A comparison of five methods for extracting DNA from paucicellular clinical samples

Leslie Cler, Dawei Bu, Cheryl Lewis, David Euhus *

Hamon Center for Therapeutic Oncology Research, UT Southwestern Medical Center at Dallas, 5323 Harry Hines Boulevard, Dallas, TX 75390-9155, USA

Received 23 February 2005; accepted for publication 20 December 2005

Abstract

Translational protocols in cancer and carcinogenesis often require isolation of genomic DNA from paucicellular clinical samples. DNA extraction methods for PCR-based applications should optimize the recovery of amplifiable DNA. We compared five methods for DNA extraction in paucicellular epithelial and lymphocyte samples using proportion of extractions producing amplifiable DNA and mean real-time PCR Ct values for GAPDH as the endpoint measures. The methods included solid-phase DNA adsorption (QIAamp), sequential protein and DNA precipitation (Puregene), magnetic bead adsorption (Dynabeads), phenol–chloroform extraction, and single-step proteinase K digestion. In general, the performance of the three commercial kits was superior to either phenol–chloroform extraction or single-step proteinase K digestion. However, QIAamp and Puregene produced amplifiable DNA more frequently than Dynabeads for starting cell numbers >50,000. GAPDH Ct values for QIAamp extractions showed the greatest dynamic range and the best linearity across the range of starting cell numbers, but QIAamp was not statistically significantly superior to Puregene. Of the three commercial kits, Puregene is the least expensive. QIAamp and Puregene DNA extraction methods are well-suited for the preparation of paucicellular clinical samples for PCR-based assays.

Keywords: DNA extraction; Paucicellular; Real-time PCR; Methods; Cost analysis

1. Introduction

Translational protocols in cancer and carcinogenesis often require isolation of genomic DNA from paucicellular clinical samples such as fine needle aspirates [1], nipple fluid aspirates [2,3], sputum [4], buccal swabs [5], or urine [6]. If high quality amplifiable DNA can be extracted from these samples, they are often suitable for LOH analysis [7], gene copy number determinations [8], genotyping [9], mutation analysis [10,11], or promoter methylation studies [12]. The classical approach to DNA extraction employs organic solvents to dissolve DNA after which it is precipitated in absolute alcohol. Though suitable for highly cellular samples, this approach requires multiple centrifugation steps and often results in poor yields of amplifiable DNA when the starting material is limited. Newer approaches for DNA extraction include single-step proteinase K digestion (without the use of organic solvents), adsorption of DNA on to silica gel membranes or magnetic beads, or simplified approaches for sequentially precipitating proteins and then DNA.

The DNA content of an extract is often estimated by measuring the absorbance of light at 260 nm (A260) and the purity of the DNA by calculating the A260/A280 ratio. A260 values between 0.1 and 1.0 correlate with DNA content in a linear fashion, but values this high are rarely achieved when the starting material is limited. Real-time PCR provides a semi-quantitative approach for estimating the content of amplifiable DNA in extracts from paucicellular samples. The Ct value determined by this method is the number of PCR cycles required to generate a specified quantity of product. Ct values correlate inversely with the quantity of starting template (e.g. lower Ct values indicate greater quantities of starting template).

We compared five DNA extraction methods using paucicellular samples from various aneuploid and diploid cells. The selected extraction methods are representative of the diverse approaches that are commonly employed in modern laboratories: organic solvent extraction, non-solvent-based
enzymatic digestion, solid-phase adsorption, sequential protein and DNA precipitation, and magnetic bead adsorption. The endpoints compared were proportion of extractions producing amplifiable DNA and Ct values for GAPDH as measured by real-time PCR.

2. Materials and methods

2.1. Cells

DNA was extracted from two aneuploid tumor cells lines (cervical cancer cell line HeLa [13] and breast cancer cell line HCC1806 [14]), two diploid Human Mammary Epithelial Cell (HMEC) cultures (UTSW991 and UTSW1004), and freshly isolated lymphocytes from two donors. The lymphocytes were isolated from whole blood using Vacutainer CPT tubes (Becton Dickinson and Company, Franklin Lakes, NJ) according to the manufacturer’s instructions. Based on hemocytometer counts 50, 500, 5000, or 50,000 cells were pelleted for DNA extraction.

