<i>C. hominis</i>	4.3
W15508	31	Not stated	England	<i>C. hominis</i>	3.5
W15519	68	F	England	<i>C. hominis</i>	5.5
W15521	14	F	Scotland	<i>C. hominis</i>	11.5
W15509	12	Not stated	England	<i>C. parvum</i>	6.7
W15511	19	F	Scotland	<i>C. parvum</i>	6.2
W15516	21	F	England	<i>C. parvum</i>	8.9
W15517	12	M	England	<i>C. parvum</i>	6.4
W15518	2	M	England	<i>C. parvum</i>	13.2

strategies for WGA have been developed including multiple displacement amplification (MDA) and OmniPlex WGA (Park *et al.* 2005). MDA is an isothermal amplification using degenerate hexamers and the bacteriophage phi-29 DNA polymerase, which possesses high processivity, strand-displacement abilities and a proof-reading activity resulting in error rates 100 times lower than the Taq polymerase (Eckert and Kunkel, 1991; Esteban *et al.* 1993; Dean *et al.* 2002; Nelson *et al.* 2002; Hawkins *et al.* 2002). MDA was first described by Blanco *et al.* (1989) and then used for WGA of different targets such as lymphoma and leukaemia clinical specimens (Luthra and Medeiros, 2004), complex mixtures of DNA (Shoaib *et al.* 2008), whole blood and tissue-culture cells (Dean *et al.* 2002), human blastomeres (Snabes *et al.* 1994), plasmid constructs and whole bacterial genomes (Detter *et al.* 2002). The OmniPlex WGA technique uses libraries of 200–2000 bp fragments created by random chemical cleavage of genomic DNA, followed by ligation of adaptor sequences to both ends and PCR amplification (Barker *et al.* 2004; Bergen *et al.* 2005). This fragmentation/ligation/PCR-based method amplifies the entire genome several 1000-fold, and could be even re-amplified to achieve a final amplification of over 1 000 000-fold without introducing inaccuracies (Langmore, 2002). Currently, several commercial kits for MDA and OmniPlex based WGA are available.

We report here the results of the first study to evaluate the suitability of WGA for the accurate expansion of genomic DNA from *Cryptosporidium* isolates using commercial kits.

## MATERIALS AND METHODS

### *Cryptosporidium* DNA

Ten clinical samples were selected from the collection of the *Cryptosporidium* reference unit (CRU),

Swansea. They originated from diarrhoea patients with confirmed cryptosporidiosis from different geographical locations in the UK. DNA was isolated from semi-purified oocyst suspensions prepared from stool samples by saturated-salt solution centrifugation and extracted using QIAamp DNA mini kit spin columns (Qiagen Ltd, Crawley, UK) as previously described by Elwin *et al.* (2001). For each sample, the speciation was performed by PCR-RFLP of the *Cryptosporidium* oocyst wall protein (COWP) gene as previously described by Spano *et al.* (1997) and by real-time PCR using simplex Lib 13 primers for *C. parvum* and *C. hominis* as described by Tanriverdi *et al.* (2003). For this study, we tested the DNA of 5 *C. hominis* and 5 *C. parvum* isolates as these two species are the cause of the majority of human cases of cryptosporidiosis. Table 1 detailed the origin and the epidemiological data of the tested isolates. All DNA samples were quantified by spectrometry using Nanodrop ND-1000 (Thermo Fisher Scientific, Leicestershire, UK). In addition, *C. parvum* IOWA reference strain DNA extracted from a purified oocyst suspension from a commercial source (ATCC/LGC Promochem, Teddington, UK) was tested as a positive control and its DNA concentration was quantified as 5.8 ng/ $\mu$ l.

### WGA kits

Three commercial WGA kits were tested during this study: illustra GenomiPhi V2 DNA amplification Kit (GE Healthcare, Buckinghamshire, UK), REPLI-g Ultra-fast Mini Kit (Qiagen, Crawley, UK) and GenomePlex<sup>®</sup> Complete WGA Kit (Sigma, Dorset, UK). Illustra GenomiPhi and REPLI-g kits are MDA-based WGA, while GenomePlex kit is based on the OmniPlex technique. All 3 kits were used according to the manufacturer's instructions. The manufacturer's recommended starting genomic DNA concentration was 1 ng for the illustra kit and 10 ng for both the REPLI-g and GenomePlex kits.

As the DNA concentrations of *Cryptosporidium* isolates ranged from 3.5 to 13.2 ng/ $\mu$ l (with a mean of 7.7 ng/ $\mu$ l), 1  $\mu$ l of each sample was used for WGA. WGA techniques are well established in the literature to give amplification levels from 10-fold less than our lowest template concentration (0.3 ng) to 20-fold higher than our highest concentration (300 ng) (Dean *et al.* 2002), which raises the prospect for them to be used widely for clinical samples. Thus, in our study all of the samples lie well within the range we would expect to give good amplification by the kits under test and the range recommended by the manufacturers.

