Template-dependent multiple displacement amplification for profiling human circulating RNA

Weihua Wang
Saint Louis University

Yi Ren
Saint Louis University

Yang Lu
Saint Louis University

Yuan Xu
Saint Louis University

Seth D. Crosby
Washington University School of Medicine in St. Louis

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

Recommended Citation
Wang, Weihua; Ren, Yi; Lu, Yang; Xu, Yuan; Crosby, Seth D.; Di Bisceglie, Adrian M.; and Fan, Xiaofeng, "Template-dependent multiple displacement amplification for profiling human circulating RNA." BioTechniques. 63, 1. 21-27. (2017).
https://digitalcommons.wustl.edu/open_access_pubs/6024

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Authors
Weihua Wang, Yi Ren, Yang Lu, Yuan Xu, Seth D. Crosby, Adrian M. Di Bisceglie, and Xiaofeng Fan
Template-dependent multiple displacement amplification for profiling human circulating RNA

Weihua Wang1,4, Yi Ren1, Yang Lu1, Yuan Xu1, Seth D. Crosby3, Adrian M Di Bisceglie1,2, and Xiaofeng Fan1,2*

1Division of Gastroenterology & Hepatology, Department of Internal Medicine, Saint Louis University School of Medicine, St. Louis, MO, 2Saint Louis University Liver Center, Saint Louis University School of Medicine, St. Louis, MO, 3Department of Genetics, Washington University School of Medicine, St, Louis, MO, and 4Wuhan Pulmonary Hospital, Wuhan, Hubei, China

Multiple displacement amplification (MDA) is widely used in whole-genome/transcriptome amplification. However, template-independent amplification (TIA) in MDA is a commonly observed phenomenon, particularly when using high concentrations of random hexamer primers and extended incubation times. Here, we demonstrate that the use of random pentamer primers with 5’ ends blocked by a C18 spacer results in MDA solely in a template-dependent manner, a technique we have named tdMDA. Together with an optimized procedure for the removal of residual genomic DNA during RNA extraction, tdMDA was used to profile circulating RNA from 0.2 mL of patient sera. In comparison to regular MDA, tdMDA demonstrated a lack of quantifiable DNA amplification in the negative control, a remarkable reduction of unmapped reads from Illumina sequencing (7 ± 10.9% versus 58.6 ± 39%, P = 0.006), and increased mapping rates of the serum transcriptome (26.9 ± 7.9% versus 5.8 ± 8.2%, P = 3.8 × 10^-4). Transcriptome profiles could be used to separate patients with chronic hepatitis C virus (HCV) infection from those with HCV-associated hepatocellular carcinoma (HCC). We conclude that tdMDA should facilitate RNA-based liquid biopsy, as well as other genome studies with biological specimens having ultralow amounts of genetic material.

METHOD SUMMARY

Random pentamer primers blocked at the 5’ end with a C18 spacer were shown to eliminate template-independent amplification, a major problem in multiple displacement amplification.
options have not been widely adopted. Increasing the amount of template (≥10 ng) and using a short incubation time, as recommended in many commercial kits such as the REPLI-g kits from QIAGEN, is a common option to minimize TIA, yet such an amount (10 ng) is not always available in many biological scenarios. Here, we have developed a robust MDA protocol that is free of TIA through the use of 5′ end-blocked random pentamer primers. Isothermal amplification occurs solely in a template-dependent manner, which we have named tdMDA. The potential of tdMDA has been demonstrated by an efficient profiling of human circulating RNA, a biological source notorious for its ultralow concentration and complex heterogeneity (17).

Materials and methods

Patient samples

A total of 30 serum samples were used in the current study. Of these, 20 were from a completed study (18), and 3 samples from blood donors were used to optimize the experimental protocols for RNA extraction, genomic DNA removal, and tdMDA. The remaining seven samples were collected from patients at the Saint Louis University Hospital. Three patients had chronic HCV infection, and the other four were diagnosed with HCV-associated hepatocellular carcinoma (HCC) by computed tomography (CT), serum α fetoprotein (AFP) measurement, and histological examination. These samples were used to test the feasibility of circulating RNA profiling through the use of tdMDA.

