High-Throughput Plasmid Purification for Capillary Sequencing

Christopher J. Elkin, Paul M. Richardson, H. Matthew Fourcade, Nancy M. Hammon, Martin J. Pollard, Paul F. Predki, Tijana Glavina, and Trevor L. Hawkins[1]

Department of Energy Joint Genome Institute, Walnut Creek, California 94598, USA

The need for expeditious and inexpensive methods for high-throughput DNA sequencing has been highlighted by the accelerated pace of genome DNA sequencing over the past year. At the Joint Genome Institute, the throughput in terms of high-quality bases per day has increased over 20-fold during the past 18 mo, reaching an average of 18.3 million `Phred 20` bases per day. To support this unprecedented scaleup, we developed an inexpensive automated method for the isolation and purification of double-stranded plasmid DNA clones for sequencing that is tailored to meet the more stringent needs of the newer capillary electrophoresis DNA sequencing machines. The protocol is based on the magnetic bead method of solid phase reversible immobilization that has been automated by using a CRS-based robotic system. The method described here has enabled us to meet our increases in production while reducing labor and materials costs significantly.

In April 1999, the Production Sequencing Facility of the Joint Genome Institute (JGI) replaced 28 ABI 377 (Applied Biosystems) slab gel systems with 84 MegaBACE 1000 DNA capillary sequencers (Amersham Pharmacia Biotech) in the realization that a massive scaleup in DNA sequencing throughput was needed to achieve the new goals of the Human Genome Project. For our part, we were charged with the draft sequencing of human chromosomes 5, 16, and 19, encompassing ~300 Mb or 10% of the human genome, which may contain up to 4000 genes. Our sequencing strategy was somewhat different than that of the other members of the G5 (Baylor College of Medicine, Washington University School of Medicine, Whitehead Institute, Sanger Center, and JGI). We were totally committed to using paired-end plasmid sequences, attributable in part to the documented advantages of using plasmids as opposed to M13-based sequencing vectors (Roach et al. 1995; Chissoe et al. 1997). In addition, the introduction of capillary-based DNA sequencing machines, coupled with the use of polyacrylamide as our sieving matrix, raised the throughput and uniformity requirements of our plasmid preparation. These facts required us to rethink our current DNA purification strategy.

In developing a scalable method for double-stranded plasmid template isolation, we faced the challenges of developing a high-throughput protocol that could be largely automated while maintaining the highest standards in terms of read length and pass rates. Internally, we set the bar at a 70% final pass rate with an average read length of more than 500 `Phred 20` bases (Ewing and Green 1998) on lanes producing 50 or more bases. In this high-throughput environment, we could not test each and every template before DNA sequencing. Therefore, the pass rate and DNA concentration had to be robust enough to allow for limited quality control testing. In practice, the method worked so well that we abandoned our gel and fluorometric analysis to gauge pass rate and average DNA concentration.

The protocol is a combination of cell lysis and cell debris removal followed by a final polishing step to remove proteins and lysis buffer components. A key component is the use of solid phase reversible immobilization (SPRI) as our polishing step. It presented itself as a scalable and automatable method that purifies plasmids to the levels required for DNA sequencing (Hawkins et al. 1994, 1997) and PCR products (DeAngelis 1995).

Our main focus was to develop a high-throughput lysis method that uniformly cracked open the bacteria while removing substances that degrade plasmids and interfere with the SPRI process. We adopted a detergent-heat lysis method that was developed at Washington University School of Medicine (Marra et al. 1999) and modified it to meet our needs. The major optimizations were the elimination of 19 steps by using direct lysis of the cells in the bacterial media, growth in shallow well plates, and the elimination of agarose gel quality control. These changes resulted in an extremely robust process that meets the DNA quality and yield requirements for DNA sequencing with linear polyacrylamide capillary systems.

