Dilute-‘N’-Go dideoxy sequencing of all DNA strands generated in multiplex LATE-PCR assays

Yanwei Jia, Adam Osborne, John E. Rice and Lawrence J. Wangh*

Biology Department, Brandeis University, 415 South St., Waltham, MA 02155, USA

Received August 20, 2009; Revised February 2, 2010; Accepted February 8, 2010

ABSTRACT

We have recently described a Dilute-‘N’-Go protocol that greatly simplifies preparation and sequencing of both strands of an amplicon generated using linear-after-the-exponential (LATE)-PCR, an advanced form of asymmetric PCR. The same protocol can also be used to sequence all limiting primer strands in a multiplex LATE-PCR, by adding back each of the depleted limiting primers to a separate aliquot of the multiplex reaction. But, Dilute-‘N’-Go sequencing cannot be used directly to sequence each of the excess primer strands in the same multiplex reaction, because all of the excess primers are still present at high concentration. This report demonstrates for the first time that it is possible to sequence each of the excess primer strands using a modified Dilute-‘N’-Go protocol in which blockers are added to prevent all but one of the excess primers serving as the sequencing primer in separate aliquots. The optimal melting temperatures, positions and concentrations of blockers relative to their corresponding excess primers are defined in detail. We are using these technologies to measure DNA sequence changes in mitochondrial genomes that accompany aging and exposure to certain drugs.

INTRODUCTION

A host of novel sequencing methods are being developed for rapidly sequencing whole genomes by in silico splicing of short overlapping DNA sequences (1–5). Nevertheless, dideoxy sequencing remains the method of choice for sequencing PCR amplicons up to 1000 bases long, even though preparation and clean-up of such amplicons prior to sequencing is still a time consuming and expensive bottleneck to high-throughput dideoxy sequencing. The conventional approach to sample preparation is to remove unused primers, nucleotides and enzyme present at the end of amplification by filtration, chromatography or gel electrophoresis and then add a standard sequencing cocktail (6). Enzymatic methods have also been employed for the removal of interfering components (7–9). None of these procedures, however, solves the problem of sequencing all of the strands generated in a single tube multiplex reaction.

Linear-after-the-exponential (LATE)-PCR is an advanced form of asymmetric PCR that uses a limiting primer (LP) and an excess primer (XP) to exponentially amplify double-stranded DNA efficiently and then predictably switches to linear amplification of single-stranded DNA (10,11). Limited amount of LP strand, the DNA strand extended from the LP, has been generated. While the LP is no longer present at the end of the reaction and a high concentration of the XP strand has accumulated, unused XP is still abundant. The LP strand can readily be sequenced by simply diluting an aliquot of the reaction into distilled water containing a few picomoles of the LP used for amplification of the double-stranded amplicon. The solution is so diluted that the LP strand as the template for sequencing the XP strand has been diluted out. Therefore, no XP strand sequence can be obtained. The added LP primer serves as the sequencing primer while the XP strand serves as template when an aliquot of the mixture is added to a cocktail for commercial dideoxy sequencing (12). We have also adapted this approach for pyrosequencing (13). In the case of dideoxy sequencing, a second aliquot of a monoplex reaction can be diluted and added to a commercial sequencing cocktail without addition of the LP. In this case, the XP still present in the reaction serves as the sequencing primer and the LP strand serves as template. We have recently shown that sequencing carried out in this way is accurate, even when PCR amplification is initiated with as little as a single DNA molecule (14).

Multiplex LATE-PCR assays generate separate single-stranded XP strand for each of the amplicons. Each of these strands can be used as a sequencing template by spiking separate aliquots of the reaction...
products with one of the LPs, all of which have been depleted. The fidelity of the resulting set of sequences can be confirmed by comparison to each of the corresponding sequences generated in its own monoplex reaction (12). But, the fidelity of the set of sequences derived from the multiplex reaction cannot be confirmed by sequencing each of the complementary strands in the same multiplex reaction using the standard Dilute-'N'-Go protocol. This is because all of the XPs are still present in the reaction product and can generate a set of overlapping sequences.

