

High-throughput plasmid DNA purification for 3 cents per sample

Marco A. Marra*, Tamara A. Kucaba, LaDeana W. Hillier and Robert H. Waterston

Washington University Genome Sequencing Center, 4444 Forest Park Boulevard, St Louis, MI 63108, USA

Received April 30, 1999; Revised July 30, 1999; Accepted August 15, 1999

ABSTRACT

To accommodate the increasingly rapid rates of DNA sequencing we have developed and implemented an inexpensive, expeditious method for the purification of double-stranded plasmid DNA clones. The robust nature, high throughput, low degree of technical difficulty and extremely low cost have made it the plasmid DNA preparation method of choice in both our expressed sequence tag (EST) and genome sequencing projects. Here we report the details of the method and describe its application in the generation of more than 700 000 ESTs at a rate exceeding 16 000 per week.

INTRODUCTION

While cost-effective and robust methods for purification of cloned DNA from bacteria impact the success of any DNA sequencing effort, they are crucial for high-throughput, large-scale projects where many thousands of DNA purifications and subsequent sequencing reactions are performed daily. These activities require methods, easily implemented, that are capable of the required throughput, have a low level of technical difficulty, yield DNA in reproducible amounts, and are applicable to different bacterial host/vector combinations. As part of ongoing efforts to improve the double-stranded plasmid DNA purification process at our Genome Sequencing Center (1), we explored different commercial and in-house methods, evaluating each with respect to the factors listed above. These investigations led to the development of the DNA purification method described here; first implemented in a production mode in October of 1997. Since that time, a team of four technicians has used it to prepare more than 700 000 DNA samples, with current throughput exceeding 16 000 preparations each week. The method, suitable for fueling high-throughput sequencing projects, is described here in detail.

MATERIALS AND METHODS

96-well microwave protocol

Two studies in the literature reported the isolation of plasmid DNA from bacterial cells by treatment with lysozyme followed by a brief exposure to microwave radiation. One study (2) used 1 ml aliquots of bacterial cultures grown to saturation in 50 ml culture volumes. Bacterial cell pellets, collected in individual microfuge tubes, were resuspended in a lysozyme-containing solution. Exposure to microwave radiation was performed, and the resulting DNA purified by isopropanol precipitation. The authors noted that DNA prepared by their protocol was of quality suitable for restriction digests and other unspecified enzymatic modifications, but the suitability of the DNA for sequencing was not reported. Another study (3) reported the use of microwave radiation in the isolation of plasmid DNA from large volumes (250 ml) of bacterial cell cultures. Again, the suitability of this DNA for sequencing was not discussed. Both the volumes and formats reported in these studies were unsuited to large-scale sequencing; hence, we sought to adapt microwave-mediated bacterial cell lysis to 96-well format, and assess the performance of the resulting DNAs in sequencing reactions.

Culturing plasmid clones and DNA purification

Culture conditions were similar to those described (4). Briefly, a 12 channel pipettor was used to transfer, from a 384-well plate of glycerol stocks, 5 µl to each well of a 96-well Beckman block (Beckman part no. 140504) containing 1 ml of Terrific Broth (Difco) with the appropriate antibiotic. Blocks were covered with polystyrene lids, placed in custom holders in a Labsystems floor shaker and incubated at 37°C for 24 h with agitation at 295 r.p.m. Following growth, blocks were removed, cell pellets collected by an 8 min centrifugation at 2700 r.p.m. (1400 *g*) in a Jouan GR-422 centrifuge, and excess culture media decanted. Draining of residual media was achieved by inverting the block over paper toweling with intermittent rapping, after which blocks were sealed with foil tape and stored at –80°C. To purify DNA from the bacterial cell pellets, blocks were first thawed by incubation on the laboratory bench or in a 37°C water bath. Sterile de-ionized water (25 µl) was added to each well with either BIOHIT Proline 50–1200 µl or Multidrop 831 (Labsystems) pipettors and bacterial cell pellets were resuspended by vigorous vortexing using a Multi-Tube Vortexer (VWR 58816–115). Resuspension of the bacterial cell pellet was confirmed by subjecting each block to a brief vortex using a Vortex-Genie 2 model G-560 (VWR) and inspecting the wells to insure they were free of 'clumps'. A lysis solution (70 µl), prepared fresh by mixing 200 ml STET-TWEEN20 [10 mM Tris–HCl pH 8.0, 1 mM EDTA pH 8.0, 100 mM NaCl, 5% Molecular Biology Grade TWEEN20