RESULTS

We began by analyzing the current cell lysis methods detailed in the literature, including alkaline lysis (Brinboim and Doly 1979), boiling lysis (Holmes and Quigley 1981), detergent-microwave lysis (Marra et al. 1999), detergent-lysozyme-based lysis (Clewell and Helinski 1970), solvent-based lysis (Raha et al. 1990), physical grinding (Zhang and Ishaque 1997), and SDS-proteinase-K (Gross-Bellard et al. 1973). The most popular method, alkaline lysis, was ruled out because of requirements for high-speed centrifugation, multiple chemical additions, accurate timing of certain steps, and multiple liquid handling issues, as pointed out by Marra et al. (1999). Several other methods, such as boiling lysis and protenase-K, were eliminated as a result of unreliable plasmid yields and purity. Toxic chemicals, lack of automatability, and variability in plasmid yield and quality eliminated the rest of the methods except for the detergent-lysozyme-based lysis.

The advantages of detergent-lysozyme–based lysis included low centrifugation speeds, nontoxic solvents, high

[1]Corresponding author.
E-MAIL tlhawkins@lbl.gov; FAX (925) 296-5710.
Article and publication are at www.genome.org/cgi/doi/10.1101/gr.167801.

DNA yields, greater tolerances in step timings, low cost, and simplicity. In the initial development stages, we also noted that the high DNA yields allowed a move from deep-well growth blocks to standard 250-µL or shallow well plates. This would result in significant cost savings of bacterial media. Normally, the evaporation of media from around the edges of shallow plates during bacterial growth at 37°C prohibits this optimization. However, the use of Hi-Gro incubators (GeneMachines) reduced the evaporation through the injection of humidified oxygen.

During the development of this protocol, we found that our initial method produced a wide range of plasmid yields ranging from 0 to 100 ng/µL. It has been well characterized that the optimum DNA concentration range for capillary electrophoresis with linear polyacrylamide is 8–25 ng/µL (data not shown). We considered diluting each well on the basis of fluorometric readings but found that this limited throughput was unreliable and caused other failures resulting from plate handling. Therefore, we began looking for points in the method that cause nonuniformity or plasmid dropouts.

Our original lysis method allowed the plates to remain at room temperature for up to 2 h during lysozyme digestion followed by an incubation step in a 95°C water bath. These steps were replaced with convection ovens that remain at 37°C and 95°C for a specific time. These temperatures and times were optimized for both yield and speed. We also tested individual wells in each plate with thermocouple probes to determine the optimum placement and number of plates that could be incubated in each oven. During these experiments we noticed that plates coming out of the 95°C oven had seals that were peeling off. Apparently, the samples would boil over sometimes, causing the tween-20 to destroy the integrity of the seal. Supernatant was lost around the edges and during the subsequent centrifugation step. We had previously seen a lower `Phred Q20` read length on the edges of our plates (Fig. 1) and this seemed to be one source of the variation. We replaced the seal with a high-strength aluminum seal and installed a custom plate sealer that we affectionately named "the stomper." The next observation was that the bacteria would not always resuspend properly after the growth media was removed by centrifugation. The bacteria would simply wrap up into a ball or long strings on the addition of water.

A recent article (Song et al. 1999) detailed the direct addition of phenol chloroform to bacterial cultures to induce lysis. We believe that our lysis was still not uniform enough because the bacterial surface area for lysis was varying on addition of water before STET-E1 addition. We modified our lysis approach by adding an optimized detergent lysis solution directly to the bacterial cultures. The bacteria were still dissolved well in the Terrific Broth, and therefore lysis should be more uniform. Concentrations of ethylenediaminetetraacetic acid (EDTA), detergent, and enzymes were increased and salt concentrations were adjusted to levels that optimized plasmid yield and uniformity. The direct lysis also eliminated 19 steps, including a centrifugation step, resulting in a 50% reduction in labor. Just the elimination of these steps was thought to have increased uniformity, along with a significant decrease in both stress and injuries on our production line. Plate failures from low plasmid yields fell from 12% to <2% (data not shown), allowing for the elimination of the agarose gel quality-control step. This resulted in even higher throughput and uniformity.

The need for stop points in the process led to experiments that tested the effects of freezing at certain steps. Freezing the lysed cells at −80°C, followed by direct centrifugation of the frozen samples, gave equivalent results compared with freshly prepared cells (data not shown). An added bonus of freezing was that the resulting lysed pellets were firmly seated at the bottom of the wells, allowing for shorter centrifugation times.