In this paper, we introduce the use of blockers in conjunction with Dilute-'N'-Go sequencing in multiplex LATE-PCR assays. A blocker (BLK) is designed to inhibit hybridization of the 3'-end of the XP to its binding site on target, thereby preventing primer elongation during each thermal cycle of the sequencing reaction. All but one of the XPs is blocked in each aliquot of the diluted multiplex reaction product. Only the XP that is not blocked in that aliquot functions as a sequencing primer. The conditions for using this strategy were optimized using a duplex LATE-PCR assay and then applied to a triplex assay comprised of three amplicons for the human mitochondrial genome. This assay is being built and tested in order to be able to monitor mutagenesis at the single genome level caused by either aging or azidothymidine (AZT), a drug used to inhibit the HIV reverse transcriptase.

MATERIALS AND METHODS

Synthesis and purification of oligonucleotides

Primers:

- **G269 LP:** 5'-CGAGGAGTATTAGTACTCTTGTGGCGGG A-3' BHQ1
- **HV1 LP:** 5'-CGAGGAGAATGATCTCTTATGCGGG A-3' BHQ1
- **Globin LP:** 5'-CGGGGTATCTAAGGAGAACTCCTGCT -3' BHQ1
- **Cytb LP:** 5'-ACCTCTATTCTTGGACAGGATGCAG -3' BHQ1

were synthesized and salt-free purified by Eurofins MWG Operon (AL, USA).

Blockers:

- **G269 BL:** 5'-CGGCTGTCGCCCTTACATAGTCTACGAC -3' BHQ1
- **HV1 BL:** 5'-GAGGAGAGATGATCTCTTATGCGGG A-3' BHQ1
- **Globin BL:** 5'-TGGGTTTGTATAGGAGACGCTTCTCTTGACCT TGGG-3' BHQ1
- **Cytb BL:** 5'-ACCTCTATTCTTGGACAGGATGCAG -3' BHQ1

were synthesized and double HPLC purified by Biosearch Technologies (CA, USA).

LATE-PCR conditions

All PCR reactions had the following reagents: 1× PCR buffer, 0.2 mM dNTPs, 3 mM MgCl2, 2 units of Platinum Taq DNA polymerase (InVitrogen, CA, USA), 50 nM of each LP, 1 μM of each XP and 0.24× SYBR Green I in a 25 μl volume. For G269 & HV1 duplex and TSD & Globin duplex, 300 nM PrimeSafe™ I (12) and 1000 genomes of purified human genomic DNA were included. For mitochondrial DNA HV1, HV2 and Cyt b triplex, 300 nM PrimeSafe™ II and 100 K copies of mitochondrial DNA from purified human DNA were included. LATE-PCR amplification and fluorescence detection were carried out in a Stratagene Mx 3005 PCR thermocycler.

Thermal cycling profile for G269 & HV1 duplex and TSD & Globin duplex included an initial denaturation step of 5 min at 95°C, followed by 55 cycles of 95°C for 10 s, 64°C for 20 s and 72°C for 1 min, then a melting from 55 to 95°C with a 1°C increase. The thermocycler temperature program for the mitochondrial DNA HV1, HV2 and Cyt b triplex consisted of denaturation at 95°C for 5 min, followed by 60 cycles of 95°C for 30 s, 62°C for 20 s and 72°C for 45 s, then a melting from 35 to 95°C with a 1°C increase. The completed LATE-PCR samples were kept at −20°C until sequenced. The fluorescence reading in all assays was taken at 72°C.

Sample preparation for dideoxynucleotide sequencing

No purification of the LATE-PCR product was performed. Dideoxy sequencing was performed by GENEWIZ, Inc (South Plainfield, NJ) using Applied Biosystems BigDye version 3.1. The reactions were then run on Applied Biosystems’s 3730x1 DNA analyzer. Standard cycle sequencing includes an initial denaturation at 96°C for 1 min, followed by 25 cycles of 96°C for 10 s, 50°C for 5 s and 60°C for 1 min. The premixed samples were diluted 2× in GENEWIZ before the reaction. For the sequencing of the limiting primer strand, 1 μl DNA from a 10× diluted LATE-PCR that has completed
~20–30 linear cycles is diluted to 12 µl in distilled water containing 1 µM of the same LP used in the LATE-PCR or another inner sequencing primer. For the sequencing of the XP strand, 1 µl DNA from LATE-PCR is diluted to 12 µl with 1 µM of each BLK, unless specified otherwise in the text. In order to obtain the whole sequence or the second half of the HV2 amplicon, the HV2 XP was blocked along with the other XPs in the triplex reaction and a separate downstream sequencing primer was added at 1 µM.