*To whom correspondence should be addressed. Tel: +1 314 286 1800; Fax: +1 314 286 1810; Email: mmarra@alu.wustl.edu
Present address:
Marco A. Marra, Genome Sequence Centre, British Columbia Cancer Research Centre, 601 West 10th Avenue, Vancouver, British Columbia V5Z 1L3, Canada

(Sigma; P9416); stored at room temperature for at least 3 months with no decrease in DNA sequence quality] with 48 mg (4.8 ml of a 10 mg/ml solution) RNAse A (Sigma; R6513) and 100 mg (2 ml of a 50 mg/ml solution) of lysozyme (Sigma; L7651), were then added to each well with a Multi-drop pipettor.

Immediately after addition of lysis solution mixing was achieved by subjecting the block to a brief vortex, with care taken to avoid the introduction of bubbles. Multiple blocks were handled serially; lysis solution was added, the block vortexed, and then the next block processed. This was followed by a 5 min incubation at room temperature, after which two blocks were placed side-by-side in a microwave oven (General Electric Model J-E693TWH 002) with a peak power output rating of 0.95 kW. The blocks were subjected to a 30 s irradiation at the maximum power setting and, to ensure uniform exposure of all wells to microwaves, removed from the microwave, rotated 180 degrees, replaced in the oven, and subjected to an additional 30 s of irradiation. The Multidrop pipettor was then used to add 350–500 µl sterile de-ionized water to each well. The volumes added varied depending upon the vector types and host cells of the clones being processed, with the optimal amount determined empirically by sequence analysis (see below). Blocks were subjected to a brief vortex to mix the well contents, sealed with foil tape, and placed in an ice water bath for up to 15 min. Bacterial cell debris was collected in the bottom of the wells by a 30 min centrifugation of the block in a Jouan GR-422 at 4000 r.p.m. (~3000 *g*), after which a Robbins Scientific 96-well Hydra pipettor was used to transfer 105 µl of DNA-containing supernatant from the Beckman blocks to non-tissue culture treated 96-well polystyrene Costar microtiter plates (VWR; 62408–189). This DNA-containing solution was used directly, without further purification, for DNA sequencing and restriction analysis.

DNA sequencing

DNA was sequenced (5) using either BigDye terminator chemistry (Perkin-Elmer/Applied Biosystems) or DYEnamic energy transfer dye primers (Amersham). Typically, 'master mixes' of sequencing reaction components were dispensed into Robbins cycleplates using a Robbins Hydra, which was also used subsequently to add the DNA. For BigDye terminator sequencing, a reaction cocktail consisting of 2.5 µl of terminator pre-mix, 1.0 µl (3 pM) oligonucleotide primer, and 3.5 µl of sterile deionized water was assembled. To this was added 1–2 µl of microwave lysate containing on average between ~130 and 260 ng of DNA. A MicroAmp Full Plate Cover (Perkin-Elmer; N801–0550) was placed onto the cycleplates and the reaction mixtures collected in the bottom of the wells by centrifugation. Cycle sequencing was performed in MJ Research PTC-200 or -225 thermocyclers using 35 cycles of 95°C for 15 s, 45°C for 5 s, 60°C for 2 min, followed by incubation at 4°C. After cycling, cycleplates were centrifuged to collect the reactions in the bottom of the wells and a mixture of 1 µl of 3 M sodium acetate, pH 5.2 and 100 µl of 100% ethanol (typically added as 101 µl of a mixture of the two components) was added to each well. After mixing by repeated pipetting with a Robbins Hydra, DNA precipitates were collected by centrifuging the cycleplate for 30 min at 4000 r.p.m. Ethanol was decanted and the reaction pellet washed with 200 µl 80% ethanol. After a second centrifugation for 15 min at 4000 r.p.m., ethanol was again decanted and the samples dried for 10 min in a SpeedVac. Dried samples were placed in light-tight containers and stored at –20°C until electrophoresis on ABI 373 or 377 sequencers.