#### Quantification of genomic DNA after WGA

After WGA, the amplified products were analysed by agarose gel electrophoresis. In addition, 3 methods were used to quantify the amplified genomic DNA: Nanodrop, Hoechst and PicoGreen. Hoechst 33258 dye exhibits enhanced fluorescence when bound to dsDNA under high ionic strength conditions (Goumenou and Macheria, 2004). Serial dilutions of calf thymus DNA stock solution (1 mg/ml) (Sigma, Dorset, UK) were performed yielding concentrations ranging from 100 ng/ml to 2500 ng/ml. Fifty  $\mu$ l of each preparation was used for measurement. All DNA samples were diluted in TNE buffer (100 mM Tris, 1 M NaCl, 10 mM Na<sub>2</sub>EDTA, pH 7.5). For each DNA sample, an equal volume of 2x Hoechst Dye solution (200 ng/ml) (Sigma, Dorset, UK) was added. Fluorescence was read using a microplate reader FLUOstar, BMG Labtech (Aylesbury, UK) after incubation for 5 min at room temperature. The blank solution was prepared by adding an equal volume of TNE buffer and 2x dye solution. The average value of the blank measurement was subtracted from the measurements made at each concentration and the results plotted. A linear regression was performed on the standard curve to allow the determination of the DNA concentrations of the tested samples. PicoGreen was also used as an ultra-sensitive fluorescent nucleic acid stain for accurate quantification of double-stranded DNA (dsDNA). Quanti-iT™ PicoGreen® dsDNA kit (Molecular Probes, Invitrogen, Paisley, UK) was used according to the manufacturer's instructions. We also used GeneTools software (Syngene, Cambridge, UK) for densitometry-based DNA quantification.

#### Integrity and fidelity of amplified genomic DNA

The integrity of the amplified DNA for the 3 kits was evaluated by PCR using Cry15/Cry 9 primers amplifying 550 bp of the COWP gene (Spano *et al.* 1997). In addition, 2 newly designed primers were also used, amplifying 270 bp and 247 bp of Cgd6\_5020 and Chro.20156 genes, respectively. COWP

and Cgd6\_5020 markers are on Chromosome 6 and Chro.20156 gene is on Chromosome 2. The primer sequences were as follows: Cgd6\_5020F (AACAG GAGCTGACGATTGCT), Cgd6\_5020R (ACAT TGTGCCATTCCAAGGT), Chro.20156F (TTC GCTTGAAGCCGTAACT) and Chro20156R (GGCATTGATACCAGGCAAGT). All DNA templates were diluted 1/25 after WGA, PCR conditions were the same for genomic DNA and for post-WGA subsequent amplifications. The PCR mix for each primer set consisted of 1.5 mM MgCl<sub>2</sub>, 0.2 mM of each dNTP (Bioline, UK), 0.6  $\mu$ M of each primer and 2.5 U of HotStar Taq DNA polymerase (Qiagen Ltd, Crawley, UK) in a 50  $\mu$ l final volume. The cycling conditions were as follows: an initial hot-start at 95 °C for 15 min, followed by 40 cycles of 94 °C for 50 s, 57 °C for 30 s and 72 °C for 50 s and a final extension at 72 °C for 10 min. The PCR reactions were performed using Techne TC-512 thermal cycler (SLS, Nottingham, UK). PCR products were run on a 2% agarose gel stained with ethidium bromide in TBE buffer and visualized under UV.

The fidelity of the amplification was assessed by PCR product sequence analysis before and after WGA of 2 *C. hominis* samples (W15507, W15519), 2 *C. parvum* samples (W15511, W15516) and the reference strain *C. parvum* IOWA. PCR products were purified using QIAquick® PCR purification Kit (Qiagen Ltd, Crawley, UK) and were sequenced using the Big-Dye Terminator cycle sequencing system and an ABI 3770 DNA sequencer at the genome lab, John Innes Centre ([www.jic.ac.uk](http://www.jic.ac.uk)). The sequences were aligned using Vector NTI Advance™ 10 (Invitrogen, Paisley, UK).

#### Comparative analysis of *Cryptosporidium* genomic DNA before and after WGA

The comparative analysis of genomic DNA and paired WGA DNA was limited to the samples which were prepared using the most promising WGA method as explained below. After WGA, the amplified products were compared to the original genomic DNA using a species-specific, semi-quantitative real-time PCR assay. Briefly, real-time PCR using simplex Lib 13 primers for *C. parvum* and *C. hominis* (Tanriverdi *et al.* 2003) was used in triplicate to amplify 2  $\mu$ l of genomic DNA and 2  $\mu$ l of its paired WGA DNA (both diluted 1 in 25 v/v). We used Corbett Rotorgene 3000 platform (Corbett Life Science, Sydney, Australia). Melt curve analysis was performed to identify *Cryptosporidium* species and C<sub>T</sub> (threshold cycle) values were recorded to compare each paired sample, before and after WGA. This assay shows a demonstrable difference of 3 C<sub>T</sub> units per 10-fold difference in DNA target copy number (CRU unpublished data). Since the WGA DNA originated from 1  $\mu$ l of genomic DNA producing 10  $\mu$ l of product (representing a 10-fold dilution in comparison

