

High-throughput sequencing of high copy number plasmids from bacterial cultures by heat lysis

Tapan Ganguly, Peiqin Chen, Rebecca Teetsel, Lan Ping Zhang, Elias Papaioannou, and Joseph Cianciarulo

University of Pennsylvania School of Medicine, Philadelphia, PA, USA

BioTechniques 39:304-308 (September 2005)

High-throughput sequencing as applied to shotgun plasmid libraries, cDNA, or other libraries requires a rapid and cost-effective method for screening large number of bacterial colonies. This is necessary to check for either the presence of the insert, the correct orientation of inserts, or in many cases, to identify clones prior to any downstream analysis. Continuous improvements in DNA sequencing technologies over the years has given rise to enhanced sequence quality, higher throughput, and a substantial reduction in cost. However, in large-scale sequencing projects, the preparation of templates involving cellular growth from bacterial colonies and the subsequent plasmid purification continues to be one of the most costly steps in terms of reagents and labor. This is particularly true for the laboratories in academic institutions. Previous attempts to sequence directly from bacterial colony (1–3) or culture (4) have yielded poor quality data and short sequence reads. The latest advancements in sequencing instrumentation, chemistry, and data analysis tools, available from Applied Biosystems, have significantly increased the quality of direct sequencing as demonstrated by a recent effort that used a proprietary buffer for cell lysis (Poster 113199; Applied Biosystems, www.appliedbiosystems.com). This prompted us to initiate direct sequencing of lysed bacterial cells harboring high copy number plasmids. We have developed a simple, high-throughput, and inexpensive method; the sequence quality and efficiency obtained by this method far surpasses the quality reported in similar efforts made earlier (1–4).

Our main objective has been to reduce or eliminate the reagent cost

involved in plasmid DNA preparation. Initially, we did a pilot study with a small number of plasmid colonies of PCR products in pCR[®]4-TOPO[®] or pCR-TOPO-XL high copy number vectors (Invitrogen, Carlsbad, CA, USA). The sizes of the inserts were between 1 and 10 kb. Cells from 10 colonies picked from different agar plates were grown in appropriate antibiotics in 3 mL 2 YT broth (Invitrogen) for approximately 18 h. The cells from 250 μ L of each culture were spun down, suspended in 125 μ L water, and lysed at 95°C for 5 min. The lysates were collected by centrifugation. The sequencing reactions were assembled in a total volume of 10 μ L with 4 μ L lysate, 2 pmols primer in 2 μ L water, and v3.1 BigDye[®] kit (Applied Biosystems) using the manufacturer's 1/8 \times reaction protocol. The primers were located either in the vector or in the inserts. Sequencing reactions were performed on an ABI GeneAmp[®] 9700 thermal cycler (Applied Biosystems) according to the manufacturer's standard protocol of 25 cycles, followed by purification using Performa[®] gel filtration short plates (Edgebio, Gaithersburg, MD, USA). The sequencing runs were performed using POP7 polymer on 36-cm capillaries in an ABI 3730 DNA sequencer (Applied Biosystems) and were analyzed by ABI Sequencing Analysis software v 5.1 and KB Basecaller (both from Applied Biosystems) capable of sequence quality value (QV) prediction on each base call.

The sequence quality obtained from these lysates was quite encouraging. Poor quality or short read length generally correlated with poor growth of bacterial cells. Few electropherograms had off-scale localized spikes over peaks

of regular height between 40 and 80 bases. These could be removed when the cell pellets obtained after the first spin were washed with water to remove any trace of growth (2 YT) medium.

Steps were then introduced to scale-up the procedure in 96-well plate format. A culture plate in a deep-well round-bottomed block was prepared from a glycerol stock plate, previously made from a library of colonies of PCR products in high copy number vectors. After growing cells in 250 μ L 2 YT medium for 18 h, the A_{600} of a 20-fold dilution of the cultures was checked randomly and was optimally between 0.2 and 0.3. The cells were pelleted by spinning for 5 min at 1000 \times g on a Marathon tabletop centrifuge (Fisher, Pittsburgh, PA, USA) equipped with plate holders, followed by two, 1-min washes with 125 μ L water/well to remove the growth medium completely. In order to avoid the aspiration steps as well as the cost of micropipet tips, the supernatants were removed each time by gently inverting the plate onto a tray and blotting on successive layers of paper towels making sure that the loosened cell pellets stayed attached to the bottom of the wells. The cells were then suspended in 125 μ L water by vortex mixing the sealed plate gently on a plate vortex unit. Care was taken so that there was no cross-contamination of cells from neighboring wells. The cell suspensions, after transfer into a PCR plate, were heated for 5 min at 95°C on the GeneAmp 9700 thermal cycler, followed by centrifugation for 10 min at 1000 \times g. Lysate (100 μ L)

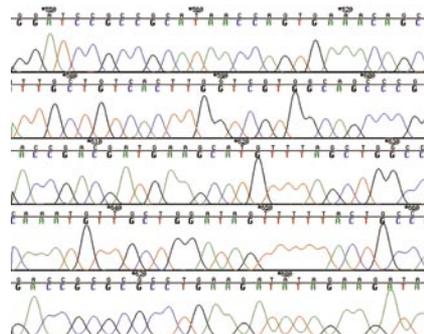


Figure 1. Partial electropherogram of the sequence of a PCR product cloned in pCR 2.1-TOPO. The template was prepared by heat lysis of cells in water, and the sequencing was done using BigDye v3.1 on a 48-capillary ABI 3730 DNA sequencer.