Ethics statement

For the collection of serum samples, written informed consent was obtained from each patient and was approved by the local Institutional Review Board in the parental studies (19). The research protocol for the collection and use of patient samples was reviewed and approved by the Saint Louis University Hospital.

Figure 1. List of primers tested in multiple displacement amplification (MDA). Group 1 primers have phosphorothioate bonds placed at different positions to test the minimum lengths of primers capable of initiating DNA polymerization. Group 2 primers are modified internally with a C3 spacer to minimize self-priming. Group 3 primers are the group 2 primers modified with the inclusion of locked nucleic acid (LNA) to stabilize the binding between primers and templates. Groups 4 and 5 are constrained primers and RNA primers (with or without a C3 spacer), respectively. Random primers blocked at the 5′ end are designated as group 6 (hexamer) or group 7 (pentamer) primers. In these two groups of primers, the blockage at the 5′ end was tested using different modifications, including inverted, 2′,3′ dideoxy-dT base (5′ inverted ddT), dSpacer, C3 spacer, C9 spacer, C18 spacer, amino modifier C6 (AmMC6) or alkyne. Degenerate bases are indicated using standard International Union of Pure and Applied Chemistry (IUPAC) codes. Each primer was tested in triplicate at a concentration of 50 µM in the standard MDA protocol with (blue) or without (red) template [1 ng of the hepatitis C virus (HCV) plasmid]. The MDA reaction with the group 7 primers was incubated at 28°C, while the other groups of primers were amplified at 30°C. The error bars indicate SD.

**Standard MDA protocol**

In order to facilitate experimental optimization, we defined a standard MDA protocol with a reaction volume of 40 µL and consisting of phi29 DNA polymerase and its corresponding buffer, 1 mM of dNTPs, and 2 µM or 50 µM of random exonuclease-resistant primers. A recombinant 12-kb plasmid (HCV plasmid) containing a 9-kb HCV genotype 1a genome from our previous study (20) was used as a positive control. The reaction was incubated at 30°C for 16 h, followed by heat-inactivation at 65°C for 10 min. In addition to visualization on a 0.8% agarose gel, MDA yield was quantitated by the Quant-iT PicoGreen dsDNA Assay kit (QIAGEN, Valencia, CA) after purification using the QIAamp DNA mini kit. Due to the extreme sensitivity of the techniques (5), all MDA experiments were performed in strict adherence to the recommendations of Champlot et al. to avoid possible contamination (21).

**Elimination of TIA in MDA**

The extreme sensitivity of MDA also causes confusion between TIA and reagent contamination, wherein the latter is caused by trace amounts of bacterial or other microbial genomes left during manufacture. Without additional experimentation, it was difficult to clarify if the positive amplification in negative controls resulted from TIA or reagent contamination. Therefore, a multi-step strategy was applied to eliminate TIA in MDA. First, using the standard MDA protocol, we evaluated the purity of phi29 DNA polymerases from various vendors, including Lucigen (Middleton, WI), Epicentre (Madison, WI), Thermo Scientific (St. Louis, MO), and New England Biolabs (Ipswich, MA). Decontamination through UV irradiation as described by Woyke et al. was also adopted for the reagents, including phi29 DNA polymerase, dNTPs, and primers (8). The procedure was conducted with the Stratagene Strata linker 2400 UV Crosslinker (Agilent Technologies, Santa Clara, CA). Next, using the standard MDA protocol with reagents that were free of contamination, elimination of TIA was evaluated by the use additives such as DMSO (0.5~5%) (Sigma-Aldrich, St. Louis, MO), Escherichia coli SSB (50~300 µg/mL) (New England Biolabs), and Fc400 (4~64 µg/mL) (Sigma-Aldrich) to emulate a crowded intracellular environment (22). The incubation temperature was tested from 25°C to 35°C.