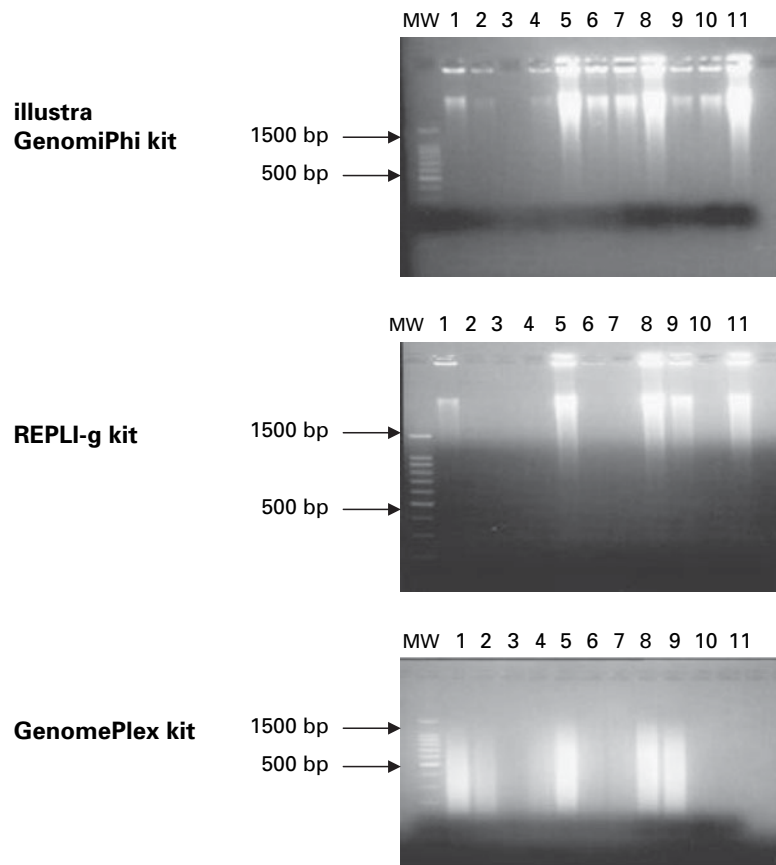


Fig. 1. Agarose gel electrophoresis of 8  $\mu$ l of amplified genomic DNA. MW: molecular weight marker, 1: W15504, 2: W15507, 3: W15508, 4: W15519, 5: W15521 (1–5: *Cryptosporidium hominis* samples), 6: W15509, 7: W15511, 8: W15516, 9: W15517, 10: W15518 (6–10: *C. parvum*), 11: *C. parvum* IOWA.

with the genomic DNA), the final  $C_T$  values for the WGA DNA were decreased by 3  $C_T$  units.

## RESULTS

### Success rate and yield of WGA kits

The reliability and robustness of WGA kits to variation in DNA acquired from clinical samples were assessed by agarose gel electrophoresis. A successful amplification was considered when genomic DNA was visible on the gel. One sample W15508 did not amplify with any of the 3 kits tested.

Illustra GenomiPhi and REPLI-g kit-amplified DNA was of high molecular weight ( $\sim 10$  kb) for all the tested samples. The success rate was 90.91% (10/11 samples) for illustra GenomiPhi kit and 45.46% (5/11) for REPLI-g kit (Fig. 1). For GenomePlex amplified samples, the generated DNA was of smaller size ranging between 200 and 1000 bp (mean size  $\sim 400$  bp) (Fig. 1). The success rate was 72.73% (8/11 samples). For illustra GenomiPhi and REPLI-g kits, samples W15521, W15516 and IOWA showed the strongest bands, suggesting better amplification. However, when the GenomePlex kit was used, W15504, W15521, W15516 and W15517 showed a high level of amplification, but, surprisingly commercially obtained IOWA DNA did not.

DNA quantification after WGA was initially assessed by 4 methods: nanodrop spectrophotometry alone, Hoechst and PicoGreen fluorimetry and agarose gel-based ethidium bromide fluorescence densitometry. Presumably because of the presence of residual random hexamers in the reaction mix (Ahn *et al.* 1996; Singer *et al.* 1997) the first 2 methods required additional purification before DNA quantification, PicoGreen quantification gave good quantitative data for 2 of the 3 kits tested (illustra GenomiPhi and GenomePlex kits) but only gel-based densitometry was able to provide DNA concentrations for all 3 kits without an additional purification step.

The results of WGA for the samples tested are summarized in Table 2. For each kit, yield was determined by calculating the amount of DNA in the final reaction volume (10  $\mu$ l for illustra GenomiPhi and 75  $\mu$ l for GenomePlex kit). Level of amplification was determined as a ratio of concentrations between template and WGA DNA for each sample. The typical yield of illustra GenomiPhi kit was 0.7–7  $\mu$ g range. The highest yield was 10  $\mu$ g achieved from IOWA DNA giving over 180-fold amplification. The typical yield of GenomePlex kit gave a 4.5–46  $\mu$ g range and the level of amplification was 10 to 70-fold. For the REPLI-g kit the yield was of a 0.6–2  $\mu$ g range

Table 2. Comparison of the performances of the 3 trialed WGA commercial kits

(Success rate was the number of samples with positive amplification as determined by agarose gel electrophoresis. The size of amplified product was determined by comparison to a molecular weight marker. PicoGreen quantification column showed the successful quantification without prior sample purification. Yield and level of amplification were determined based on PicoGreen results for illustra GenomiPhi and GenomePlex kits. \* for REPLI-g kit, densitometry was used to determine the yield and the level of amplification. Secondary amplification was the success rate of post-WGA PCR reactions. Fidelity was determined by the number of identical nucleotides before and after WGA.)