The use of SPRI to further purify the crude supernatant was shown to increase read length by over 130 bases on average and improve pass rate by ~20%. Our average read length in production, based on `Phred 20` scores and lanes producing 50 or more bases, has been ~550 bp with maximum read lengths exceeding 800 bp (Fig. 2). As a demonstration of this increased plasmid quality, 10 (96 well plates) human genomic plasmid clones from shotgun libraries were chosen at random and duplicated. To one duplicate, the full lysis/SPRI protocol was applied as described, and the clones were sequenced with –40M13 Forward (Integrated DNA Technologies). To the other duplicate, only the lysis protocol was applied, after which the samples were ethanol precipitated before the same DNA sequencing was performed. The results are shown in Table 1.

The final optimization of the protocol was the implementation of an automation platform for SPRI purification. Our current automation systems from CRS (CRS Robotics), as shown in Figure 3, have made significant improvements in our reproducibility, yields, and labor requirements for this part of the process. The system was chosen because we had previous experience with the platform and it was specifically designed for the SPRI process.

The CRS system is a mechanical arm that moves plates from different stations and magnetic platforms. The stations are smaller robotic systems like the Robbins Hydra Cavro dispenser and the Bio-Tek Plate Washer. In this way, the platform is very similar to the manual hand method and allows us to easily troubleshoot and optimize the protocol offline. Currently the CRS systems are capable of producing 11 plates of purified

397	275	263	165	265	251	289	132
131	284	358	298	327	410	274	298
241	343	297	457	371	428	359	258
259	403	380	361	345	363	229	332
253	267	323	365	327	346	300	333
299	317	326	391	438	383	335	287
200	276	360	400	450	403	361	272
278	398	379	382	439	392	382	278
321	288	286	397	384	386	262	319
329	250	333	410	489	314	368	273
212	323	360	358	361	369	266	291
217	215	250	249	247	286	296	257

Figure 1 Average `Phred Q20` read length of 225 (96-well plates) plotted according to the well position. The A1 location is indicated for orientation. Wells producing on average <301 bp are shaded green. A1 is a pUC18 control location for the sequencing chemistry step.

