CEBP Focus: Biorepository and Biospecimen Science

Approaches to DNA/RNA Extraction and Whole Genome Amplification

Regina M. Santella

Mailman School of Public Health, Columbia University, New York, New York

Abstract

Analysis of DNA and/or RNA is an important component of most epidemiologic studies. The methods used for their preparation vary depending on the number of samples in the study as well as the amount of tissue or cells available, the specific downstream assay, and the resources available. They range from classic phenol/chloroform extractions to robotic methods suitable for large-scale studies. Several methods are also available for the quantitation of DNA/RNA, including UV or fluorescence measurements and PCR. For DNA samples of limited quantity, it is also now possible to use whole genome amplification to increase the amount of DNA suitable for genotyping. Finally, a major issue in quality control is the careful labeling and handling of samples. (Cancer Epidemiol Biomarkers Prev 2006;15(9):1585–7)

The Preparation and Storage of DNA for Molecular Epidemiology Studies

The use of DNA for both adduct measurement and genetic studies has rapidly expanded in molecular epidemiology studies. DNA can be prepared from blood (total WBC or fractionated cells, plasma or serum, and blood clots), exfoliated buccal or bladder cells, or frozen tissue samples as well as paraffin tissue blocks. DNA is generally stable under appropriate storage conditions and thus can be prepared in batches as samples are received or, alternatively, all samples can be prepared when needed. The method of DNA extraction is dependent on the amount of sample available, the number of samples to be extracted, the availability of specific equipment such as robotics, and, in some situations, the ultimate use of the DNA (1). The goal is to remove lipids and other proteins from the nucleic acids and remove or degrade the RNA. The oldest procedure is the classic phenol/chloroform extraction method. In this method, the cells or tissue homogenates are lysed with a detergent such as SDS. Buffers contain EDTA as a chelating agent to tie up the divalent cations required by endogenous nucleases and thus inactivate them. Proteinase K digests proteins and RNase digests RNA. Samples are then extracted with phenol and choloform/isoamyl alcohol, which denatures the proteins. Centrifugation results in an upper aqueous layer containing the DNA and an interface above the organic layer that contains the precipitated proteins. Repeated extraction can be carried out until there is no interface. The salt concentration is then increased and the DNA precipitated with ethanol, washed with additional ethanol, and then redissolved in an appropriate buffer, usually containing EDTA. This method gives high yields of good-quality DNA. However, a major disadvantage is the use of highly toxic reagents and the extensive labor required. In addition, if the phenol is not completely removed, it will interfere with quantitation of the DNA using UV absorbance because phenol has a very high extinction coefficient at 260 nm. The development of gel lock tubes has simplified the process because, after centrifugation, the organic layer is locked under the gel. The aqueous layer can thus be easily removed simply by pouring off. In our laboratory, we only use this method when the samples must be run in a competitive ELISA because we have found that other methods result in contaminants that interfere with the assay. Similarly, for studies in which oxidative DNA damage will be measured, specific precautions must be taken, such as the addition of antioxidants, to ensure that the isolation procedure itself does not result in *in vitro* oxidation.

A simple procedure we have routinely used when DNA is needed for genotyping is salting out. In this method, cells are lysed and treated with proteinase K and RNase. Saturated NaCl is added to precipitate the proteins and the sample is centrifuged. DNA is again precipitated from the supernatant with ethanol. This method provides a good yield of high molecular weight, pure DNA, avoids toxic chemicals, and is relatively fast and cheap.

There are also a number of different types of commercially available kits for the extraction of DNA from Arcturus, Invitrogen, Qiagen, Stratagene, and others with detailed information available on the web. These kits are also used for extracting DNA from plasma or serum. Most take advantage of the DNA binding to a silica-gel membrane allowing the elution of contaminants followed by the release of DNA. Whereas the yield is sometimes slightly lower than other methods and there may be some degradation of DNA, the major advantage is that the use of phenol and chloroform is avoided and these methods are fast and easy. If small amounts of DNA are needed, 96-well DNA extraction kits can be used to isolate ~6 μg DNA from 200 μL of whole blood. However, a centrifuge that can take 96-well plates is required equipment. Methods, including specific commercial kits, are also available for extracting DNA from clotted bloods but the yields are lower than comparable blood samples and the methods are more labor intensive (2).

Received 5/3/06; accepted 6/16/06.

Grant support: NIH grants P30ES09089, R01ES05116, and P30 CA013696.