2.2. DNA extraction

The salient features of each of the five DNA extraction methods are summarized in Table 1. Separate DNA extractions were performed in triplicate for each of the six cell types, each of the four starting cell counts and each of the five methods (360 extractions). Three methods employed commercially available kits (QIAamp, Puregene, and Dynabeads, respectively) in which case extractions were performed according to the manufacturer’s instructions. A standard phenol/chloroform extraction method was tested as well. Briefly, the cells were suspended in 350 µl TE buffer with 20 µl of 10% SDS. Ten microliter of proteinase K at 10 µg/ml was added and the tubes incubated at 37 °C for 36 h. The contents of the tubes were mixed with 20 µl of 5 M NaCl after which 400 µl of phenol:chloroform:isoamyl alcohol step was repeated twice and the supernatant transferred to a fresh tube. The phenol:chloroform:isoamyl alcohol step was repeated twice more and then the DNA was precipitated with 800 µl of cold absolute ethanol. For the single-step proteinase K method, the cell pellet was suspended in 50 µl of extraction solution prepared by combining 0.5 M EDTA pH 8.0 (20 µl), 1 M Tris pH 8.0 (200 µl), Tween-20 (50 µl), Proteinase K at 20 mg/ml (100 µl), and ultrapure water (9.63 ml). The tubes were incubated at 37 °C for 36 h after which the proteinase K was inactivated by heating to 95 °C for 10 min.

Each extract was brought to a final volume of 20 µl in ultrapure water and the A260 and A280 measured spectrophotometrically after diluting 2 µl of the extract in 98 µl of water (Beckman DU-64, Beckman Instruments, Inc.). Samples were stored at 4 °C until analysis (a maximum of 3 weeks).

2.3. Real-time PCR

Genomic DNA for GAPDH was amplified by real-time PCR after combining 20 µl SYBR Green Jumpstart Taq ReadyMix (Sigma, 20 mM Tris–HCl at pH 8.3, 100 mM KCl, 7 mM MgCl2, dNTP’s at 0.4 mM each, 0.05 units/µl Taq DNA Polymerase, JumpStart Taq antibody, and SYBR Green I), 0.4 µl internal reference dye, 18 µl of water and 2 µl of the DNA extract. GAPDH primers were used at a final concentration of 0.2 µM; fwd: GCCTGCTTCAC-CACCTTCTTG, rev: GTCCACCTGGCCTTTCACAC. The PCR was run on an ABI GeneAmp 5700 Sequence Detection System (Perkins–Elmer Applied Biosystems) as follows: after a 1 min preincubation at 95 °C, amplification cycles of 95 °C for 15 s and 60 °C for 1 min were repeated 50 times. DNA extractions were scored as successful if the GAPDH Ct was less than that of the water blank.

2.4. Cost

The cost per extraction was calculated for the commercially available kits by dividing the cost of the kit by the number of extractions that could be performed with the kit. The costs of phenol/chloroform and single-step proteinase K extractions were calculated based on the cost of the consumable supplies required for each extraction. The per extraction cost of each method was compared relative to the cost of phenol/chloroform alcohol extraction (i.e. per extraction cost for Method X divided by the per extraction cost for phenol/chloroform extraction).

2.5. Validation using alcohol-fixed clinical samples

Six alcohol-fixed (CytoLyt, Cytyc Corporation, Marlborough, MA) nipple duct lavage (NDL) samples containing <1000 cells and four NDL samples reported as acellular were extracted using the Puregene method in order to determine whether this method was suitable for alcohol-fixed samples. A260 and A280 values were measured, and real-time PCR performed as described in Section 1. Each extract was diluted

---

Table 1

<table>
<thead>
<tr>
<th>Method</th>
<th>Name</th>
<th>Vendor</th>
<th>Catalog no.</th>
<th>Salient features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>QIAamp</td>
<td>Qiagen</td>
<td>51104</td>
<td>Silica-gel membrane binds DNA, contaminants pass through spin column</td>
</tr>
<tr>
<td>2</td>
<td>Puregene</td>
<td>Genta</td>
<td>D-5500A</td>
<td>Protein precipitation followed by DNA precipitation</td>
</tr>
<tr>
<td>3</td>
<td>Dynabeads</td>
<td>Dynal</td>
<td>630.06</td>
<td>DNA adsorbed onto magnetic beads</td>
</tr>
<tr>
<td>4</td>
<td>Phenol/</td>
<td>N/A</td>
<td>N/A</td>
<td>Proteinase K, followed by organic extraction, followed by DNA precipitation</td>
</tr>
<tr>
<td></td>
<td>chloroform</td>
<td></td>
<td></td>
<td>Proteinase K without DNA precipitation</td>
</tr>
</tbody>
</table>

N/A, not applicable.
twofold from 1:2 to 1:16 and the real-time PCR was run in triplicate for each dilution using 1 μl of template.