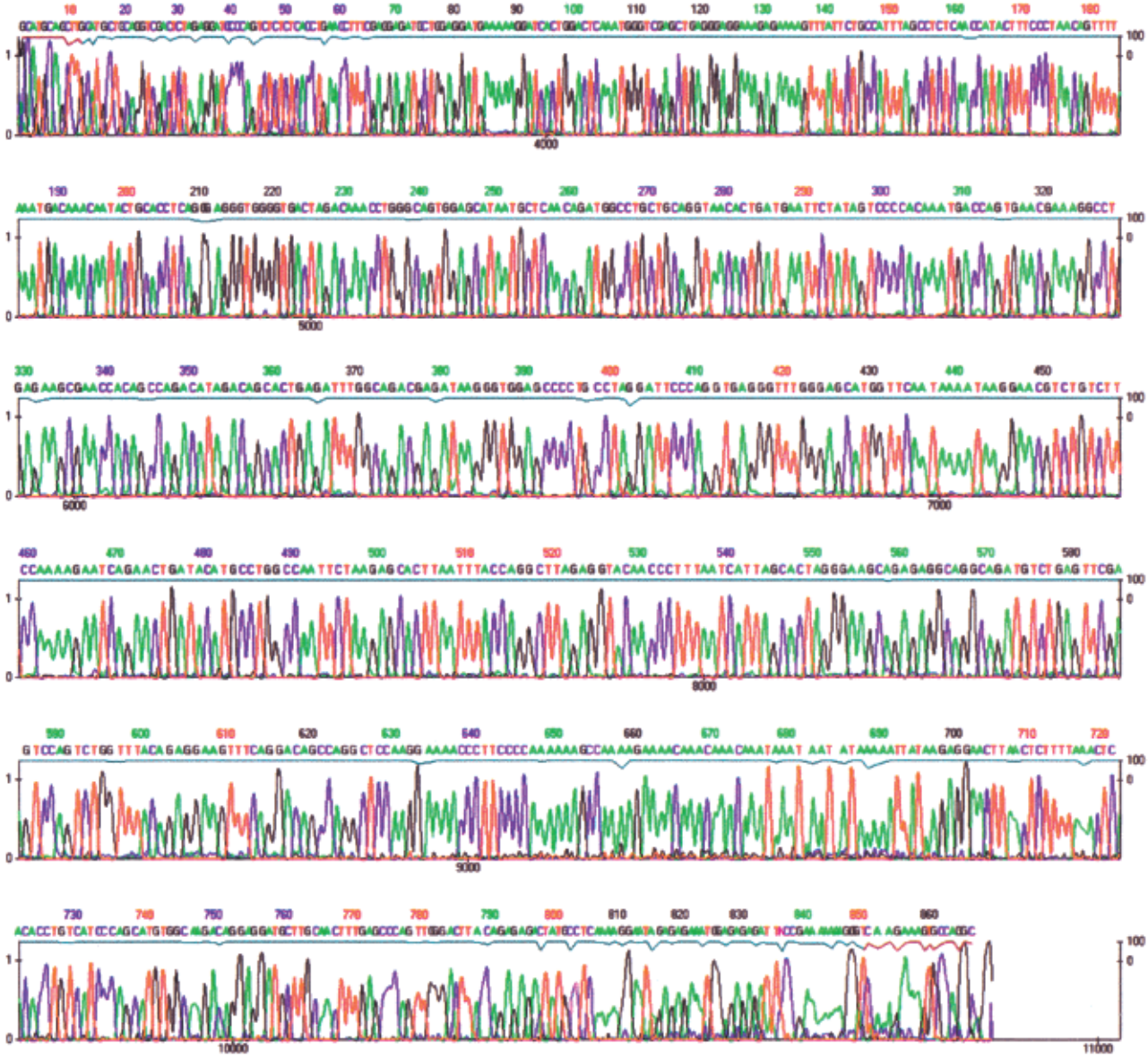

Figure 2 Electropherogram of a pUC-18 subclone with a random 3- to 4-kb human DNA insert purified with this protocol. Plasmid was sequenced with ET dye-terminator chemistry from Amersham Pharmacia Biotech on a Molecular Dynamics MegaBACE 1000 with standard conditions. Read length was 805 bp, scoring a `Phred Q20` or greater.

plasmid DNA per hour, per machine, with an off-the-shelf instrument.

DISCUSSION

This automation-based protocol provides a high-throughput, low-cost method for the isolation and purification of double-stranded plasmid template samples for use on capillary-based DNA sequencing devices. Even though this method works well manually, the development and use of automation throughout the protocol provides a critical advantage in ensuring precision and accuracy through each step of this large-scale process in which the tedium of manual repetition would become a significant variable. We have shown that plasmid purification by SPRI yields a very clean template sample suitable for capillary electrophoresis on both the MegaBACE DNA sequencers as described here and on the ABI 3700 systems as used in our microbial DNA sequencing facility. We have calculated that the reagent cost of this template isolation method is approximately $0.17 per sample, with sufficient DNA being isolated to generate multiple sequencing reactions. As the graph in Figure 4 shows, we have achieved significant increases in throughput at the JGI as well as improvements in pass rates and read lengths.

The introduction of this protocol has played a major part in the success of our sequencing facility. Looking ahead, we are in the process of adapting SPRI as a replacement for ethanol precipitation for sequencing chemistry clean-up and we are also designing 384-well versions of SPRI that can be used for both applications (manuscripts in preparation).

Table 1. Comparison of Crude Lysis Purification Using Either 70% Ethanol Precipitation or Solid Phase Reversible Immunobilization

	Pass rate (Phred 20 >50 bp)		Average read length (Phred 20 >50 bp)	
Plate Name	+SPRI	−SPRI	+SPRI	−SPRI
ker0005_fw	87.5%	67.7%	584	495
ker0006_fw	93.8%	76.0%	649	439
ker0007_fw	86.5%	78.1%	582	506
ker0008_fw	95.8%	47.9%	516	281
kfn0010_fw	95.8%	95.8%	614	536
kfn0012_fw	78.1%	61.5%	491	492
kfn0013_fw	94.8%	82.3%	614	518
kfn0014_fw	96.9%	70.8%	597	429
kfn0015_fw	92.7%	44.8%	537	347
kfn0016_fw	88.5%	65.6%	632	421
AVERAGE	**91.0%**	**69.1%**	**582**	**446**
Std Dev*	**5.9%**	**15.4%**	**51.6**	**81.2**

Plates are 96-well and contain pUC-18 plasmid with random 3–4 kb human DNA inserts from chromosomes 5 and 16. Pass rates are based on lanes producing >50 bases scoring Phred 20 or better. Read lengths are the average of the Phred 20-qualified bases in the passed lanes. Total samples, 960.
*Std Dev, Standard Deviation.