from each well was transferred into a fresh plate. Experience has shown that concentrating the lysates at this point to approximately 50% volume on a Savant SpeedVac® (Thermo Savant, Holbrook, NY, USA) ensures better sequencing result for the clones that grew poorly overnight. The procedure, starting from spinning down cells grown overnight to final concentration of the cell lysates, takes only about 2 h for a 96-well plate. The sequencing was done following the protocol mentioned earlier. A representative electropherogram of the partial sequence of a PCR product cloned in pCR 2.1-TOPO (Invitrogen) has been shown in Figure 1.

For comparison, a parallel study was done with the DNA templates made from 96 colonies obtained from the same glycerol stock plate using a SprintPrep™ kit (Agencourt, Beverly, MA, USA). The number of successful templates having at least 600 bases with QV 20 (corresponding to 1% probability of error in base call) was similar in both methods: 85% by heat lysis method versus 87% by SprintPrep method (data not shown). Attempts were made to avoid the overnight growth of cells, by growing 10 high copy number plasmids from the same glycerol stock plate for only 3 and 4 h in duplicate followed by heat lysis in

water and sequencing. The successful templates (approximately 40%) had between 300 and 400 bases with QV 20 (data not shown), while 30% of templates generated between 20 and 100 bases. The rest failed almost completely. The sequence quality did not depend significantly on the time of growth of the cells.

The percentage of templates with at least 250–300 QV 20 bases was only 50% when the heat lysis method was tried with 10 low copy number pET plasmids grown overnight (data not shown) in duplicate. The same low copy plasmids, however, generated good sequencing results (80% of templates with more than 600 QV 20 bases) when a very small portion from each colony was directly amplified, also in duplicate, for 4 h using the rolling circle amplification kit (TempliPhi™; Amersham Biosciences, Piscataway, NJ, USA) and processed for sequencing.

The heat lysis method described here will be particularly beneficial to the sequencing core facilities as well as research laboratories in academic institutions. The protocol is being routinely used at the DNA Sequencing facility at the University of Pennsylvania for large-scale sequencing of high copy number plasmids from different libraries. Figure 2 summarizes the quality of single-pass sequence data obtained from the first set of ten 96-well plates that were prepared and sequenced using this method on a routine basis. The average percentage of templates (averaged over 10 plates) that generated at least 600 QV 20 bases is 75 ± 13.0 , and the corresponding number (85%) obtained from the previous plate used for comparison falls within this range. We are currently trying to find ways to reduce the inter-plate variation in quality. The cellular lysates can also be utilized for PCR using primers located either in the vector or in the inserts.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. Gladden, J.G., R. Ray, R.D. Gietz, and M.R. Mowat. 2000. Rapid screening of plasmid DNA by direct sequencing from bacterial colonies. *BioTechniques* 29:436-437.
2. Kilger, C., M. Krings, H. Poinar, and S. Paabo. 1997. "Colony sequencing": direct sequencing of plasmid DNA from bacterial colonies. *BioTechniques* 22:412-418.
3. Krishnan, B.R., R.W. Blakesley, and D.E. Berg. 1991. Linear amplification DNA sequencing directly from single phage plaques and bacterial colonies. *Nucleic Acids Res.* 19:1153.
4. Chen, Q., C. Neville, A. MacKenzie, and R.G. Korneluk. 1996. Automated DNA sequencing requiring no DNA template purification. *BioTechniques* 21:453-457.

Received 8 April 2005; accepted 3 June 2005.

Address correspondence to Tapan Ganguly, DNA Sequencing Facility, University of Pennsylvania School of Medicine, B1 Richards, 3700 Hamilton Walk, Philadelphia, PA 19104, USA. e-mail: gangulyt@mail.med.upenn.edu

To purchase reprints
of this article, contact
apfeffer@BioTechniques.com

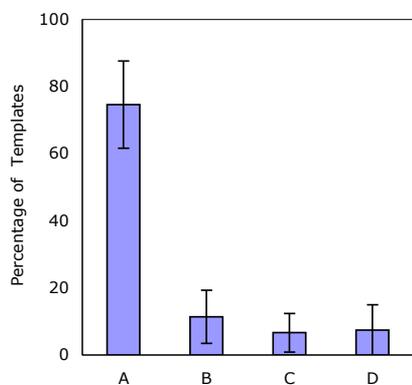


Figure 2. Quality of high-throughput sequencing obtained from ten 96-well plates. The percentage of templates from each plate that generated (A) at least 600 quality value (QV) 20 bases; (B) between 400 and 600 QV 20 bases; (C) below 400 QV 20 bases; and (D) very few usable bases was averaged over 10 plates. $QV = -10\log_{10}(Pe)$, where Pe is the probability of error. QV is calculated by KB Basecaller in the ABI Sequencing Analysis software v 5.1. QV 20 corresponds to 1% probability of error in base call.

ACKNOWLEDGMENTS

We thank Dr. Frederick Bushman for providing a majority of the clones.