\[ T_m \]

\( T_m \) calculation

\( T_m \) is defined as the temperature at which, for a given concentration of primer or BLK at the start of the reaction, half of the target is bound to that primer or BLK to form a heterodimer. The \( T_m \) of each primer or BLK can be obtained from the corresponding binding curve to its target simulated by VisualOMP V7 (DNA software, CA) as shown in Figure 1b. The simulation is performed under sequencing conditions with all target concentration set as 2.5 nM, all XP concentrations set at 25 nM and all BLK concentrations set at 500 nM unless described otherwise in the text.

RESULTS

Blockers: rationale and proof of concept

All of the multiplex LATE-PCR assays described here contained 50 nM of each LP and 1 µM of each XP at the start of the reaction. At the end of the exponential phase of amplification all LPs are depleted, but each XP is still present at ~500 nM after 20–30 additional cycles of linear amplification. If an aliquot of a multiplex reaction is simply diluted into a commercial sequencing cocktail, each XP will hybridize to its specific target and will generate a ladder of oligonucleotides. All such ladders will overlap making it impossible to discern individual sequence. Therefore, to use the Dilute-'N'-Go protocol for sequencing each of the XP strands generated in a multiplex LATE-PCR assay it is necessary to block extension of all-but-one of the XPs present in the reaction. This is accomplished by adding all possible sets of all-but-one BLKs to parallel aliquots of the reaction products. In order to be effective all BLKs have to: (i) specifically hybridize to their intended LP strand; (ii) overlap the 3'-end of the corresponding XP and (iii) bind to their target sequences more rapidly than the corresponding XP.

Figure 1a shows the simplest way to design BLKs to meet these criteria. The BLKs for both G269 and HV1 are both oligonucleotides 5–10 bases longer on their 3'-ends than the XPs with which they compete. The 3'-end of both BLKs is capped with a BHQ-1 (Black Hole Quencher-1) when ordered from Biosearch Technologies although a carbon linker is equally effective (data not shown). Each BLK is perfectly matched to its target and therefore has a \( T_m \) at least 5°C higher than that of its corresponding XP. \textit{In silico} melting profiles show that the two BLKs saturate the targets before the XPs as the temperature drops (Figure 1b).

The Dilute-'N'-Go sequencing results, Figure 1c, show that the addition of the HV1 BLK resulted in a clean sequence for G269 XP strand, while addition of the G269 BLK to a second aliquot resulted in a clean sequence for HV1 XP strand. Omission of both BLKs, but addition of the LP for G269 to a third aliquot generated a clean sequence for the G269 LP strand, while addition of the LP for HV1 to a forth aliquot generated a sequence for the HV1 LP stand. The sequences of the two HV1 and the two G269 strands are complementary when aligned right-to-left/left-to-right. It is worth emphasizing that preparation of the four aliquots by dilution with distilled water took only seconds to carry out.

Dissecting the elements of blocker design

\textit{Position effect}. Both BLKs in Figure 1 out compete their respective XPs by binding to the entire target sequence plus additional bases at the 3'-end. But, the data shown in Figure 2 demonstrate that several BLKs with similar melting temperatures are equally effective provided that they overlap with at least one nucleotide on the 3'-end of the XP. In the examples shown here the lengths of all the BLKs were adjusted to keep the \( T_m \) of each BLK ~10°C higher than that of the XP, 57.8°C.
This guarantees that as the temperature is lower the BLK is positioned on the target before the XP has a chance to bind.

Figure 2 shows the G269 XP strand sequence in the absence or presence of one of several possible HV1 BLKs. When no BLK was added, both the G269 and HV1 sequences are generated and overlap. When only a single nucleotide of the HV1 BLK overlaps the 3' end of XP the base call of the G269 sequence is accurate, although there is a slight background of the HV1 sequence. Additional experiments, however, demonstrate that blocking does not occur if the blocker binds to the target beyond the 3' end of the XP (data not shown).

