BioTechniques

ISSN: 0736-6205 (Print) 1940-9818 (Online) Journal homepage: www.tandfonline.com/journals/ibtn20

Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol

Adam R Platt, Robert W. Woodhall & Alfred L. George Jr

To cite this article: Adam R Platt, Robert W. Woodhall & Alfred L. George Jr (2007) Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol, BioTechniques, 43:1, 58-62, DOI: 10.2144/000112499

To link to this article: https://doi.org/10.2144/000112499

Improved DNA sequencing quality and efficiency using an optimized fast cycle sequencing protocol

Adam R. Platt, Robert W. Woodhall, and Alfred L. George, Jr.
Vanderbilt University, Nashville, TN, USA

BioTechniques 43:58-62 (July 2007)
doi 10.2144/000112499

Technological advances in DNA sequencing over the past 20 years have given rise to the science of genomics and have enabled critical advances in other fields, including epidemiology, forensics, evolutionary biology, and medical diagnostics. Further advances in this technology will ensure the proliferation of new applications, such as comparative sequencing, and stimulate a paradigm shift in biology (1). Unfortunately, several factors, primarily the high cost of sequencing reactions, have slowed progress (2). In large genome sequencing centers, reduction of reaction volume along with standardization of template type and quality have contributed to lowered costs (3). Emphasis is now on developing novel sequencing methods to enable genome-wide analysis (4), and these efforts promise to further reduce the cost of large-scale sequencing projects. However, most DNA sequencing projects are of modest scale (5) and still use the dideoxy chain termination method described by Sanger et al. (6). Therefore, improvements to conventional methods that increase efficiency and reduce costs will have immediate benefits to a large number of researchers, especially when implemented in the setting of centralized facilities in which template heterogeneity and variability provide challenges to cost-cutting measures that are effective in genome sequencing centers.

Applied Biosystems (Foster City, CA, USA) recently introduced the BigDye® Fast cycle sequencing protocol to be used with their new 9800 Fast Thermal Cyclers. This protocol aims to increase throughput by expediting thermal cycling, the most time-consuming reaction step. Typical cycle sequencing protocols require 2–4 h, whereas the BigDye Fast cycle sequencing protocol requires only 50 min. The commercial protocol requires 4 µL undiluted BigDye Terminator v3.1 Ready Reaction mix to be used per reaction (docs.appliedbiosystems.com/pebiodocs/00114162.pdf). The commercial fluorescent terminator mix is the most expensive reagent used in sequencing; therefore the heavy consumption recommended by Applied Biosystems when using the BigDye Fast protocol potentially eliminates the savings gained by increased throughput. We investigated using the BigDye Fast protocol with the terminator mix reduced to more economical levels and found it performed well on short PCR fragments, but could not reliably sequence long PCR fragments (>500 bases) or plasmids. Here, we describe an improved cycle sequencing protocol that enables robust and reliable sequencing of a variety of DNA templates using an 8-fold reduction in fluorescent terminator mix.

To optimize our fast cycle sequencing protocol, we systematically evaluated the following parameters: (*i*) reaction volume, (*ii*) DNA concentration, (*iii*) dilution of fluorescent terminator mix, (*iv*) cycling temperatures, (*v*) and durations. The components of the various sequencing reactions we tested are provided in Table 1. Cycle sequencing was performed using an ABI 9800 Fast Thermal Cycler. Reaction products were purified using the Performa® DTR V3 column system (Edge