For each dye primer sequencing reaction a reaction cocktail consisting of 0.4 µl Thermo Sequenase reaction buffer, 1.6 µl Thermo Sequenase Nucleix, 1 µl DYEnamic ET primer mix (at 0.1 µM for A and C reactions and 0.2 µM for G and T reactions) and 1 µl (1.5 U) Thermo Sequenase enzyme was prepared. Reaction cocktails were aliquoted into cycleplates and stored at –20°C until required. When needed, a set of four cycleplates, representing a set of 96 reactions, were thawed and centrifuged and a Robbins Hydra used to add to each well 2 µl of DNA. Cycleplates were again centrifuged, placed in thermocyclers and subjected to 15 cycles of 95°C for 4 s, 55°C for 10 s, 70°C for 1 min; followed by 15 cycles of 95°C for 4 s, 70°C for 1 min. Precipitation of completed sequencing reactions was performed with a Robbins Hydra using a pooling scheme which involved addition of 100 µl of 100% ethanol and 3 µl of 5 M ammonium acetate to all 'A' reactions, followed by serial transfer of the ethanol/salt/DNA mixture first to the cycleplate containing the 'C' reactions, then the 'G' reactions and finally the 'T' reactions. Pooled reactions in the 'T' tray were placed on ice for 15 min and then centrifuged at 4000 r.p.m. for 30 min. Following centrifugation, trays were inverted to decant ethanol and blotted onto paper towels to facilitate draining. Reaction pellets were washed by addition of 200 µl of 70% ethanol and then centrifuged for 15 min at 4000 r.p.m. to collect the reaction pellet. Ethanol was decanted, the tray rapped gently on paper towels to remove excess ethanol, and the reaction pellets dried in a SpeedVac with the rotor removed. Reactions were stored in light-tight containers until required. Immediately prior to electrophoresis, reactions were resuspended in 2 µl of formamide/blue dextran dye and denatured by heating at 95°C for 5 min. Reactions were subjected to electrophoresis on ABI 373 or 377 sequencers with 36 cm well-to-read glass plates. For the 373, sequence data were collected on 8.3 M urea, 4.75% SeaQuate gels (Sooner Scientific, SQL 4750) run at 3000 V, 25–30 W, 40 mA for 10 h. For the 377, sequence data were collected on 6 M urea, 5% Gene Page Plus (Amresco J722–1L-UWA; degassed before use). Both Seq Run 36A-2400 (3000 V, 60 mA, 200 W; 3 h electrophoresis) and Seq Run 36A-1200 (1750 V, 50 mA, 150 W; 8 h electrophoresis) run modules were used for dye primer sequencing reactions; for BigDye terminator reactions, both Seq Run 36E-2400 and Seq Run 36E-1200 run modules (conditions described above) were used. Sequencing gel images were processed using Gelminder (J.Mullikan, M.Holman and A.Chinwalla, unpublished) which incorporates the signal processor PLAN (B.Ewing and P.Green, unpublished).

DNA sequence analysis

We conducted qualitative evaluations of sequence electropherograms ('traces') generated from microwave-prepared DNA using the UNIX tool VTRACE (B.Ewing, unpublished), which facilitates manual inspection of sequence trace data. Initial sequencing results revealed traces which were 'top heavy'-sequence-read lengths were shorter than expected, with a fragment size distribution weighted towards short sequencing extension products. These trace profiles often indicate excessive