METHODS

Bacterial Inoculation

Plasmid clones (pUC-18 based) were picked by using a variety of automated pickers into Nalge/Nunc International 384 well plates that were previously filled with Terrific Broth + glycerol and then incubated at 37°C in Hi-Gro shakers (Genomic Instruments Services Incorporated). After the initial growth, 5 µL of bacterial inoculum was transferred from the 384-well plates into a new 96-well Falcon polystyrene plate (N3918–3071) containing 185 µL of Terrific Broth w/ 1.6% glycerol (Teknova) with antibiotic added. These inoculated plates were subsequently loaded into microtiter plate holders (GeneMachines) without lids, loaded into Hi-Gro incubators (GeneMachines), and incubated at 37°C for 18 h at an agitation of 450 rpm. The 384-well glycerol plates were then stored at −80°C.

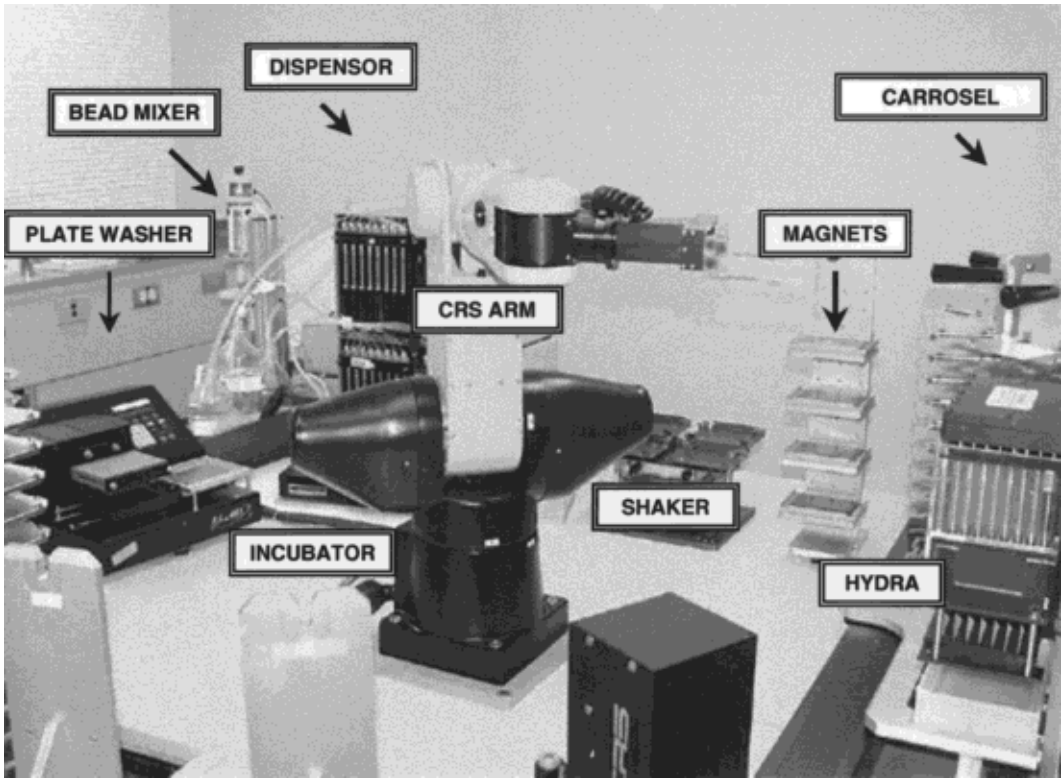

Figure 3 One of the automation systems for solid phase reversible immobilization purification of crude DNA lysates in use at the Joint Genome Institute. The current stand-alone system is based on a CRS platform and produces up to 11 (96-well plates) per hour.