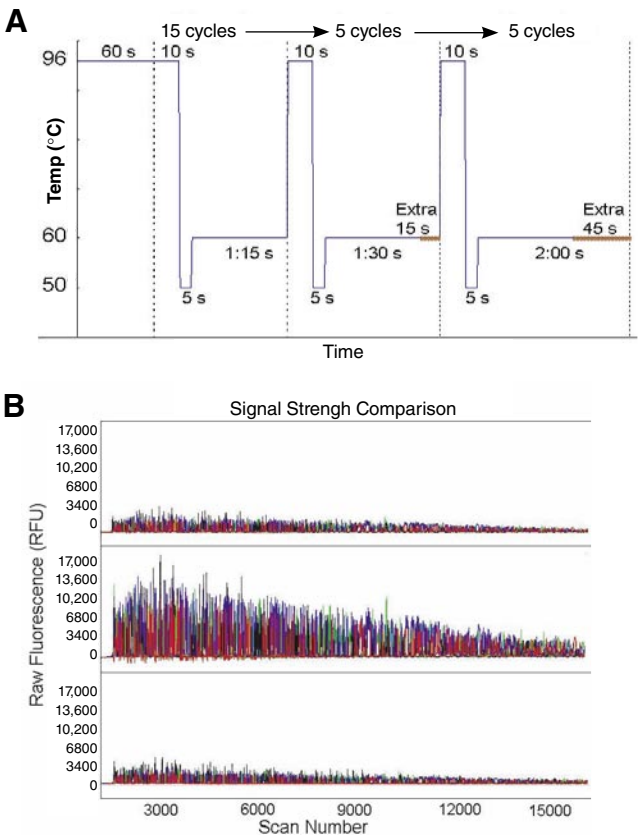

Figure 1. The design and performance of the stepped elongation time (ST_eP) cycle sequencing protocol. (A) The ST_eP cycling protocol. Applied Biosystems' BigDye Fast protocol consists of 25 cycles of the section highlighted in blue. ST_eP consists of 25 cycles, but elongation time (T_e) is extended by 15 s after cycle 15 and extended by an additional 30 s after cycle 20. (B) Plots of raw fluorescent signal for pGEM 3Zf(+) plasmid primed by M13F(-40) using three different cycling protocols. Upper panel, BigDye Fast; middle panel, ST_eP; and lower panel, standard protocol with a 3.5-h cycling time.

Table 1. Comparison of DNA Sequencing Reaction Components

Reaction: Type	BigDye Fast	Standard	ST_eP
BigDye Terminator v3.1	4.0 µL	0.67 µL	0.5 µL
5× Sequencing Buffer	0.0 µL	7.33 µL	2.0 µL
Primer	1.0 µL (18 ng)	4.0 µL (40 ng)	2.0 µL (20 ng)
DNA	1.0 µL (200 ng)	4.0 µL (100 ng)	2.0 µL (200 ng)
Deionized Water	4.0 µL	4.0 µL	3.5 µL
Total Volume	10 µL	20 µL	10 µL
Cycling Time	50 min	210 min	55 min

ST_eP, stepped elongation time (T_e) protocol.

Biosystems, Gaithersburg, MD, USA). Electrophoresis was performed using an ABI 3730*xl* DNA analyzer and POP-7™ polymer.

We found that adjustment of the elongation time (T_e) had the most profound effect on data quality. This modification is based on the observation that cycle sequencing creates short Sanger fragments in excess. These short fragments consume more dideoxynucleotide triphosphates (ddNTPs) in relation to deoxynucleotide triphosphates (dNTPs) than long fragments. We hypothesized that the majority of fragments produced during the initial cycles are small, causing ddNTP concentration to decrease more rapidly than dNTP concentration, progressively dropping the ddNTPP:dNTP ratio. With a lower ddNTP:dNTP ratio, the polymerase is less likely to ligate a terminating nucleotide and this increases the probability of creating longer fragments. Our protocol allows more time for elongation at the end of cycling when the likelihood of producing long fragments is highest.

The BigDye Fast protocol consists of an initial 60 s incubation at 96°C; followed by 25 cycles of 96°C for 10 s; 50°C for 5 s; and 60°C for 75 s. Our cycling protocol (Figure 1A) is identical to the BigDye Fast protocol through the first 15 cycles, but we increased the duration of T_e in the last 10 cycles. Specifically, T_e was extended by 15 s for cycles 16–20 and further extended an additional 30 s for cycles 21–25. We evaluated other modifications to the T_e, including lengthening the interval uniformly for all 25 cycles, as well as increasing T_e in a single step for the last 5 or 10 cycles. No changes were made to the denaturation or annealing temperatures or times. We found that the two-step T_e modification produced the best combination of signal-to-noise ratio and runtime. Our protocol, which we call stepped T_e protocol (ST_eP), has a runtime of 55 min, an increase of only 5 min over BigDye Fast, but produces much stronger signals and requires less fluorescent terminator mix. ST_eP produced a 4-fold increase in raw fluorescent signal strength over the BigDye Fast protocol using identical reaction conditions (Figure 1B). We found that this increased signal strength directly translates into a more robust and reliable reaction capable of sequencing small PCR fragments and large plasmids without adjustments.