2.6. Statistical analysis

The proportion of extractions producing amplifiable DNA was compared using Fisher’s exact test. Mean Ct values are reported with 95% confidence intervals and compared using Student’s t-tests. All statistical comparisons are two-tailed. No adjustments were made for the multiple comparisons.

3. Results

3.1. DNA yield for different cell types

In general, amplifiable DNA yield, as measured by Ct for GAPDH, was best for the cancer cells intermediate for HMEC and worst for lymphocytes for all methods and for all starting cell numbers. Analyzing the data for lymphocytes and epithelial cells separately did not affect the conclusions concerning the relative efficiencies of the five methods, so data for the six cell sources were combined.

3.2. Number of successful extractions

The QIAamp and Puregene extractions produced amplifiable DNA more frequently than the other methods when the starting cell counts were <50,000 (Fig. 1). The proportion of extractions producing amplifiable DNA was similar for QIAmp and Puregene across all starting cell numbers. For starting cell numbers of 50,000, the proportion of extractions producing amplifiable DNA was similar for all methods except for the single-step proteinase K extraction which performed poorly for all starting cell numbers.

3.3. Ct values by real-time PCR

Mean GAPDH Ct values, 95% confidence intervals, and statistical comparisons for the five methods and four starting cell numbers are depicted in Table 2. Of note, mean GAPDH Ct value for the 50-cell extractions was lowest for Puregene, but this result was not statistically significant. GAPDH Ct values for QIAamp extractions showed the greatest dynamic range and the best linearity ($R^2 = 0.919$) across the range of starting cell numbers (Fig. 2).

3.4. Absorbance at 260 nm

A260 values ranged from 0.029 to 0.034 for QIAamp, 0.024 to 0.060 for Puregene, 0.180 to 0.663 for phenol/chloroform, and 0.341 to 0.542 for single-step proteinase K. GAPDH Ct did not correlate with A260 for any method. The A260/A280 ratios ranged from 2.82 to 4.83 for QIAamp, 1.81 to 2.26 for Puregene, 1.41 to 1.46 for phenol/chloroform, and 0.64 to 0.70 for single-step proteinase K extractions.

3.5. Relative costs

Compared to phenol/chloroform extraction (relative cost = 1.0), the relative supply costs for the other methods were 6.2 for QIAamp, 0.3 for Puregene, 1.9 for Dynabeads, and 0.6 for single-step proteinase K. On the average, QIAamp and Puregene required 2 h of technician time from start to finish, Dynabeads 2.5 h, phenol/chloroform 3 h and single-step proteinase K 30 min.
specifically, there was excellent linearity between GAPDH Ct and starting template quantity (Fig. 3). The A260 for these samples was similar to that observed for the unfixed test samples.

3.6 Alcohol-fixed clinical samples

Performance of the Puregene method for alcohol-fixed NDL samples was similar to that observed for the unfixed test samples. Specifically, there was excellent linearity between GAPDH Ct and starting template quantity (Fig. 3). The A260 for these samples correlated well with GAPDH Ct ($R^2 = 0.436, P = 0.03$).

4 Discussion

Modern clinical and translational research protocols often require PCR amplification of DNA obtained from paucicellular clinical samples. We used benign and malignant epithelial cells as well as lymphocytes for this comparative evaluation as these are representative of the samples that are frequently obtained for translational studies. The primary endpoint was amplifiable DNA as measured by real-time PCR. The Puregene method, which employs sequential protein and DNA precipitation steps, and the QIAamp method, which is based on solid-phase adsorption of DNA, outperformed the other methods for extraction of DNA from paucicellular samples. The Puregene method provided the additional advantage of lower cost. Our laboratory is primarily interested in DNA from alcohol-fixed breast epithelial cells obtained by nipple duct lavage or random periareolar fine needle aspiration biopsy. There are a variety of additional DNA extraction methods that we did not evaluate and a variety of modifications available for the assays we did evaluate. Our conclusions may not be generalizable to other applications, particularly those that use formalin-fixed tissues. Fifty cells was the lowest cell count we evaluated. Some clinical samples will contain fewer than 50 cells of interest. We are primarily interested in quantitative real-time PCR for gene copy number determinations and promoter region methylation studies and have found that the reproducibility of these assays declines precipitously when starting template is generated from fewer than 50 cells. Reproducible methods for DNA extraction and analysis of ultra-paucicellular clinical samples are needed.