Note: This article is one of a series of articles that were presented at a methods workshop, "Sample Collection, Processing and Storage for Large Scale Studies: Biorepositories to Support Cancer Research," held during the AACR 97th Annual Meeting in 2006.

Requests for reprints: Regina M. Santella, Columbia University, 630 West 168th Street, New York, NY 10032. Phone: 212-305-1996; Fax: 212-305-5328. E-mail: rps1@columbia.edu

doi:10.1158/1055-9965.EPI-06-0631

Table 1. Comparison of methods for quantitation of DNA samples

Method	Absorbance at 260 nm	Picogreen	RT-PCR SYBR green
Sensitivity	>5 μg/mL	25 pg/mL-1 μg/mL	<25 pg
Type of DNA detected	Single and double stranded	Double stranded	Single and double stranded
Interference by RNA	Yes	No	No

For those laboratories that carry out large-scale DNA extractions, the cost of investing in an automated system may be worthwhile. For example, Gentra[1] has a system that can handle whole blood samples ranging from 0.05 to 10 mL and yield DNA typically of ~50 kb in length within 30 minutes to 1 hours.

Simple methods are available for extraction of DNA from paraffin (e.g., ref. 3), as well as commercial kits.[2] When DNAs are extracted from paraffin sections, the size of the product is much smaller than for other types of samples. This may require the development of new PCR primers that produce smaller products or a poor success rate will be obtained.

Methods for Quantitation of DNA Yield and Purity

Three primary methods are available for quantitation of DNA yield (Table 1). When DNA yields are sufficient, the easiest method is absorbance at 260 nm, the peak for DNA. Using the extinction coefficient for double-stranded DNA (ε = 6,500) and Beer's Law, absorbance is converted into molar concentration and then μg/mL. One absorbance unit in a 1-cm cuvette corresponds to ~50 μg/mL. Both single-stranded and double-stranded DNAs contribute to absorbance at 260 nm but ssDNA has a higher absorption due to the hypochromic effect in double-stranded DNA. Any RNA present will also contribute to the absorbance. The absorbance at 320 nm should also be determined and be almost zero because DNA does not absorb at this wavelength. Significant absorption may indicate insoluble material in the solution that results in light scatter. Centrifugation or further purification is necessary under these circumstances.

When insufficient amounts of DNA are available for UV absorbance, a method based on the generation of fluorescence when picogreen intercalates into double-stranded DNA can be used. This method is able to measure pg/mL amounts of DNA but will only detect double-stranded DNA. A standard curve must be generated using known quantities of DNA. This method will not be influenced by RNA contamination. The most sensitive method of quantitation is real-time PCR with SYBR green, a dye which binds to double-stranded DNA but not to ssDNA. Again a standard curve is generated by using known amount of DNA and plotting the threshold cycle versus DNA concentration. The threshold cycle for the unknown DNA is used to determine its concentration. Whereas the dye fluoresces only when binding to double-stranded DNA, both single-stranded and double-stranded DNAs can be detected. If the DNA is to be used for genotyping, this method of quantitation also shows that the sample is of appropriate quality for PCR. With good-quality DNA, the three methods generally agree but discrepancies can be found with some samples.

Whole Genome Amplification

When DNA amounts are limited, the technique of whole genome amplification can be used. In this method, random hexamer primers anneal at multiple sites on denatured DNA and Φ29 DNA polymerase initiates replication. Strand displacement of downstream replicated DNA generates new ssDNA to which additional primers bind. This isothermal process produces large quantities of high molecular weight, double-stranded DNA, from nanogram quantities to 40 μg. Qiagen reports that the average product length using their REPLI-g kit is typically >10 kb, with a range between 2 and 100 kb. A number of publications have shown that whole genome amplification gives reliable genotyping data compared with the original DNA sample including that of plasma DNA (3, 4). In contrast, specific methylation of cytosine is lost during whole genome amplification. It is also less clear that whole genome amplification can be done with paraffin DNA (3).

Quality Control of DNA

The most frequently used method to determine DNA quality is the ratio of absorbance at 260 and 280 nm ($A_{260\ nm}/A_{280\ nm}$). Ratios around 1.8 indicate good-quality DNA whereas lower values indicate protein contamination because proteins have a peak in absorption at 280 nm resulting from the aromatic amino acids. Higher values indicate RNA contamination. DNAs can also be run on a gel to determine size. As an identity check of the DNA sample, highly polymorphic microsatellite repeats can be determined and compared with another source of DNA from the same individual. For example, if a filter card blood spot was made from the initial blood sample, DNA can be isolated from it and alleles compared. This method can also be used to confirm that a cell line made from a blood sample corresponds to that individual. The presence of the Y chromosome can also be determined as a partial identity check.