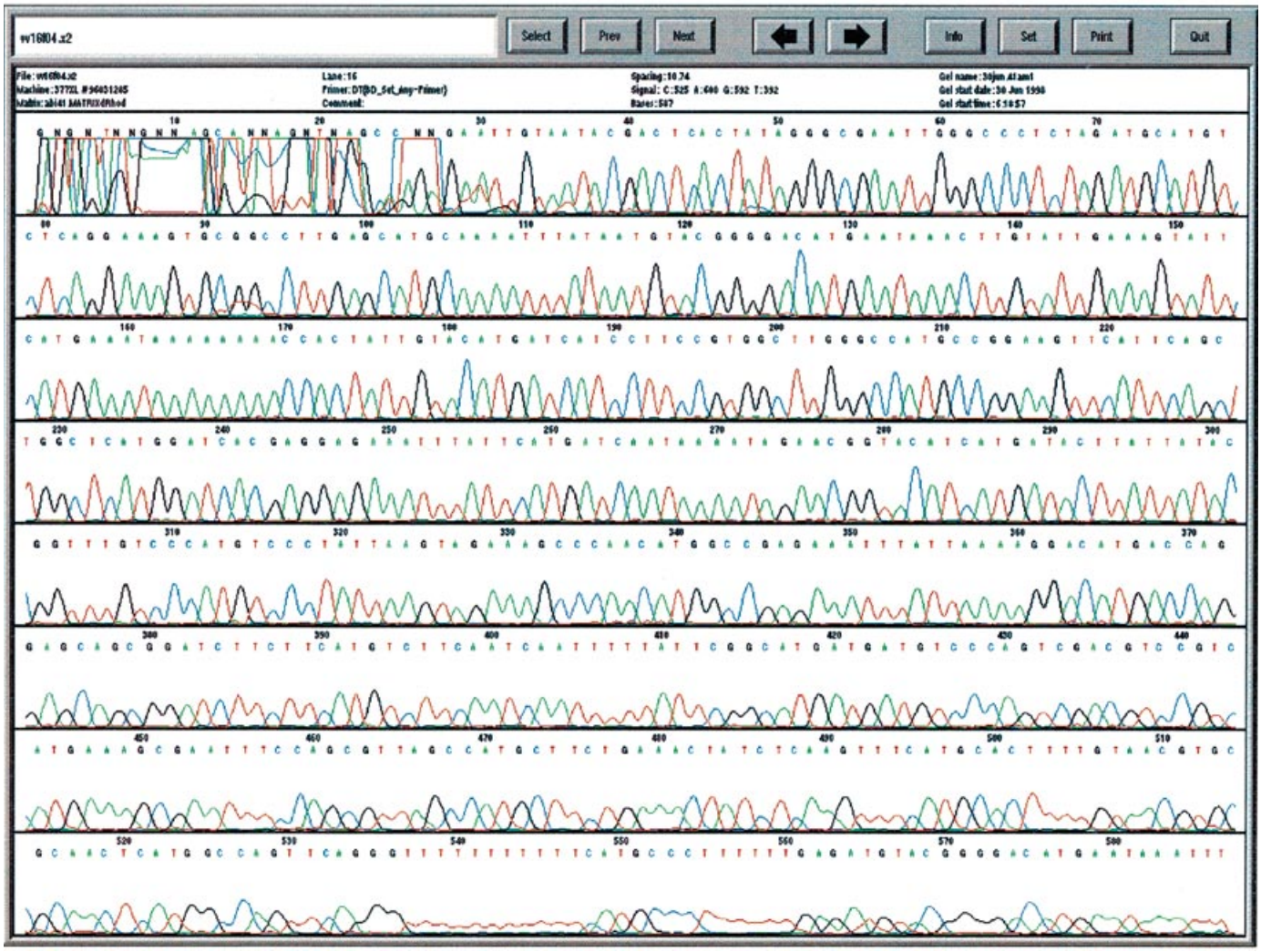

Figure 1. VTRACE output for a DNA sample prepared with the 96-well microwave method and sequenced as described with BigDye terminators. Electrophoresis was performed on an ABD 377 using 4XA run module parameters.

DNA in the sequencing reaction. To determine the optimal dilution of the microwave lysate a series of trial-and-error experiments were performed in which lysates were diluted with water and re-sequencing performed. Traces were then manually inspected using VTRACE and analyzed by EST_OTTO (described below). Optimal dilutions were plasmid- and host cell-dependent; for pT7T3-based plasmids (6) and pZERO plasmids (Invitrogen) propagated in DH10B cells (Life Technologies), best results were obtained by addition of 500 µl of water to the DNA preparation following microwave lysis. For pBluescript-based plasmids propagated in SOLR cells (Stratagene), equivalent sequencing results were obtained when 360 µl of water were added. After satisfactory and reproducible sequencing results had been obtained, we measured the concentration of pZERO microwave-prepared plasmids using PicoGreen stain (Molecular Probes) and a Fluorimager (Molecular Dynamics). The average concentration of 192 diluted samples, from two 96-well microtitre plates, was calculated to be 135.7 ng/µl with a standard deviation of 9.8 ng/µl (DNA concentrations ranged from 10.9 ng/µl to 448 ng/µl). A VTRACE image of a DNA sample prepared by the microwave method and sequenced with BigDye terminators is shown in Figure 1. Peaks are well resolved, and background signal does not confound the base-calling. For this sequence trace, the high quality length cut-off was 400 bases, the maximum allowed by the software. This cut-off is conservative; useful data extends well beyond, with resolvable peaks in the run of T residues near position 540. These results typified those achieved with DNA prepared using the microwave protocol and our standard sequence electrophoresis conditions for ESTs. The performance of the microwave lysates under specific conditions designed to achieve longer read lengths was not assessed.