96-Well Lysis

After incubation, the 96 well plates were moved to Multidrop units (Lab Systems Incorporated), which added 60 µL of STET-E1 solution (3.0 M NaCl, 165 mM Tris-HCl at pH 8.0, 45mM EDTA at pH 8.0, 16.5% Tween-20, 200 mg/mL RNase A, 1 mg/mL lysozyme) to each well of the plate. The plates were then sealed with foil seals (Beckman Coulter)

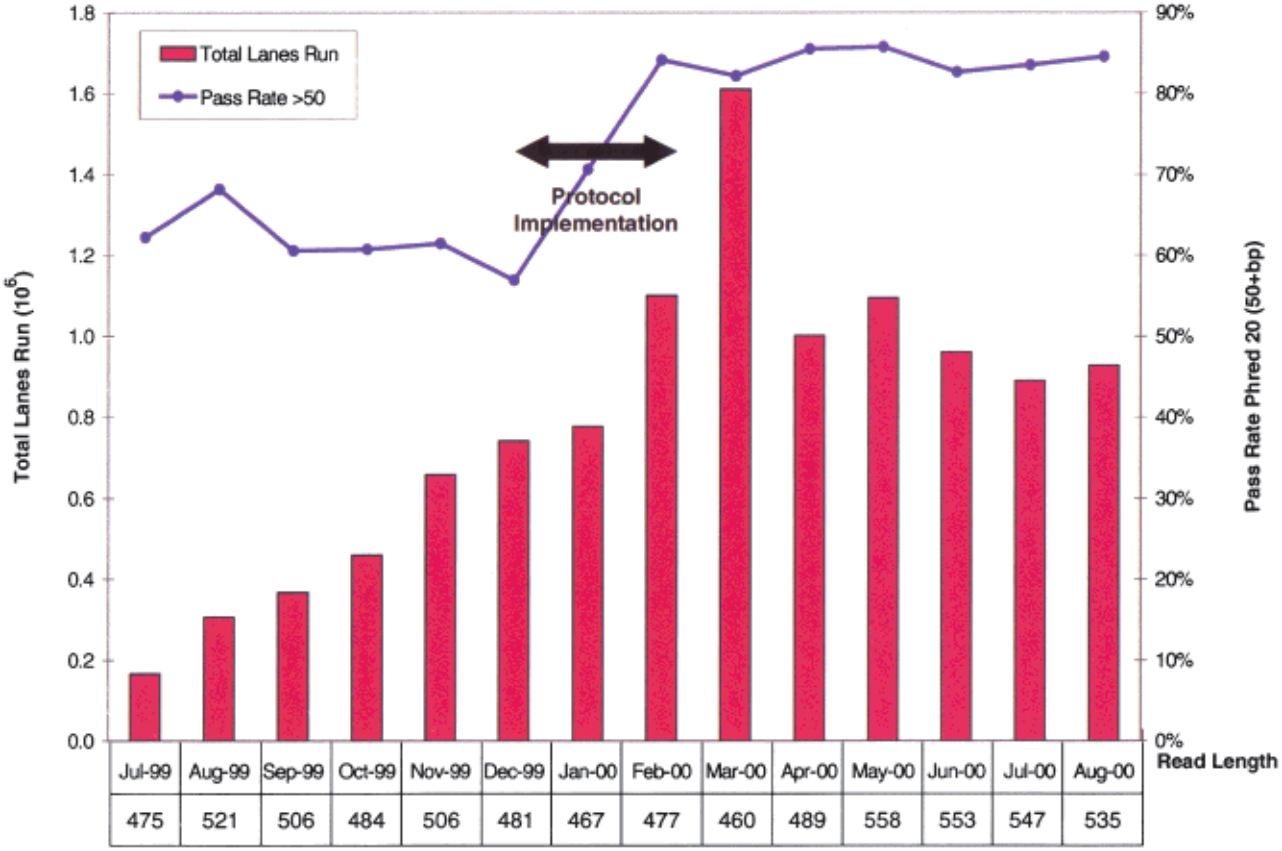

Figure 4 Total lanes run is the total of lanes loaded in the corresponding month multiplied by 1,000,000. Read length is the average read length for all lanes loaded that produced 50 or more bases with a `Phred` score of 20 or higher. Pass rate is the percentage of lanes loaded that produced 50 or more bases with a `Phred` score of 20 or higher.