Blocker $T_m$ effect. In order to examine the relationship between BLK $T_m$ and XP $T_m$ independently of other factors, we designed four BLKs of the HV1 XP that have zero, one, two or three mismatches with the corresponding target, Figure 3. Each of the mismatches causes the $T_m$ of the BLK to decrease a few degrees. One HV1 BLK, HV1_BL, is perfectly complementary to the target and has a $T_m$ 14°C higher than that of the HV1 XP. The target is completely bound by the BLK prior to binding the XP. The $T_m$ of the BLK with one mismatch, HV1_BL6, is 66°C. While BLKs with two mismatches, HV1_BL7, and with three mismatches, HV1_BL8, have $T_m$s of 60 and 57.2°C respectively, which are just 2°C higher or even lower than the $T_m$ of HV1 XP.

As shown in Figure 3, a clean G269 sequence was obtained when the $T_m$ of the BLK was more than 8°C above the $T_m$ of HV1 XP. But, the G269 sequence was not free of interfering sequences when the $T_m$ of the BLK was only 2°C higher than the $T_m$ of the HV1 XP.

Concentration effect. The melting temperature of two complementary oligonucleotide strands depends in part on their concentrations (15–17). $T_m$ of an XP is higher at high concentration than it is at low concentration.

Each XP in the HV1 & G269 duplex reaction is present at 1000 nM and is estimated at ~500 nM after 20–30 cycles of linear amplification. When the amplification products are then diluted and sequenced using the Dilute-N-Go protocol the final concentration of the XP is ~25 nM. Overall this translates into a decrease in the $T_m$ of the HV1 XP from 64 to 57.8°C.

We designed an HV1 BLK, HV1-EB, which has the same sequence as the XP but capped at the 3' end (Figure 4a). The $T_m$ of HV1-EB increases from 53.5 to 65.9°C when its concentration is raised from 2.5 nM to 5 μM. Figure 4b demonstrates that the G269 sequence from the G269 XP becomes cleaner as the concentration of HV1-EB increases. At a concentration of 500 nM the BLK $T_m$ is 4.6°C higher than the XP $T_m$ and the G269 sequence is clean. These results are consistent with those of Figure 3 and allow us to conclude that a $T_m$ difference of ~5°C between the BLK and the XP is sufficient for full blocking, except as described below.

Secondary structure effect. In yet another LATE-PCR assay we used the TSD & Globin amplicons (see ‘Materials and Methods’ section) to test the efficacy of blockers in a duplex reaction, Figure 5. The $T_m$ of TSD XP is 67°C while the $T_m$ of Globin XP is still ~61°C. BLKs, TSD-BL and Globin-BL, were designed with $T_m$s about 10°C higher than their corresponding XPs. The sequencing results for the two XP strands in this duplex are shown in Figure 5a. As expected both sequences are present in the absence of any BLK. In contrast, TSD XP generated a clean sequence in the presence of the Globin-BL. But, Globin XP did not generate a clean sequence in the presence of the TSD-BL. Further analysis revealed that the TSD-BL forms a hairpin at 50°C, Figure 5b, which renders it ineffective as a blocker. We therefore modified the blocker sequence by introducing a mismatch which destabilized the stem. Blocker TSD_BLm, Figure 5c, does not form a hairpin at 50°C and efficiently blocks the TSD excess primer as shown in Figure 5a.
Demonstration of the utility of blockers in a triplex LATE-PCR assay

We are interested in detecting mutations that accumulate in mitochondrial genomes due to aging, diabetes and environmental damage. Because these mutations are rare it is important to sequence both strands of each amplicon to verify that changes in one strand are not due to sequencing errors (14). In order to pursue this inquiry we have built a triplex LATE-PCR assay for three genes in the human mitochondrial genome HV1, HV2 and Cyt b (see ‘Materials and methods’ section). The amplicons generated in this reaction are 425–541 base pairs long and Figure 6 shows the results of sequencing all six strands using the Dilute-'N'-Go protocol with appropriate BLKs. Three aliquots of the reaction were used to sequence the LP strands by adding back the corresponding LPs without any BLK. Three additional aliquots were used to sequence the XP strands by adding two BLKs that prevented extension of the remaining two XPs in the reaction. But, sequencing the whole HV2 amplicon failed because it contains a string of Gs (or Cs) followed by a string of Ts (or As). Only the sequence before the Gs (or Cs) were obtained. Extensive trials with Genewiz, Inc established that this was a sequencing problem independent of the Dilute-'N'-Go protocol. To circumvent this problem we designed an inner HV2 sequencing primer, HV2_LPs2, which hybridizes to the HV2 XP strand downstream of the string of Ts in its template. The sequence of the whole HV2 amplicon can be obtained by combining the first part driven by HV2 LP (or XP) and the second part driven by HV2 LPS2 (or XPS2). Clean sequences for all strands in the triplex LATE-PCR reaction are obtained as shown in Figure 6.
DISCUSSION