	Success rate (electrophoresis)	Size of amplified product	PicoGreen quantification	Yield	Level of amplification	Secondary amplification (Post-WGA PCRs)	Fidelity (nucleotide sequence match)	Specific amplification of <i>Cryptosporidium</i> DNA (Real-Time PCR)
Illustra GenomiPhi	90.91 %	> 10 kb	Yes	0.7–10 µg	6–180 fold	87.88 %	100 %	Yes (11/11 samples)
REPLI-g	45.46 %	> 10 kb	No	*0.6–2 µg	*50–160 fold	63.64 %	99.62 %	Not tested
GenomePlex	72.73 %	< 1 kb	Yes	4.5–46 µg	10–70 fold	93.94 %	99.87 %	Not tested

(final volume 20 µl), corresponding to 50 to 160-fold amplification.

#### *Integrity of amplified DNA*

The integrity of the amplified genomic DNA was assessed by PCR using 3 primer sets. For Cry 15/9 primers, amplifying a 550 bp of the COWP gene, all 11 samples were positive before WGA (data not shown). After WGA, 10/11 illustra GenomiPhi amplified samples (91%), 6/11 REPLI-g amplified samples (54.6%) and 9/11 GenomePlex amplified samples (81.9%) were PCR positive (Fig. 2). For Cgd6\_5020 primers, amplifying 270 bp, all 11 samples were also positive before WGA. After WGA, 10/11 illustra GenomiPhi amplified samples (91%), 8/11 REPLI-g amplified samples (72.8%) and 11/11 GenomePlex amplified samples (100%) were PCR positive (data not shown). For Chro.20156 primers, amplifying 247 bp, 10/11 samples were positive before WGA, only W15519 sample was negative (data not shown). After WGA, 9/11 illustra GenomiPhi amplified samples (81.9%), 7/11 REPLI-g amplified samples (63.7%) and 11/11 GenomePlex amplified samples (100%) were PCR positive. The overall post-WGA PCR success rates from the 3 kits were 87.88%, 63.64% and 93.94% for illustra GenomiPhi, REPLI-g and GenomePlex kits, respectively (Table 2).

#### *Fidelity of WGA kits*

The fidelity of amplification was assessed by PCR product sequence analysis, generated with and without a WGA intermediate step. For Cry 15/9 primers, PCR product sequences were identical using all 3 kits for W15507, W15511 and IOWA isolates (supplemental data). For W15516 isolate, WGA using REPLI-g kit produced 3 nucleotide errors and the use of GenomePlex kit produced 1 error, corresponding to 99.34% and 99.78% sequence concordance, respectively. For Cgd6\_5020 and Chro.20156 primers, PCR product sequences were identical before and after WGA for all the samples. The overall fidelity of the 3 kits was 100%, 99.62% and 99.87% for illustra GenomiPhi, REPLI-g and GenomePlex kits, respectively (Table 2).

Interestingly, one of the samples tested (W15519) failed to amplify using Chro.20156 primers without WGA or after WGA using illustra GenomiPhi and REPLI-g kits, but did amplify when GenomePlex amplified DNA was used. The identity of the PCR product was confirmed by sequencing.

#### *Comparative analysis of *Cryptosporidium* genomic DNA before and after WGA*

Melt curve analysis of genomic DNA and WGA DNA amplified using the Illustra GenomiPhi confirmed that the same species of *Cryptosporidium*