The performance of ST_eP was tested in a sequencing core facility that processes an array of DNA samples. A diverse set of plasmid DNA samples originating in 26 different laboratories were sequenced using ST_eP and then compared with a similar set previously sequenced using a standard protocol. The standard protocol is more time-consuming (3.5 h) and consumes more fluorescent terminator mix than ST_eP (Table 1). All sequences were base called using phred, and each base call was assigned an error probability (7,8). Sequence sets were compared by the number of base calls with phred-quality values (Q) >20 (Q20; error probability <1%). The set sequenced with the standard cycling time averaged 850 ± 60.8 (mean ± SD; n = 870 samples) Q20 base calls per read. The ST_eP-cycled set averaged a significantly greater 871 ± 64.8 (n = 586) Q20 bases per read ($P < 0.0001$). Our sequencing facility recently implemented ST_eP as its standard sequencing protocol. To date, sequencing reactions utilizing ST_eP had a significantly higher success rate (96.8%; n = 11,289) as compared with the standard cycle sequencing protocol (92.9%; n = 11,270; $P < 0.0001$). Sequencing reactions performed using ST_eP utilized either standard (i.e., M13F, M13R, T7, T3, SP6; 27% of reactions) or user-designed custom (73%) oligonucleotide primers, and the majority of templates were either plasmids (5–10 kb; 58% of samples) or PCR products (average length 640 bp; 42%) provided by 367 different users. We also sequenced 34 bacterial artificial chromosome (BAC) templates yielding 822 ± 66 Q20 bases per read.

In conclusion, we have developed an optimized fast cycling protocol (ST_eP) for DNA sequencing. Sequencing results generated by ST_eP exhibited stronger signals than the commercial protocol, with only a 5 min increase in runtime. When compared with the more time-consuming standard protocol previously used by our sequencing facility, ST_eP had a significantly lower failure rate and produced significantly more Q20 base calls. The ST_eP cycle sequencing protocol is applicable for universal use (i.e., short and long DNA templates) and can be implemented as a means to reduce costs, increase efficiency, and increase sequence data quality.

ACKNOWLEDGMENTS

The authors would like to acknowledge Peggy Zahn and Deana Blevins for their work on initial testing. We thank Robert Taylor and Clay Miller for critical review of this manuscript. This work was supported in part by the Vanderbilt Diabetes Research and Training Center funded by a grant from the National Institutes of Health (no. DK20593).

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. **Chain, P., S. Kurtz, E. Ohlebusch, and T. Slezak.** 2003. An applications-focused review

of comparative genomics tools: Capabilities, limitations and future challenges. Brief. Bioinformatics *2*:105-123.
2. **Chan, E.Y.** 2005. Advances in sequencing technology. Mutat. Res. *573*:13-40.
3. **Smailus, D.E., A. Marziali, P. Dextras, M.A. Marra, and R.A. Holts.** 2005. Simple, robust methods for high-throughput nanoliter-scale DNA sequencing. Genome Res. *15*:1447-1450.
4. **Metzker, M.L.** 2005. Emerging technologies in DNA sequencing. Genome Res. *15*:1767-1776.
5. **Witonsky, J.** 2007. BioMarket trends: DNA sequencing. Genet. Eng. News *27*:1.
6. **Sanger, F., S. Nicklen, and A.R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *12*:5463-5467.
7. **Ewing, B., L. Hillier, M.C. Wendl, and P. Green.** 1998. Base-calling of automated sequencer traces using phred I. Accuracy assessment. Genome Res. *8*:175-185.
8. **Ewing, B. and P. Green.** 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. *8*:186-194.

Received 12 April 2007; accepted 29 May 2007.

Address correspondence to Alfred L. George, Jr. M.D., Vanderbilt University Division of Genetic Medicine, 529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232, USA. e-mail: al.george@vanderbilt.edu