Quantitative evaluations of sequence quality for large numbers of traces were produced by the computer program EST_OTTO. These reports included information on the fraction of high quality (or 'successful') sequences and the average high quality read lengths attained for a set of sequences. Both overall sequence trace quality and the position of the last high-quality base in a sequence trace can be determined. A measure of the noise-to-signal ratio (the ratio of the height of the highest non-called base to the height of the called base at that position)

Table 1. EST_OTTO analysis of ESTs generated from microwave-prepared DNA

Library (vector/host)[a]	Sequencing chemistry[b]	High quality sequences[c]	Attempts	Percent	High quality lengths[c]
Stratagene mouse skin (pBluescript/SOLR)	E-T dye primer	209	288	72	368
Soares parathyroid tumor (pT7T3-Pac/DB10B)	E-T dye primer	299	384	77	363
SAGE HCT116 (pZERO/DH10B)	BigDye terminator	330	384	86	367
Gessler Wilm's tumor (pSPORT1/DH10B)	E-T dye primer	34	48	71	342
		37	**48**	**77**	**367**

[a]Libraries are human cDNA or SAGE libraries except as indicated. The vector and host bacterial cell line are shown for each library.
[b]DNA templates were sequenced with energy transfer (E-T) dye primers or BigDye terminators. Gessler Wilm's tumor templates were used in the stability assay (text). Numbers in bold are sequencing results obtained in the test for DNA stability, after passage of the DNA preparations through five freeze–thaw cycles.
[c]See text for definition of 'high quality'.

was used to determine overall trace quality. At least ten 16-base windows (each window overlapping by 8 bases) of the trace were required to display fewer than six 'problem' regions, where a problem region was defined as a peak having a noise-to-signal ratio of >0.27. To set the high quality cut-off for trace data a 'sideband' measure was used. This was defined as the ratio of the height of a peak's shoulder at a position in a trace halfway between adjacent peaks to the maximum height of the peak. The side band measure was used starting from the 5′ end of the trace and moving towards the 3′ end. Traces were scanned with a window of 16 bases, moved in 8 base increments, until four consecutive 'problem' regions were found. Here, a problem region was defined as one in which a single base in a window of 16 bases had a sideband ratio of 0.95 or more. After this position was determined, a noise-to-signal measure was applied to trim the trace back towards the 5′ end. The position of the last high-quality base in the trace was then set at the point where the noise-to-signal ratio was <0.25 for every base in two consecutive windows. These values were determined empirically: the derivation of trace quality values that described high quality sequence data resulted from efforts to identify values which best approximated a dataset of manually selected cutoffs.

Examples of some of the output EST_OTTO reported for sequences generated from microwave lysates are shown in Table 1. From these data it is evident that the microwave method can be used to prepare sequence-quality DNA from cDNA libraries employing different vectors and propagated in different bacterial host strains. This flexibility is relevant to EST efforts, which sequence libraries obtained from diverse sources. The data in Table 1 also show that microwave lysates can be sequenced with different chemistries; we have used successfully BigDye terminator (Perkin-Elmer/Applied Biosystems) and Energy-Transfer dye primer (Amersham) technologies. Finally, Table 1 shows that the DNA prepared using this method is of sufficient stability for our purposes. Shown is an assay for stability in which 48 samples from a nephroblastoma library, constructed by M.Gessler, were subjected to five consecutive freeze–thaw cycles with temperature range from –20°C to room temperature. Subsequent resequencing resulted in no degradation of sequence quality. Week-long incubations at 4°C could be performed without affecting the integrity of the DNA (not shown) but DNA was found to degrade after a 4 day incubation at room temperature.