Finally, we focused on primer optimization by designing a total of 36 different types of primers that could be categorized into 7 groups (Figure 1). Primers containing locked nucleic acid (LNA) were synthesized by Exiqon (Woburn, MA). All remaining primers were purchased from Integrated DNA Technologies (IDT, Coralville, IA). Each primer was tested in triplicate at 50 µM, a typical concentration used in MDA (1,2).

**Serum RNA extraction, genomic DNA removal, and real-time RT-tdMDA**

Total RNA was extracted from 200 µL serum using the miRNeasy Serum/Plasma kit (QIAGEN) and then eluted into 14 µL water. Extracted RNA was quantitated using the Quant-iT RiboGreen RNA kit (Life Technologies,
Purified RNA (9.4 µL) was reverse transcribed in a 20 µL reaction consisting of 1× SuperScript III buffer (Life Technologies), 10 mM DTT, 2 mM dNTPs, 20 U RNaseIn RNase Inhibitor (Promega, Madison, WI), 200 U SuperScript III reverse transcriptase and 80 µM primers. The reaction was incubated at 37°C for 30 min, 50°C for 30 min, and then inactivated at 70°C for 15 min. An aliquot of 4 µL of this reverse transcription (RT) reaction served as the template in the subsequent MDA. Both RT and MDA used the same random primer capable of eliminating TIA. Using three serum samples with different RNA concentrations, we conducted real-time RT-tdMDA with 0.1× SYBR Green I (Thermo Fisher Scientific) included in the tdMDA step. The reaction was incubated at 28°C for 24 h on the ABI 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA). The resulting curves of fluorescent intensity were used to determine an optimal incubation time for RT-tdMDA applied to HCV and HCC patient samples.

Illumina sequencing

The optimized procedures from RNA extraction and purification were applied to serum samples from seven HCV and HCC patients. The cDNA from the RT reaction was amplified by both tdMDA and the regular MDA protocol using random hexamer primers (primer 6NS2 listed in Figure 1). Based on the product size (~20 kb), the purified RT-MDA product was diluted to 0.4 ng/µL and enzymatically sheared by incubating the DNA at 55°C for 5 min using the Nextera XT transposome from the Nextera XT DNA Sample Preparation kit (Illumina, San Diego, CA), which added adapter sequences simultaneously. Fragmented DNA was amplified with 12 cycles of PCR using the Index 1 and Index 2 primers from Illumina, followed by purification and size selection using AMPure XP beads (Beckman Coulter, Beverly, MA). Size-selected libraries were normalized using the normalization beads from Illumina. In this step, the quantity of each library was normalized to ensure equal representation in the pooled sample. An aliquot of 2 µL normalized and pooled library was added to 998 µL pre-chilled hybridization buffer, denatured at 98°C for 2 min, and then sequenced on the Illumina NextSeq 500 Sequencing platform (1 × 250-bp single reads and mid-output) at MOgene (St. Louis, MO).

Data analysis

Raw sequence reads in fastq format were first filtered in PRINSEQ (v0.19.5) (18) for quality control, including read length ≥70 bp, mean read quality score ≥25, low complexity with DUST score ≤7, ambiguous bases ≤1%, and all duplicates (18). By subtracting the human reference genome, NCBI reference sequences for bacteria, viruses and fungi, and NCBI nt database (6), we calculated and compared the rates of unmapped sequencing reads for each sample amplified by either tdMDA or regular MDA. Potential gene detection and differential gene expression (DGE) between HCV and HCC patients were analyzed by the TopHat-featureCount-edgesR pipelines (23,24). In brief, filtered reads were mapped in TopHat onto the annotated human reference genome (Ensembl GRCh 38.84). Gene-counting data were summarized at the transcript-level from TopHat output in featureCounts (25), followed by DGE analysis in edgeR with the false discovery rate (FDR) set at 10%.