This study demonstrates that Dilute-'N'-Go sequencing with added blockers or LPs is a convenient and accurate method for rapidly preparing and sequencing all of the oligonucleotide strands generated in a LATE-PCR multiplex reaction. In contrast, all other protocols for sample preparation prior to dideoxy sequencing are time consuming and expensive (6).

The concept of preventing primer extension by addition of a competitive oligonucleotide has been used before. For instance, Gillespie turned symmetric PCR into asymmetric by stopping the reaction after 10–50 thermal cycles and adding a blocker to one of the primers (18). Nordström et al. (19) carried out pyrosequencing from a double-stranded DNA template by adding two non-extendable oligonucleotides to prevent hybridization of both the forward and backward primers and also added an inner sequencing primer. In addition, allele specific PCR assays have been described in which a primer that is perfectly matched to the mutant but mismatch to the wild type is enhanced by addition of a blocker that is perfectly matched to the wild type strand but mismatched to the mutant strand, thereby making amplification of wild type less efficient than that of the mutant (20,21).

In contrast to these previous reports the study reported here investigated the design criteria of blockers needed to achieve efficient dideoxy sequencing of all strands in a multiplex LATE-PCR assay. Our results clearly show that efficient blocking of each XP depends on three independent factors. First, the position of the blocker relative to the primer is important because the 3’-end of the XP must be prevented from hybridization to its target to prevent elongation by the DNA polymerase. At least a single nucleotide overlap is required, but may not be sufficient even if the BLK T_m is much higher than the XP T_m. Because hybridization is dynamic, reliable blocking requires an overlap of at least three nucleotides. Second, the T_m difference between the BLK and the XP is also critical. Blocking is efficient when the T_m of the BLK is at least 5°C higher than that of the T_m of the XP. Finally, a blocker having a secondary structure with a T_m higher than the annealing temperature, typically 50°C for dideoxy sequencing, should be avoided.

We are interested in detecting changes in mitochondrial genomic sequences that result from aging or drugs. But, detecting such changes is difficult for two reasons. First, they are rare and second, the large number of mitochondrial genomes in a single cell obscures sequence changes in a fraction of these genomes. The solution is to amplify and then sequence portions of individual mitochondrial genomes. But, multiplex single molecule amplification is not all that is required. The sample must also be handled free of contamination and sequence changes that are detected in an amplified strand of DNA need to be verified since they could also be the results of errors during amplification or during sequencing. We have recently reported that LATE-PCR is a reliable method of amplification starting from single molecules and that Dilute-'N'-Go of the two complementary strands of DNA generated in a monoplex reaction can be used to validate rare sequence changes (14). Different polymerases, Platinum Taq and Tfi(–), have been used for LATE-PCR amplification with no difference in the sequencing results. The HV1 and HV2 genes of the human mitochondrial genome are part of the D-loop and are known to be variable from person to person, while the Cyt b gene is conserved within a species. Preliminary data indicate that treatment of human liver carcinoma cells with AZT results in accumulation of mutations in HV2 (Osborne and Wangh, unpublished results). The present report paves the way for extending that study to the use of multiplexed LATE-PCR assays. This approach will be far more cost effective than amplifying one target at a time.

FUNDING

A research grant from Smith Detection, Inc. (to L.J.W.). Funding for open access charge: Smiths Detection, Inc.

Conflict of interest statement. None declared.