A long-term, larger-scale assessment of sequence quality has been performed with the tool GASP (7; Fig. 2). An evaluation of 40 329 sequences, generated by BigDye terminator sequencing of microwave lysates, was performed. Here, we have used a function provided by GASP to plot the fraction of sequences (ordinate) against the number of high quality bases per read (abscissa), where base quality has been determined using the program Phred (8,9). From this analysis, 7% of the sequences have 50 or fewer high quality bases (Phred values of at least 20). The average high quality sequence length, excluding those reads with 75 or fewer high quality bases, was 427 bases, while 70.52% of the sequences had at least 400 bases of high quality data.

Restriction digestion of microwave lysates

In addition to sequencing, another focus of several projects has been the estimation of cDNA insert size by restriction digestion of DNA samples followed by analysis on agarose gels (4). A restriction digestion mix for cDNA libraries, constructed as described (6), consisted of 7.75 µl deionized water, 1.0 µl Buffer B (Boehringer Mannheim) and 0.125 µl (5 U) each of *Hin*dIII and *Eco*RI (Boehringer Mannheim). cDNA libraries constructed using different cloning strategies required different choices of enzymes to liberate the cDNA insert from the vector. Typically, sufficient 'master mix' for multiple 96-well digests was assembled on ice. Nine microliters of the mix were then dispensed into each well of a Robbins Scientific cycleplate using a Robbins Scientific Hydra 96 channel pipetting device, which was also used to add ~200 ng of microwave-prepared DNA to each well. Cycleplates were sealed with foil tape, centrifuged briefly to collect the reaction components in the bottom of the wells, and incubated for 1 h at 37°C in a water bath. Two microliters of 6× loading dye (10) were added to each well and the cycleplate subjected to a second centrifugation. Ten microliters of this mixture were then loaded with a 12-channel pipettor (Oxford) into each well of a 96-well 2% agarose gel cast in 1× TAE (10). Electrophoresis in 1× TAE buffer was for ~3 h at 80 V, or until complete as judged by monitoring the migration of the dyes in the gel. DNA was visualized by incubating gels in 0.5 µg/ml ethidium bromide and then scanning the gel using a Molecular Dynamics Fluorimager (Model SI or Model 595). 16-bit images were collected at normal sensitivity with a pixel size of 200 microns and a PMT voltage of 850.

Shown in Figure 3 is an image, generated on a Molecular Dynamics Fluorimager, of an ethidium bromide stained 2%