and vortexed for 30 sec by using a Mini Orbital Shaker (Bellco Glass Inc.) at a low speed to ensure sufficient mixing of the lysis solution with the cellular suspension. After vortexing, the plates were incubated at 37°C for 15 min. The cell/STET-E1 mixtures are then incubated at 90°C in a horizontal flow oven (VWR Scientific, model 1350FM) for 10 min. At this point, plates were stored at −80°C for at least 30 min and often overnight. The cellular debris was then separated by centrifuging at 3200×*g*-force for 90 min in an Eppendorf 5810 centrifuge. (Two plates can be stacked together in each bucket of the centrifuge.) Plate samples should be frozen when transferred to the centrifuge for optimal debris removal and minimal handling. Centrifugation was then followed by the immediate transfer of 100 µL of crude supernatant to a 96-well polystyrene Falcon plate by using a Hydra-96 (Robbins' Scientific) 96-well pipettor. The DNA-containing supernatant can then be stored at −20°C, if needed, before proceeding with DNA purification.

Automated DNA Purification Protocol

The DNA must be further purified to prepare it for sequencing chemistry and subsequent capillary electrophoresis. DNA purification was accomplished by using SPRI. This used carboxylate-coated magnetite, 0.768 + 0.008 um-diameter Sera-Mag magnetic microparticles (Seradyn, part no. 44152105050350). The sodium azide suspension, in which the beads are preserved, was removed before purification by washing 5 mL of beads in 45 mL of 10mM Tris (diluted from 1M Tris-HCl at pH 8.0) in a 50-mL conical tube. The beads were separated from the Tris/bead suspension by using a MagneSil Magnetic Separation Unit (Promega). The beads were then washed in this fashion three times before resuspending the beads in a final 45 mL of Tris (1:10 dilution). Fifty-two millimeters of washed beads were then mixed with 448 mL of PEG/NaCl (2.5M NaCl, 20% PEG-8000; Teknova) to create a suspension for use with automated purification.

The remainder of purification was completed through the use of an automated platform centered on a CRS-articulated robot. The main components of the robot included an 8-tipped reagent dispenser, a 96-well plate washer (Elx405 Magna; Bio-Tek), a horizontal plate shaker (Vibrax; IKA), a 96-well Hydra (Robbins' Scientific), four microtiter plate hotels (each with a 15-plate capacity), and a carousel of 120 microtiter plate nests.

All plates that were processed through the 96-well lysis protocol were loaded onto the CRS DNA purification robotic platform into individual plate nests on the carousel. Each plate was first transferred from its initial position on the carousel to a sliding microtiter plate nest, below the eight-tipped reagent dispenser, where 145 µL of the beads/PEG solution was added to the already existing 100 µL of lysate in each well of the plate. The plate was then transferred to the horizontal plate shaker by the CRS arm and shaken for 5 min at 800 rpm to thoroughly mix the lysate, beads, and PEG together. The plate was then replaced to its initial nest on the carousel to incubate at room temperature for 20 min, allowing the DNA present in the lysate to hybridize to the magnetic microparticles. The CRS arm then transferred the microtiter plate to a

magnetic plate made up of 96 separate doughnut-shaped magnets on which the plate was incubated for 10 min at room temperature. The magnets draw the magnetic-bead-bound DNA to the bottom of the plate, where it is held in a circular conformation, providing a way to retain the bead-hybridized DNA within the microtiter plate while impurities are washed from the DNA. The DNA was washed six times with 200 µL of 70% ethanol by a 96-well plate washer on which was mounted a magnetic plate similar to that described earlier.

The washed plate, containing only bead-hybridized DNA and a small amount of residual ethanol, was again replaced to its initial nest and allowed to dry for 10 min at room temperature. The dry plate was transferred back to the sliding microtiter plate nest, where 50-µL Millipore filtered water was added to each well of the plate to dehybridize the bound plasmid DNA from the beads. To ensure that all bound plasmid DNA was exposed to water and subsequently disassociated from the beads, we transferred each plate to the horizontal plate shaker, where the plate was shaken for 5 min at 800 rpm. On resuspension in water, the plates were all replaced to their initial positions on the plate nest carousel by the CRS robot.