REFERENCES

1. Kling, J. (2003) Ultrafast DNA sequencing. Nat. Biotechnol., 21, 1425–1427.
2. Braslavsky, I., Hebert, B., Kartalov, E. and Quake, S.R. (2003) Sequence information can be obtained from single DNA molecules. Proc. Natl Acad. Sci. USA, 100, 3960–3964.
3. Shendure, J., Mitra, R.D., Varma, C. and Church, G.M. (2004) Advanced sequencing technologies: methods and goals. Nat. Rev. Genet., 5, 335–344.
4. Pihlak, A., Bauren, G., Hersoug, E., Lonnerberg, P., Metsis, A. and Linnarsson, S. (2008) Rapid genome sequencing with short universal tiling probes. Nat. Biotechnol., 26, 676–684.
5. Guo, J., Xu, N., Li, Z., Zhang, S., Wu, J., Kim, D.H., Marma, M.S., Meng, Q., Cao, H., Li, X. et al. (2008) Four-color DNA sequencing with 3’-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. Proc. Natl Acad. Sci. USA, 105, 9145–9150.
6. Roe, B.A., Crabtree, J.S. and Khan, A.S. (1996) DNA Isolation and Sequencing, chapter IV. John Wiley & Sons, Hoboken, NJ, pp. 74–117.
7. Li, H., Cui, X. and Arnheim, N. (1991) Eliminating primers from completed polymerase chain reactions with exonuclease VII. Nucleic Acids Res., 19, 3139–3141.
8. Hanke, M. and Wink, M. (1994) Direct DNA sequencing of PCR-amplified vector inserts following enzymatic degradation of primer and dNTPs. Biotechniques, 17, 855–860.
9. Nordström, T., Nourizad, K., Ronaghi, M. and Nyérő, P. (2000) Method enabling pyrosequencing on double-stranded DNA. Anal. Biochem., 282, 186–193.
10. Sanchez, J.A., Pierce, K.E., Rice, J.E. and Wangh, L.J. (2004) Linear-After-The-Exponential (LATE)-PCR: an advanced method of asymmetric PCR and its uses in quantitative real-time analysis. Proc. Natl Acad. Sci. USA, 101, 1933–1938.
11. Pierce, K.E., Sanchez, J.A., Rice, J.E. and Wangh, L.J. (2005) Linear-After-The-Exponential (LATE)-PCR: primer design criteria for high yields of specific single-stranded DNA and improved real-time detection. Proc. Natl Acad. Sci. USA, 102, 8609–8614.
12. Rice, J.E., Sanchez, J.A., Pierce, K.E., Reis, A.H., Osborne, A. and Wangh, L.J. (2007) Monoplex/multiplex linear-after-the-exponential-PCR assays combined with PrimeSafe and Dilute-'N'-Go sequencing. Nat. Protoc., 2, 2429–2438.
13. Salk, J.J., Sanchez, J.A., Pierce, K.E., Rice, J.E., Soares, K.C. and Wangh, L.J. (2006) Direct amplification of single-stranded DNA for pyrosequencing using linear-after-the-exponential (LATE)-PCR. Anal. Biochem., 353, 124–132.
14. Osborne, A., Reis, A.H., Bach, L. and Wangh, L.J. (2009)
Single-molecule LATE-PCR analysis of human mitochondrial
Genomic sequence variations. PLoS ONE, 4, e5636.
15. Santalucia, J. (1998) A unified view of polymer, dumbbell, and
Oligonucleotide DNA nearest-neighbor thermodynamics.
Proc. Natl Acad. Sci. USA, 95, 1460–1465.
16. Le Novere, N. (2001) Melting, computing the melting temperature
Of nucleic acid duplex. Bioinformatics, 17, 1226–1227.
17. SantaLucia, J. (2007) Physical principles and visual-OMP software
For optimal PCR design. Methods Mol. Biol., 402, 3–33.
18. Gillespie, D. (1997) Competitor primer asymmetric polymerase
Chain reaction. Patent No. 5627054.
19. Nordström, T., Alderborn, A. and Nyrén, P. (2002) Method for
One-step preparation of double-stranded DNA template applicable
For use with Pyrosequencing™ technology. J. Biochem. Biophys.
Methods, 52, 71–82.
20. McKinzie, P.B. and Parsons, B.L. (2002) Detection of rare K-ras
Condon 12 mutations using allele-specific competitive blocker
PCR. Mutat. Res., 517, 209–220.
21. Morlan, J., Baker, J. and Sinicropi, D. (2009) Mutation detection by
Real-time PCR: a simple, robust and highly selective method.
PLoS ONE, 4, e4584.