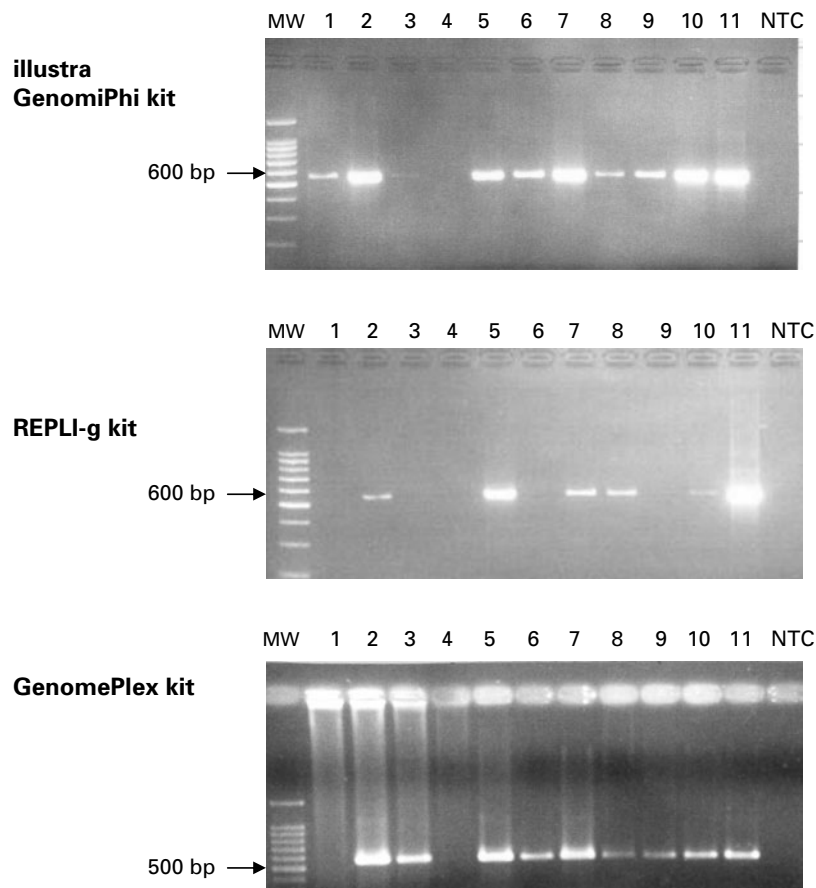


Fig. 2. PCR products of WGA amplified *Cryptosporidium* DNA isolates using Cry15/9 primers. MW: molecular weight marker, 1: W15504, 2: W15507, 3: W15508, 4: W15519, 5: W15521, 6: W15509, 7: W15511, 8: W15516, 9: W15517, 10: W15518, 11: *C. parvum* IOWA, NTC: non-template control.

was present before and after WGA in each of the samples tested (data not shown). The unadjusted  $C_T$  values show that 8/11 of the samples had a lower  $C_T$  after WGA than before, indicating that in these samples the WGA did amplify *Cryptosporidium* DNA. After adjustment to allow for the 10-fold dilution applied through the WGA process, all of the samples had a lower  $C_T$  value after WGA confirming that *Cryptosporidium* DNA was present in higher copy numbers in the samples post-WGA than before (Table 3). The highest difference of  $C_T$  before and after WGA was 10.13 for sample W15516, which corresponds to over 30-fold increase in DNA target copy number. This is in accordance with the overall level of amplification. For the other samples, the difference in  $C_T$  values before and after WGA ranged between 2.63 (for sample W15504) and 8.11 (for sample W15511), corresponding to an 8 to 27-fold increase in copy numbers. For sample W15508, the difference in  $C_T$  value was 0.73 supporting a poor amplification.

#### DISCUSSION

In this study, we successfully used WGA commercial kits for amplification of *Cryptosporidium* genomic

DNA from clinical isolates. The Illustra GenomiPhi and GenomePlex kits successfully amplified the majority of the tested isolates (90.91% and 72.73%, respectively). The REPLI-g Kit, however, amplified less than half of the samples. One sample W15508 did not amplify with any of the 3 kits tested, while the reason for this cannot be known and the template concentration was well within the recommended range for the kits under test, it did correlate with the sample having the lowest concentration of template DNA tested and thus may reflect some degree of degradation of the DNA in that sample or relative enrichment of inhibitors carried through the DNA purification process.

For MDA-based kits (illustra GenomiPhi and REPLI-g), the amplified DNA was of high molecular weight. It was noticeable that most of the DNA remains in the well when run out on an agarose gel, this is likely due to the formation of very high molecular weight DNA forms, independent of the genome size, as previously reported by Detter *et al.* (2002). OmniPlex-amplified DNA was of smaller size, which is due to the method used, involving fragmentation of the genomic DNA followed by linker ligation to enable amplification (Thorstenson *et al.* 1998; Fiegler *et al.* 2007).

Table 3. Real-time analysis of *Cryptosporidium* DNA before and after WGA and estimation of the increase in copy numbers after WGA using the illustra GenomiPhi kit

Sample	Genomic DNA/ WGA DNA	Mean C <sub>T</sub> (adjusted)	C <sub>T</sub> value difference	Increase in copy numbers
15504	Genomic DNA WGA DNA	31·63 29·00	2·63	8·8
15507	Genomic DNA WGA DNA	22·62 18·09	4·53	15
15508	Genomic DNA WGA DNA	29·06 28·33	0·73	2·5
15519	Genomic DNA WGA DNA	32·56 30·02	2·54	8·5
15521	Genomic DNA WGA DNA	23·72 19·31	4·41	14·8
15509	Genomic DNA WGA DNA	28·82 24·73	4·09	13·7
15511	Genomic DNA WGA DNA	23·57 15·46	8·11	27
15516	Genomic DNA WGA DNA	29·57 19·44	10·13	33·8
15517	Genomic DNA WGA DNA	28·87 24·79	4·08	13·6
15518	Genomic DNA WGA DNA	25·99 18·18	7·81	26
IOWA	Genomic DNA WGA DNA	19·15 11·46	7·96	26·6

PicoGreen DNA quantification of the amplified DNA proved effective for Illustra GenomiPhi and GenomePlex kits. The yield range was 0·7–10 µg and 4·5–46 µg, respectively, which is in accordance with the manufacturer's claims. The apparent higher yield of GenomePlex kit corresponds with a higher reaction volume (10 µl versus 75 µl). In our hands, the PicoGreen quantification of REPLI-g kit amplified samples was not effective without prior purification of the amplified DNA.