Statistical analysis

Except for built-in statistical analyses in programs such as Cufflinks and edgeR, other comparisons were done using the two-tailed Student’s t-test. Data were expressed as the mean ± SD, and P < 0.05 was considered statistically significant.
Results and discussion

Reagent contamination and TIA are two major concerns in MDA. While these issues are negligible when using large amounts of template, both concerns become significant shortcomings when working with low or ultralow amounts of template (typically <1 ng) since they compress the product space of target molecules and/or aggravate qualitative biases (2,6,7,26). While reagent contamination could be minimized or even eliminated through the use of clean enzymes and reagents (27,28), TIA remains a consistent issue in MDA, despite many attempts to overcome it. The current study provides a simple, but efficient solution in which the use of a specially designed primer (C28) (Figure 1) eliminates TIA.

In a stepwise optimization process, we first identified that phi29 DNA polymerase from Epicentre was considerably free of reagent contamination in comparison to other enzymes (Supplementary Figure S1). UV treatment of the Epicentre phi29 DNA polymerase eliminated contamination, but MDA yield was reduced, while UV treatment of dNTPs or random hexamer primers abolished or compromised the reaction (Supplementary Figure S2). In the standard MDA protocol using the phi29 enzyme from Epicentre, all additives had no apparent role in eliminating TIA (Supplementary Figure S2). Incubation temperatures >31°C reduced yield, while temperatures <30°C enhanced TIA (Supplementary Figure S2). These results led us to focus on primer design to eliminate TIA in MDA.

As expected from the strong 3´→5´ exonuclease activity of phi29 DNA polymerase, which digests oligonucleotides (29), Primer 6N of the Group 1 primers produced no DNA amplification. Primers C5 and C6 failed to amplify the HCV plasmid, while the performance of Primer C7 was similar to the random hexamer primer (Primer 6NS2) (Figure 1). Internal modification of the C3 spacer caused steric hindrance. However, such hindrance, as shown with the Group 2 primers, cannot differentiate between self and primer-template priming, which was not improved by the inclusion of LNA bases (Group 3). Despite the theoretical possibility (30), RNA primers (Group 4) were incapable of inhibiting TIA. The constrained primers in Group 5 resulted in an overall reduction of MDA yields, and TIA was still observed.

Substantial suppression of TIA was found in MDA using 5´ end-blocked random hexamer (Group 6) or pentamer (Group 7) primers (Figure 1). Of the eight options to block 5´ ends of random pentamer primers, consecutive dSpacer (1´,2´-dideoxyribose) (Primer C22), amino modifier C6 (Primer C30) and hexynyl (Primer C31) produced no amplification at all. While other options showed a substantial suppression of TIA, consistent suppression of TIA resulting in increased reproducibility was only achieved using the C18 spacer modification (Primer C28), suggesting the subtle influence of the different blocking options (Figure 1). Given the decreased melting temperatures of random pentamer primers in comparison to random hexamer primers, the incubation temperature in MDA with the Group 7 primers was initially reduced from 30°C to 28°C. In subsequent experiments with Primer C28, this temperature appeared to be optimal with regard to the complete elimination of TIA and provided MDA yields comparable to the use of the random hexamer primers with or without C18 spacer–blocked 5´ ends (Figure S3).

The tdMDA reaction could tolerate primer concentrations up to 400 µM (Supplementary Figure S4). Beyond this threshold, yield was reduced dramatically, perhaps due to decreased Mg2+ concentration caused by binding to primer (31). For HCV plasmid amounts ranging from 1 ng to 50 pg, tdMDA yields were maximized with primer concentrations of 80–100 µM. When the input amount was lower than 50 pg in a 40-µL reaction, tdMDA required templates to
be freshly prepared. After being resuspended in water, Primer C28 with the C18 spacer modification was considerably stable. No apparent reduction in MDA yields was seen when using Primer C28 subjected to 30 thaw-freeze cycles (Supplementary Figure S5). Similarly, tdMDA was reproducible when using Primer C28 stored up to 6 months at -20°C (Supplementary Figure S5).