The final step in the automated process ensured that none of the beads were transferred into subsequent PCR or thermal cycling reactions. The plates containing resuspended DNA (donor plates) were removed from their positions on the plate carousels by the CRS arm and placed onto a magnetic plate. This magnetic plate had been mounted onto a computer-controlled sliding platform, which functioned in conjunction with the 96-well Hydra. This sliding platform had three plate nests used as positions for an empty destination plate, a plate containing resuspended DNA and magnetic beads, and a reservoir containing Millipore-filtered water. An appropriately labeled, empty destination plate was placed onto the sliding platform adjacent to the donor plate by the CRS arm. The platform then moved to orient the donor plate directly under the 96-well Hydra, which aspirated 40 µL of resuspended DNA uncontaminated by beads. The platform then oriented the destination plate under the Hydra and dispensed the resuspended DNA into the plate. The platform moved again to orient the water reservoir under the Hydra to wash the Hydra needles. The donor plate was then disposed of into a large plastic waste container and the destination plate was placed into the carousel on the nest from which the donor plate originated. Typical yields are 100 ng/µL in 40 µL.

ACKNOWLEDGMENTS

We acknowledge the help of the JGI's DNA sequencing group for their hard work in generating the 3.8 billion `Phred 20` bases by using this method. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Berkeley National Laboratory under contract No. DE-AC03–76SF00098, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48, and Los Alamos National Laboratory under contract No. W-7405-ENG-36.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Birnboim, H.C. and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7:** 1513–1523.

Chissoe, S.L., Marra, M.A., Hillier, L., Brinkman, R., Wilson, R.K., and Waterston, R.H. 1997. Representation of cloned genomic sequences in two sequencing vectors: Correlation of DNA sequence and subclone distribution. *Nucleic Acids Res.* **25:** 2960–2966.

Clewell, D.B. and Helinski, D.R. 1970. Properties of a supercoiled deoxyribonucleic acid-protein relaxation complex and strand specificity of the relaxation event. *Biochemistry* **9:** 4428–4440.

DeAngelis, M.M., Wang, D.G., and Hawkins, T.L. 1995. Solid-phase reversible immobilization for the isolation of PCR products. *Nucleic Acids Res.* **23:** 4742–4743.

Ewing, B. and Green, P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8:** 186–194.

Gross-Bellard, M., Oudet, P., and Chambon, P. 1973. Isolation of high-molecular-weight DNA from mammalian cells. *Eur. J. Biochem.* **36:** 32–38.

Hawkins, T.L., O'Connor-Morin, T., Roy, A., and Santillan, C. 1994. DNA purification and isolation using a solid-phase. *Nucleic Acids Res.* **22:** 4543–4544.

Hawkins, T.L., McKernan, K.J., Jacotot, L.B., MacKenzie, J.B., Richardson, P.M., and Lander, E.S. 1997. A magnetic attraction to high-throughput genomics. *Science* **276 :** 1887–1889.

Holmes, D.S. and Quigley, M. 1981. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114:** 193–197.

Marra, M.A., Kucaba, T.A., Hillier, L.W., and Waterston, R.H. 1999. High-throughput plasmid DNA purification for 3 cents per sample. *Nucleic Acids Res.* **27:** e37.

Raha, S., Merante, F., Proteau, G., and Reed, J.K. 1990. Simultaneous isolation of total cellular RNA and DNA from tissue culture cells using phenol and lithium chloride. *Genet. Anal. Tech. Appl.* **7:** 173–177.

Roach, JC, Boysen, C, Wang, K, and Hood, L. 1995. Pairwise end sequencing: A unified approach to genomic mapping and sequencing. *Genomics* **26:** 345–353.

Song, J., Geltinger, C., Sun, K., Kanazawa, I., and Yokoyama, K.K. 1999. Direct lysis method for the rapid preparation of plasmid DNA. *Anal. Biochem.* **271:** 89–91.

Zhang, Z.Q. and Ishaque, M. 1997. Evaluation of methods for isolation of DNA from slowly and rapidly growing mycobacteria. *Int. J. Lepr. Other Mycobact. Dis.* **65:** 469–476.

Received October 18, 2000; accepted in revised form March 22, 2001.

High-Throughput Plasmid Purification for Capillary Sequencing

Christopher J. Elkin, Paul M. Richardson, H. Matthew Fourcade, et al.

Genome Res. 2001 11: 1269-1274
Access the most recent version at doi:10.1101/gr.167801