4.1 Traditional approaches to DNA extraction

Phenol–chloroform–isoamyl alcohol extraction is the classical method for extracting DNA from clinical samples. A Proteinase K treatment step is often incorporated as the enzyme degrades proteins into sub-tetrameric fragments and has been shown to improve the efficiency of PCR-based applications by destroying DNases and RNases [15]. The multiple centrifugation steps required for phenol–chloroform–isoamyl alcohol extraction is cumbersome prompting some to evaluate proteinase K digestion with [16] or without [17,18] a single salting-out step rather than solvent extraction.

4.2 Previously published comparisons of DNA extraction methods

A study that compared methods for extracting DNA from bone marrow cells scraped from Giemsa-stained slides used amplification of β-globin and β-actin genes as the endpoint measure [16]. The PCR products were electrophoresed on 4% agarose gels and extraction scored as successful if a band could be detected by ethidium bromide staining. Bands were detected in 20 of 20 samples extracted using the QIAamp or proteinase K salting-out methods, and 19 of 20 samples using classical phenol–chloroform–isoamyl alcohol extraction (with proteinase K digestion). The three non-enzymatic boiling methods that were tested produced bands in only 35–80% of samples. The traditional phenol–chloroform–isoamyl alcohol extraction did not perform nearly as well in our hands. It is conceivable that the use of a DNA/RNA carrier such as glycogen and the use of phase-separating tubes would have improved the efficiency of this approach for paucicellular samples. Nevertheless, the multiple pipetting and centrifugation steps required by this approach limits its utility for larger studies.

In another study, DNA suitable for PCR amplification was obtained from 13 of 14 fine needle aspiration samples of putative renal cell carcinoma metastases extracted using a single-step proteinase K method [17]. In this study, the extraction was scored as successful if PCR-amplified
microsatellite repeats resolved on a 6% acrylamide gel could be visualized by autoradiography. The single-step proteinase K method (without salting-out) was also found to be superior to phenol–chloroform extraction when applied to formalin-fixed paraffin-embedded thymoma tissue as determined by A260 values and intensity of SYBR Green I stained bands obtained by electrophoresing β-globin gene PCR products [19].

In contrast to these results are those obtained for DNA extracted from buccal cells using three different methods [20]. Amplification of the β-globin gene detected by ethidium bromide staining of electrophoresed PCR products was the endpoint measure in this study. Phenol–chloroform–isoamyl alcohol extraction (with proteinase K) was successful in 16/17 (94%) samples, QIAamp extraction in 12/16 (75%) and single-step proteinase K extraction in 2/16 (13%). A comparison of five commercially available solid-phase adsorption kits that used A260 as the endpoint concluded that the Genomic DNA Isolation Kit provided by Sigma produced the best DNA yields when the starting sample was whole blood, buccal swabs or muscle [21]. The QIAamp kit, which was included in this study, ranked in the middle for these five kits, but the analysis cannot be considered valid as A260 was the only measure of DNA yield employed.

4.3. A260 determinations as measure of DNA content

Quantities of nucleic acids in solution are often estimated based on the absorbance of light at a wavelength of 260 nm. An A260 of 1.0 correlates roughly with a double stranded DNA content of 50 µg/ml. A260 values between 0.10 and 1.00 are thought to correlate in a linear fashion with nucleic acid content. The A280 is traditionally taken as a measure of protein content in a solution (though nucleic acids absorb a considerable amount of light at 280 nm) and the A260/A280 ratio as a measure of the purity of the nucleic acid extract. A260/A280 ratios of 1.8–2.0 are generally considered relatively free of protein contamination, though in reality a solution with an A260/A280 of 1.8 may represent a 60/40 mixture of protein and nucleic acids [22]. A pure nucleic acid solution should have an A260/280 of 2.0.

QIAamp, Puregene and Dynabeads all produced A260 values $<0.10$, even with the 50,000 cell extractions. This is not unexpected when DNA is extracted from paucicellular samples using methods that effectively exclude protein carryover. Phenol–chloroform extraction and the single-step proteinase K method both produced A260 values well above 0.10 but these values must be viewed as artifactual based on the poor PCR performance that was observed when these methods were applied to paucicellular samples. It is likely that both methods were compromised by significant protein carryover or carry-over of factors that inhibit PCR [23].

4.4. Conclusions

Prior comparisons of DNA extraction methods have largely relied on the detection of PCR products on electrophoresis gels as the endpoint measure. Real-time PCR is a more sensitive and specific measure of amplifiable DNA. Each of the methods we tested yielded amplifiable DNA, but the QIAamp and Puregene methods were successful more frequently when starting cell numbers were low and were associated with the lowest Ct values by real-time PCR. These methods would seem best suited to paucicellular clinical samples.

References