The mechanisms for TIA generation in MDA are not fully understood. However, based on the observations from the current study, we speculate there are two possibilities: (i) polymerase jumping and (ii) self-priming. It has long been known that Taq DNA polymerase is capable of jumping among templates (32). We believe that this kind of activity might be enhanced by phi29 DNA polymerase due to the polymerase’s inability to be used as a molecular brake in comparison to Taq DNA polymerase (Sahu, S. PhD thesis, Duke University, 2007). Polymerase jumping among random primers could form artificial templates for the generation of TIA. This process was abolished by blocking the 5’ ends of random primers. On the other hand, self-priming is thought to be a common mechanism responsible for artifacts in PCR and MDA (12). Our data showed no product was generated in MDA with Primer C5, which was actually 4-bp in length after digestion owing to the exonuclease activity of phi29 DNA polymerase. Thus, efficient polymerization by phi29 DNA polymerase may require 5-bp priming, which would potentially allow self-priming among random hexamer (6 bp) primers. This may explain why TIA cannot be eliminated entirely through the use of random hexamer primers with 5’ ends blocked by the C18 spacer (Figure 1). Primers C1 and C2 have an internal modification of the C3 spacer that theoretically prevents self-priming. Primer C25 is a random pentamer primer without modification. The use of these primers produces a significant amount of TIA (Figure 1). Based on these observations, polymerase jumping, rather than self-priming, might be the critical factor for TIA formation in MDA.

Finally, tdMDA was evaluated for profiling circulating RNA. The average RNA recovery from 200 µL serum was 10.76 ± 2.8 ng using the QIAGEN miRNeasy Serum/Plasma kit. After optimizing genomic DNA removal (Supplementary Table S1), 3 serum samples with various RNA yields were selected for real-time RT-tdMDA with a 24-h incubation. While DNA accumulation was not observed in the negative control (water), amplification of all 3 samples with 10-fold serial dilutions plateaued after an approximately 18-h incubation (Figure 2). The protocol was then applied to the seven HCV and HCC patient samples (Figure 3A). Illumina sequencing of RT-tdMDA product provided an average of 1.45 million reads per sample. Unmapped reads were significantly lower in tdMDA than regular MDA with random hexamer primers (7 ± 10.9% versus 58.6 ± 39%; P = 0.006) (Figure 3B). Data generated via tdMDA showed a much higher mapping rate than regular MDA in TopHat (26.9 ± 7.9% versus 5.8 ± 8.2%; P = 3.8 × 10^-4) (Figure 3B). The TopHat-feature Counts-edgeR analysis detected 3459 genes in human circulating RNA. Of these 3459 genes, 457 showed ≥2 fold changes, either up- (n = 424) or down-regulated (n = 23) in the HCC group (Figure 2C). Notably, unsupervised clustering analysis showed a clear separation of HCV from HCC based on the leading log-fold-change, which is the root-mean-square average of the largest log2-fold-changes between each pair of samples (Figure 2D) (24).

Here, we reported a simple but reliable solution for the elimination of TIA in MDA. Template-dependent amplification should extend MDA’s application to biological specimens with low or ultralow amounts of genetic material. By applying tdMDA, we demonstrated that efficient profiling of circulating RNA can be achieved from as little as 0.2 mL patient serum. Substantial differences in transcriptome profiles between HCV and HCC patients further demonstrate the promise of tdMDA for RNA-based liquid biopsy applications.
Author contributions
X.F. and A.M.D. conceived and designed the experiments. W.W., Y.R., Y.L., and Y.X. performed the experiments. S.G., A.M.D., and X.F. interpreted the data. X.F. analyzed the data and wrote the paper.

Acknowledgments
This work was supported by the US National Institutes of Health (NIH) grants DK080711 (X.F), AI111047 (X.F.), and a seed grant from the Saint Louis University Liver Center (X.F.). This paper is subject to the NIH Public Access Policy.

Competing interests
The authors declare no competing interests.

References

Received 01 March 2017; accepted 25 May 2017.

Address correspondence to Xiaofeng Fan or Adrian M. Di Bisceglie, Division of Gastroenterology & Hepatology, Department of Internal Medicine, Saint Louis University School of Medicine, St. Louis, MO 63104. E-mail: fanx@slu.edu (X.F.) or dibiscam@slu.edu (A.M.D.)

To purchase reprints of this article, contact: biotechniques@fosterprinting.com