The integrity of the amplified DNA was assessed by the ability to generate PCR product for 3 genetic loci and 87·88% (29/33), 63·64% (21/33) and 93·94% (31/33) of *Cryptosporidium* samples, amplified by illustra GenomiPhi, REPLI-g and GenomePlex kits, respectively, were PCR positive. For 1 sample, the W15519 isolate, no PCR amplification using Chro.20156 primers was observed before WGA or after WGA using illustra GenomiPhi and REPLI-g kits. However, GenomePlex amplified DNA of the same sample was PCR positive and the identity of the PCR product was confirmed by sequencing. These results suggest that the use of WGA-amplified DNA as PCR template can sometimes actually increase PCR sensitivity from clinical samples. Similar findings were reported for the detection of *Trypanosoma* species from blood samples (Pinchbeck *et al.* 2008). Further work to more thoroughly test the integrity of the amplified DNA and assessing amplification biases should focus on amplification of longer sequences and broader genomic coverage utilizing sequences from each of the 8 nuclear chromosomes,

from the telomeres, centromeres, ribosomal DNA, mitochondrial DNA, and repetitive regions.

The fidelity of the amplification was assessed by PCR products sequence analysis before and after WGA. In our hands, for this sample set, the overall error rate observed was 0% for illustra GenomiPhi kit, 0·38% for REPLI-g kit and 0·13% for GenomePlex kit. Interestingly, all the errors arose from the same clinical sample using the same set of primers. This could be explained by a variety of factors such as the presence of impurities affecting the enzyme proof-reading activity, the secondary structure of the DNA, or by a low concentration of the starting material, which can decrease the amplification fidelity as reported by Bergen *et al.* (2005). It is also important to bear in mind that although the sequences analysed were short ones, both the secondary amplification step and the sequencing step have the potential to introduce errors which would not be discriminated from errors arising during the WGA.

Sequence analysis of WGA amplified PCR products using Cry 15/9 primers showed the preservation of 6 species-specific Single Nucleotide Polymorphisms (SNP)s, one of which at position 66 is of particular interest as it corresponds to an *RsaI* restriction site used for *Cryptosporidium* genotyping as previously described (Spano *et al.* 1997). In addition, sequence analysis showed the preservation of 1 species-specific SNP for Cgd6\_5020 gene and 5 species-specific SNPs for Chro.20156.

Comparative analysis of *Cryptosporidium* genomic DNA and paired WGA DNA using real-time PCR

assay confirmed that *Cryptosporidium* DNA was specifically amplified using illustra GenomiPhi kit, resulting in higher copy numbers post-WGA than before for all the samples tested.

These results were obtained from semi-purified oocyst suspensions, themselves requiring reasonable numbers of oocysts in the original sample. They carry significant contamination of bacterial, fungal and even animal DNAs also amplified by the process. Real-time PCR for faecal marker DNAs such as genes from *E. coli* might be useful to determine whether there is any predisposition to amplifying contaminating DNAs rather than the cryptosporidium DNA in the samples. Additional investigation of the effect of other oocyst purification methods (immunomagnetic separation) on the performance of WGA should also be performed. Importantly, since many investigators use DNA extracted from raw stool without prior oocyst purification, independent validation of the suitability of this material for WGA and downstream analysis should also be undertaken.

For our collection of *Cryptosporidium* clinical isolates, illustra GenomiPhi WGA kit had the best performance, with 90.91% success rate, generating a high concentration of high molecular weight DNA with 100% fidelity. The additional cost of WGA is not prohibitive for clinical usage – using illustra GenomiPhi kit based on the 2009 recommended retail price in the UK the cost is roughly £3.60 per sample added to the PCR cost for routine detection of *Cryptosporidium* DNA. These are preliminary results, highlighting the usefulness of MDA based WGA for the accurate amplification of *Cryptosporidium* genomic DNA for the purposes of immortalization of clinical isolates and enabling extensive genetic testing.

This study investigated the suitability of *Cryptosporidium* DNA after WGA for genotyping purposes. We tested COWP marker as it is routinely used for *Cryptosporidium* speciation by RFLP, together with 2 novel markers. Our results showed efficient and specific amplification of *Cryptosporidium* DNA. Further validation of these WGA techniques for routine subtyping of *Cryptosporidium* (GP60, double-stranded RNA and mini and micro-satellite repeats) would be necessary and desirable before adopting WGA for routine characterization of clinical and environmental isolates of *Cryptosporidium* species.

#### ACKNOWLEDGEMENT

Thanks are extended to Guy Robinson for helpful suggestions.

#### FINANCIAL SUPPORT

This work was partially supported by funds from the European Commission for the HEALTHY WATER project (FOOD-CT-2006-036306). The authors are solely responsible for the content of this publication. It does not

represent the opinion of the European Commission. The European Commission is not responsible for any use that might be made of data appearing therein.