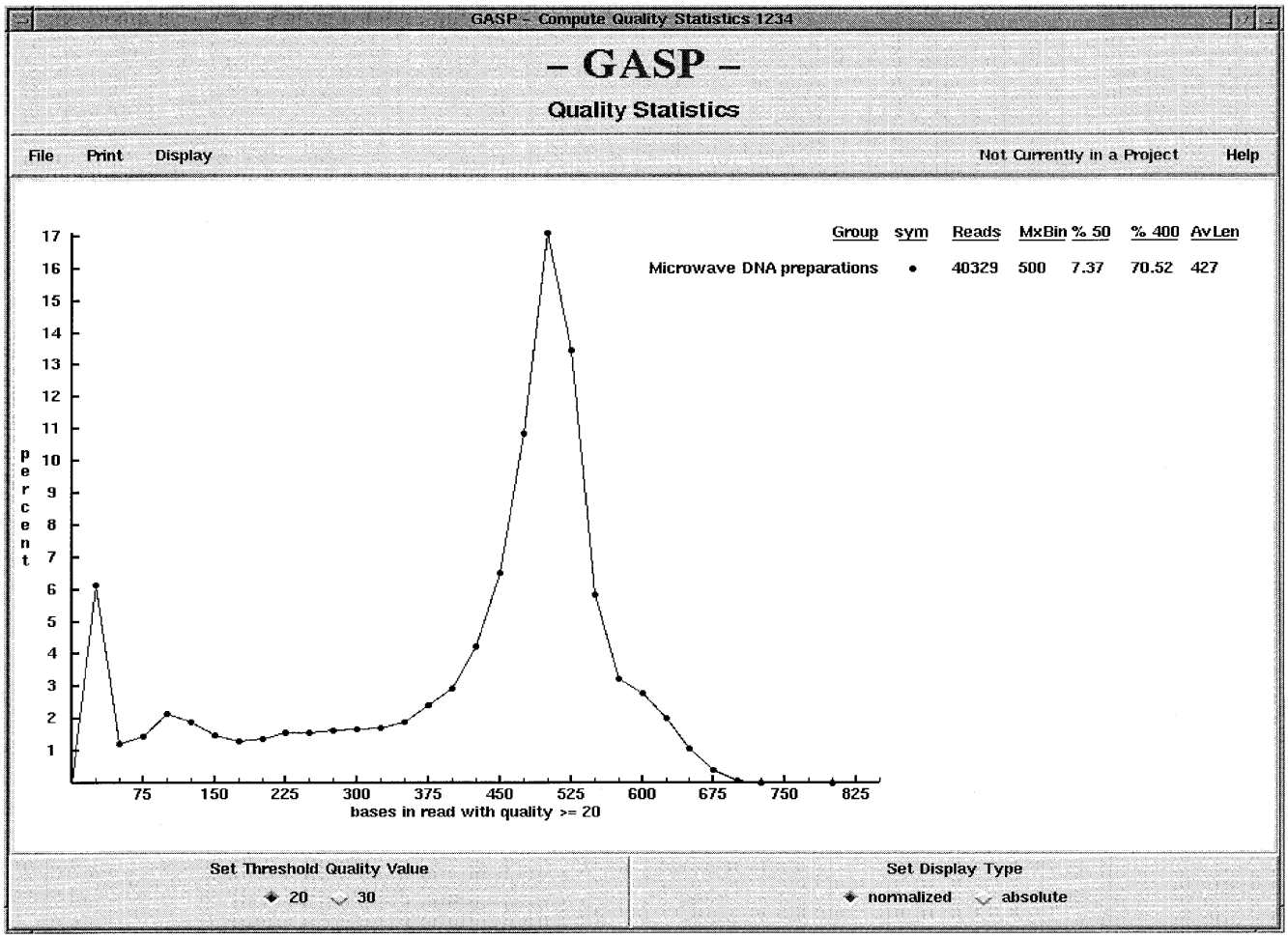

Figure 2. GASP quality analysis of EST sequences prepared using the 96-well microwave method. 40 329 sequences were analyzed with GASP using a Phred cutoff score of 20. A base having at least a score of 20 is classified here as a high quality base. MxBin, the *x*-axis coordinate where the distribution reaches a maximum. % 50, the percentage of reads having 50 or fewer high quality bases. % 400, the percentage of reads having 400 or more high quality bases. AvLen, the average length of the sequences analyzed, excluding sequences with 75 or fewer high quality bases.

agarose gel containing 96 restriction-digested microwave lysates. These results are of a quality comparable to that achieved with DNA prepared by other methods (not shown). We noted that digested lysates often exhibited evidence of high molecular weight material, perhaps bacterial genomic DNA, migrating above the vector-specific restriction fragment. The amount of this material correlated with the extent to which vortexing was conducted (after addition of water and prior to centrifugation as described above). More vigorous vortexing produced more of this material, while in cases where vortexing was brief and gentle this material was sometimes absent. Whatever its source, this material did not affect sequencing results or confound restriction analysis.

DISCUSSION

The microwave protocol described here has become the preferred method for purification of double-stranded DNA in our sequencing projects. The DNA is very easily produced, of adequate stability and generally performs very well in the sequencing and restriction analysis protocols used in our laboratory. The protocol, a favorite of our technical staff, is markedly less tedious than others due primarily to the reduced number of required manipulations. For example, growth of bacterial cultures and subsequent DNA isolation are performed in the same 96-well block, and no further purification of the DNA, by precipitation or other means, is necessary. The 96-well blocks can be re-used indefinitely provided they are cleaned between uses. Importantly for our applications, the DNA can be manipulated with the Robbins Hydra pipettor without fear of clogging or otherwise affecting the instrument. Further, the lysis solution is easily made and is stable at room temperature for a minimum of 3 months, allowing litre-quantity batches to be made and stored. And, unlike commercially available DNA preparation kits, knowledge of the reagent composition should facilitate troubleshooting when necessary. Finally, the preparation is remarkably inexpensive. We calculate that reagents cost 1.6 cents, with plasticware adding another 1.4 cents for a total estimated cost of 3 cents per sample. This compares very favorably with the cost of commercial preparation methods, which can exceed 1 US dollar per sample. We first implemented the method in October 1997 and, through March 1999, have used it to purify more than 700 000 EST templates. During this period, application of the microwave method described here