DNA is usually stored in a Tris-EDTA solution for maximum stability. Short-term storage can be at 4°C, but for long-term storage, samples should be kept at −80°C. Several reports have indicated variable stability of dilute and small volume samples (5, 6) but we have not had difficulties with long-term storage at −80°C of DNAs around 200 ng/μL. Usually storage is in either sealed vials or 96-well deep-well storage plates. Storage in well-sealed vials gives minimal evaporation and thus maintains concentration. However, the handling of large numbers of samples is more easily done robotically with 96-well plates. For most of our studies, especially when genotyping is to be carried out, the stock DNAs are kept in 1.5-mL vials at its initial concentration. But at the same time, 96-well storage plates are prepared at a standard concentration appropriate for making replica plates for genotyping.

RNA Extraction

RNA is a much less stable molecule than DNA and may be better isolated when needed. Care must be taken to avoid

[1] http://www.gentra.com.
[2] For example, http://www.arctur.com/.

contamination with RNases during its preparation and storage. Trizol, a monophasic solution of phenol and guanidine isothiocyanate, is commonly used. During sample homogenization or lysis, it maintains the integrity of the RNA while disrupting cells and dissolving cell components. Addition of chloroform followed by centrifugation separates the solution into an aqueous phase containing the RNA and an organic phase. RNA is recovered by precipitation from the aqueous phase with isopropyl alcohol. After removal of the aqueous phase, the DNA and proteins in the sample can also be recovered by sequential precipitation with ethanol and additional isopropanol. The most frequently used methods for RNA extraction including from paraffin sections employ commercial kits such as the RNAeasy kit from Qiagen. Arcturus also produces a line of products that can be used for RNA isolation from paraffin sections as well as the amplification of as little as 5 ng of RNA. Quantitation is by absorption at 260 nm, with 1 absorbance unit equal to 40 μg/mL. Because absorbance is influenced by pH, samples should be in 10 mmol/L Tris-HCl (pH 7.5).

The quality of the RNA can be evaluated from the $A_{260\ nm}/A_{280\ nm}$ with values of 1.9 to 2.1 being ideal. Analysis by gel electrophoresis should show discrete bands of high molecular weight RNA between 7 and 15 kb in size, two predominant rRNA bands at ~5 kb (28S) and 2 kb (18S) in a ratio of ~2:1, and low molecular weight RNA between 0.1 and 0.3 kb (tRNA, 5S). Long-term storage should be at −80°C or as an ethanol precipitate at −20°C.

For laboratories that deal with large numbers of samples, a major concern is errors in sample identification and swapping. For these reasons, bar code labels should always be used with no hand labeling. The use of robotics also minimizes sample swapping.

References

1. Austin MA, Ordovas JM, Eckfeldt JH, et al. Guidelines of the National Heart, Lung, and Blood Institute Working Group on Blood Drawing, Processing, and Storage for Genetic Studies.[erratum appears in Am J Epidemiol 1997 Mar 15;145(6):570]. Am J Epidemiol 1996;144:437–41.
2. Everson RB, Mass MJ, Gallagher JE, et al. Extraction of DNA from cryopreserved clotted human blood. Biotechniques 1993;15:18–20.
3. Sjoholm MIL, Hoffman G, Lindgren S, et al. Comparison of archival plasma and formalin-fixed paraffin-embedded tissue for genotyping in hepatocellular carcinoma. Cancer Epidemiol Biomarkers & Prev 2005;14:251–5.
4. Lu Y, Gioia-Patricola L, Gomez JV, et al. Use of whole genome amplification to rescue DNA from plasma samples. Biotechniques 2005;39:511–5.
5. Sozzi G, Roz L, Conte D, et al. Effects of prolonged storage of whole plasma or isolated plasma DNA on the results of circulating DNA quantification assays. J Natl Cancer Inst 2005;97:1848–50.
6. Madisen L, Hoar DI, Holroyd CD, et al. DNA banking: the effects of storage of blood and isolated DNA on the integrity of DNA. Am J Med Genet 1987; 27:379–90.

Approaches to DNA/RNA Extraction and Whole Genome Amplification

Regina M. Santella

Cancer Epidemiol Biomarkers Prev 2006;15:1585-1587.