#### REFERENCES

- Ahn, S., Costa, J. and Emanuel, J. R.** (1996). PicoGreen quantitation of DNA: effective evaluation of samples pre- or post-PCR. *Nucleic Acids Research* **24**, 2623–2625.
- Anderson, T., Haubold, B., Williams, J. T., Estrada-Franco, J. G., Richardson, L., Mollinedo, R., Bockarie, M., Mokili, J., Mharakurwa, S., French, N., Whitworth, J., Velez, I. D., Brockman, A. H., Nosten, F., Ferreira, M. U. and Day, K. P.** (2000). Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Molecular Biology and Evolution* **17**, 1467–1482.
- Barker, D., Hansen, M. S., Faruqi, A. F., Giannola, D., Irsula, O. R., Lasken, R. S., Latterich, M., Makarov, V., Oliphant, A., Pinter, J. H., Shen, R., Sleptsova, I., Ziehler, W. and Lai, E.** (2004). Two methods of whole-genome amplification enable accurate genotyping across a 2320-SNP linkage panel. *Genome Research* **14**, 901–907.
- Bergen, A., Haque, K. A., Qi, Y., Beerman, M. B., Garcia-Closas, M., Rothman, N. and Chanock, S. J.** (2005). Comparison of yield and genotyping performance of multiple displacement amplification and OmniPlex whole genome amplified DNA generated from multiple DNA sources. *Human Mutation* **26**, 262–270.
- Blanco, L., Bernad, A., Lázaro, J. M., Martín, G., Garmendia, C. and Salas, M.** (1989). Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. *The Journal of Biological Chemistry* **264**, 8935–8940.
- Bouzid, M., Steverding, D. and Tyler, K. M.** (2008). Detection and surveillance of waterborne protozoan parasites. *Current Opinion in Biotechnology* **19**, 302–306.
- Burgos, M., Méndez, J. C. and Ribon, W.** (2004). Molecular epidemiology of tuberculosis: methodology and applications. *Biomedica* **24** (Suppl. 1), 188–201.
- Cacciò, S., Homan, W., Camilli, R., Traldi, G., Kortbeek, T. and Pozio, E.** (2000). A microsatellite marker reveals population heterogeneity within human and animal genotypes of *Cryptosporidium parvum*. *Parasitology* **120**, 237–244.
- Cama, V., Arrowood, M. J., Ortega, Y. R. and Xiao, L.** (2006). Molecular characterization of the *Cryptosporidium parvum* IOWA isolate kept in different laboratories. *The Journal of Eukaryotic Microbiology* **53** (Suppl. 1), S40–S42.
- Cheung, V. and Nelson, S. F.** (1996). Whole genome amplification using a degenerate oligonucleotide primer allows hundreds of genotypes to be performed on less than one nanogram of genomic DNA. *Proceedings of the National Academy of Sciences, USA* **93**, 14676–14679.
- Dean, F., Hosono, S., Fang, L., Wu, X., Faruqi, A. F., Bray-Ward, P., Sun, Z., Zong, Q., Du, Y., Du, J., Driscoll, M., Song, W., Kingsmore, S. F., Egholm, M. and Lasken, R. S.** (2002). Comprehensive human genome amplification using multiple displacement