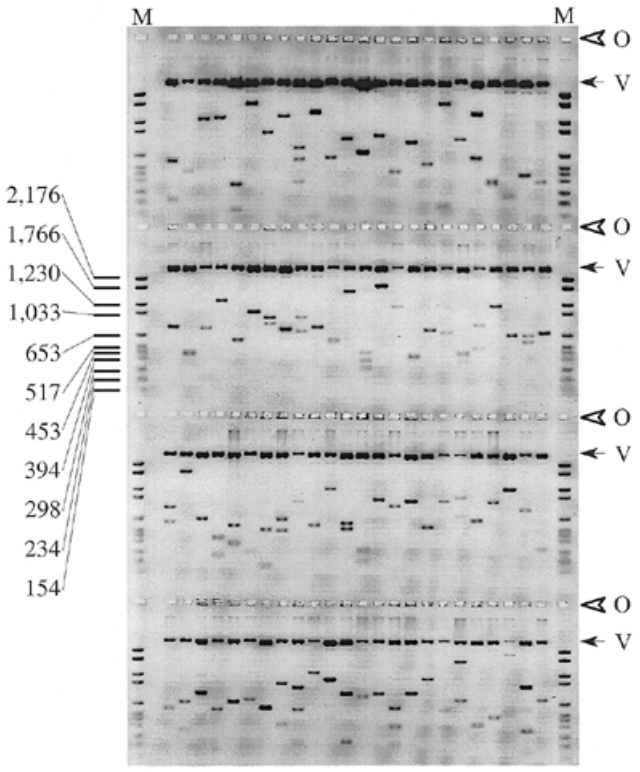

Figure 3. 2% agarose gel stained with ethidium bromide showing restriction digests of 96 microwave lysates. Samples were processed as described (Materials and Methods). M, marker DNAs (Boehringer Mannheim Marker VI). The sizes of the individual marker fragments are given. O, origins of the gel. V, vector-specific restriction fragments.

has resulted in savings of ~$700 000 over the cost of commercially obtained kit reagents—a sum which we have applied to the generation of additional ESTs.

ACKNOWLEDGEMENTS

We are indebted to Elaine Mardis, Eric Green, Sharon Gorski and Stephanie Chissoe for comments on the manuscript and to Elizabeth Boatright for assistance in quantification of microwave lysates. For excellent technical support in all aspects of EST generation at Washington University Genome Sequencing Center we are grateful to John Martin, Cathy Beck, Karen Underwood, Michele Steptoe, Brenda Theising, Todd Wylie, Missy Allen, Yvette Bowers, Barry Person, Tim Swaller, Marilyn Gibbons, Deana Pape, Njata Harvey, Rebecca Schurk, Erica Ritter, Sophie Kohn, Tanya Shin, Marco Cardenas, Tammy Yount, Angie Blistain, Yolanda Jackson, Rhonda McCann, Ann Chamberlain, Julie Chapell, Richard Morales, Louise Bowles and Steve Geisel, and all the staff at Washington University Genome Sequencing Center. Thanks also to Christa Prange (LLNL) and Greg Lennon (GeneLogic and LLNL) for making available cDNA clones for sequence analysis.

REFERENCES

1. Marra,M.A., Hillier,L. and Waterston,R.H. (1998) *Trends Genet.*, **14**, 4–7.
2. Hultner,M.L and Cleaver,J.E. (1994) *Biotechniques*, **16**, 990–994.
3. Wang,B., Merva,M., Williams,W.V. and Weiner,D.B. (1995) *Biotechniques*, **18**, 554–555.
4. Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W. *et al.* (1996) *Genome Res.*, **6**, 807–828.
5. Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.A. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
6. Bonaldo,M.F., Lennon,G. and Soares,M.B. (1996) *Genome Res.*, **6**, 791–806.
7. Wendl,M.C., Dear,S., Hodgson,D. and Hillier,L. (1998) *Genome Res.*, **9**, 975–984.
8. Ewing,B., Hillier,L., Wendl,M. and Green,P. (1998) *Genome Res.*, **8**, 175–185.
9. Ewing,B. and Green,P. (1998) *Genome Res.*, **8**, 186–194.
10. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.