- amplification. *Proceedings of the National Academy of Sciences, USA* **99**, 5261–5266.
- Detter, J., Jett, J. M., Lucas, S. M., Dalin, E., Arellano, A. R., Wang, M., Nelson, J. R., Chapman, J., Lou, Y., Rokhsar, D., Hawkins, T. L. and Richardson, P. M.** (2002). Isothermal strand-displacement amplification applications for high-throughput genomics. *Genomics* **80**, 691–698.
- Eckert, K. and Kunkel, T. A.** (1991). DNA polymerase fidelity and the polymerase chain reaction. *PCR Methods and Applications* **1**, 17–24.
- Elwin, K., Chalmers, R. M., Roberts, R., Guy, E. C. and Casemore, D. P.** (2001). Modification of a rapid method for the identification of gene-specific polymorphisms in *Cryptosporidium parvum* and its application to clinical and epidemiological investigations. *Applied and Environmental Microbiology* **67**, 5581–5584.
- Esteban, J., Salas, M. and Blanco, L.** (1993). Fidelity of phi 29 DNA polymerase. Comparison between protein-primed initiation and DNA polymerization. *The Journal of Biological Chemistry* **268**, 2719–2726.
- Fiegler, H., Geigl, J. B., Langer, S., Rigler, D., Porter, K., Unger, K., Carter, N. P. and Speicher, M. R.** (2007). High resolution array-CGH analysis of single cells. *Nucleic Acids Research* **35**, e15.
- Goodgame, R., Genta, R. M., White, A. C. and Chappell, C. L.** (1993). Intensity of infection in AIDS-associated cryptosporidiosis. *The Journal of Infectious Diseases* **167**, 704–709.
- Goumenou, M. and Macherka, K.** (2004). Measurement of DNA single-strand breaks by alkaline elution and fluorometric DNA quantification. *Analytical Biochemistry* **326**, 146–152.
- Han, S., Zschausch, H. C., Meyer, H. G., Schneider, T., Loos, M., Bhakdi, S. and Maeurer, M. J.** (2000). *Helicobacter pylori*: clonal population structure and restricted transmission within families revealed by molecular typing. *Journal of Clinical Microbiology* **38**, 3646–3651.
- Hawkins, T., Detter, J. C. and Richardson, P. M.** (2002). Whole genome amplification – applications and advances. *Current Opinion in Biotechnology* **13**, 65–67.
- Langmore, J.** (2002). Rubicon Genomics, Inc. *Pharmacogenomics* **3**, 557–560.
- Leav, B., Mackay, M. R., Anyanwu, A., O'Connor, R. M., Cevallos, A. M., Kindra, G., Rollins, N. C., Bennish, M. L., Nelson, R. G. and Ward, H. D.** (2002). Analysis of sequence diversity at the highly polymorphic Cpgp40/15 locus among *Cryptosporidium* isolates from human immunodeficiency virus-infected children in South Africa. *Infection and Immunity* **70**, 3881–3890.
- Leoni, F., Gallimore, C. I., Green, J. and McLauchlin, J.** (2003) Molecular epidemiological analysis of *Cryptosporidium* isolates from humans and animals by using a heteroduplex mobility assay and nucleic acid sequencing based on a small double-stranded RNA element. *Journal of Clinical Microbiology* **41**, 981–992.
- Luthra, R. and Medeiros, L. J.** (2004). Isothermal multiple displacement amplification: a highly reliable approach for generating unlimited high molecular weight genomic DNA from clinical specimens. *The Journal of Molecular Diagnostics* **6**, 236–242.
- Mallon, M., MacLeod, A., Wastling, J. M., Smith, H. and Tait, A.** (2003). Multilocus genotyping of *Cryptosporidium parvum* Type 2: population genetics and sub-structuring. *Infection, Genetics and Evolution* **3**, 207–218.
- Nelson, J., Cai, Y. C., Giesler, T. L., Farchaus, J. W., Sundaram, S. T., Ortiz-Rivera, M., Hosta, L. P., Hewitt, P. L., Mamone, J. A., Palaniappan, C. and Fuller, C. W.** (2002). TempliPhi, phi29 DNA polymerase based rolling circle amplification of templates for DNA sequencing. *Biotechniques* (Suppl) 44–47.
- Park, J., Beaty, T. H., Boyce, P., Scott, A. F. and McIntosh, I.** (2005). Comparing whole-genome amplification methods and sources of biological samples for single-nucleotide polymorphism genotyping. *Clinical Chemistry* **51**, 1520–1523.
- Paunio, T., Reima, I. and Syvänen, A. C.** (1996). Preimplantation diagnosis by whole-genome amplification, PCR amplification, and solid-phase minisequencing of blastomere DNA. *Clinical Chemistry* **42**, 1382–1390.
- Pinchbeck, G., Morrison, L. J., Tait, A., Langford, J., Meehan, L., Jallow, S., Jallow, J., Jallow, A. and Christley, R. M.** (2008). Trypanosomiasis in The Gambia: prevalence in working horses and donkeys detected by whole genome amplification and PCR, and evidence for interactions between trypanosome species. *BMC Veterinary Research* **4**, 7–13.
- Shoab, M., Baconnais, S., Mechold, U., Le Cam, E., Lipinski, M. and Ogryzko, V.** (2008). Multiple displacement amplification for complex mixtures of DNA fragments. *BMC Genomics* **9**, 415–428.
- Singer, V., Jones, L. J., Yue, S. T. and Haugland, R. P.** (1997). Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation. *Analytical Biochemistry* **249**, 228–238.
- Smith, H., Cacciò, S. M., Tait, A., McLauchlin, J. and Thompson, R. C.** (2006). Tools for investigating the environmental transmission of *Cryptosporidium* and *Giardia* infections in humans. *Trends in Parasitology* **22**, 160–167.
- Smith, H., Cacciò, S. M., Cook, N., Nichols, R. A. and Tait, A.** (2007). *Cryptosporidium* and *Giardia* as foodborne zoonoses. *Veterinary Parasitology* **149**, 29–40.
- Snabes, M., Chong, S. S., Subramanian, S. B., Kristjansson, K., DiSepio, D. and Hughes, M. R.** (1994). Preimplantation single-cell analysis of multiple genetic loci by whole-genome amplification. *Proceedings of the National Academy of Sciences, USA* **91**, 6181–6185.
- Spano, F., Putignani, L., McLauchlin, J., Casemore, D. P. and Crisanti, A.** (1997). PCR-RFLP analysis of the *Cryptosporidium* oocyst wall protein (COWP) gene discriminates between *C. wrairi* and *C. parvum*, and between *C. parvum* isolates of human and animal origin. *FEMS Microbiology Letters* **150**, 209–217.
- Strong, W., Gut, J. and Nelson, R. G.** (2000). Cloning and sequence analysis of a highly polymorphic *Cryptosporidium parvum* gene encoding a 60-kilodalton glycoprotein and characterization of its 15- and

- 45-kilodalton zoite surface antigen products. *Infection and Immunity* **68**, 4117–4134.
- Tanriverdi, S., Arslan, O. M., Akiyoshi, D. E., Tzipori, S. and Widmer, G.** (2003). Identification of genotypically mixed *Cryptosporidium parvum* populations in humans and calves. *Molecular and Biochemical Parasitology* **130**, 13–23.
- Telenius, H., Carter, N. P., Bebb, C. E., Nordenskjöld, M., Ponder, B. A. and Tunnacliffe, A.** (1992). Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics* **13**, 718–725.
- Thorstenson, Y., Hunicke-Smith, S. P., Oefner, P. J. and Davis, R. W.** (1998) An automated hydrodynamic process for controlled, unbiased DNA shearing. *Genome Research* **8**, 848–855.
- Zhang, L., Cui, X., Schmitt, K., Hubert, R., Navidi, W. and Arnheim, N.** (1992). Whole genome amplification from a single cell: implications for genetic analysis. *Proceedings of the National Academy of Sciences, USA* **89**, 5847–5851.