Optimal yield between 6 and 8 weeks after birth. There was no significant association with the number of cells recovered and milk volume, method of collection, or the number of samples in a 24-h period per individual. This study demonstrates that breast milk can be used as a source of DECs for biomarker studies of gene-environment interaction and that sufficient DNA can be recovered to evaluate carcinogen-DNA adducts and to perform genotyping assays. Using this approach, exfoliated DECs may serve as a source of representative cells for studies of breast carcinogenesis and biomarkers of exposure, susceptibility, and effect.

Introduction
Breast cancer is the most commonly occurring cancer among women in the United States, yet, to date, little is known regarding its etiology (1). Although there are numerous studies investigating breast cancer risk factors and the carcinogenic process itself on a cellular and molecular level, they may be hampered by the lack of representative cells from which those cancers arise, i.e., epithelial cells that line the lumen of milk ducts. Numerous assays could be performed with such representative cells, investigating many different hypotheses regarding breast cancer etiology. Furthermore, DNA extracted from these representative cells could be used for molecular epidemiological studies of markers of exposure on a molecular level and the interaction of environmental factors with genetic susceptibility.

We were particularly interested in identifying and characterizing DNA adducts in human breast cells. Characterization of these adducts could lend insight into etiological factors in breast carcinogenesis, endogenous or exogenous, and could clarify the possible role of chemical carcinogens in breast cancer etiology. Adduct levels could be evaluated in relation to genetic polymorphisms in those enzymes that metabolize putative breast carcinogens as well as to exposures as measured by questionnaire data.

There have been several studies in which carcinogen-DNA adducts were assessed in human breast tissue. One study found no detectable adduct levels among six autopsy specimens studied (2). Seidman et al. (3) found adducts in mammary epithelial cells from 3 of 10 human donors. Recently, other investigators have examined breast tissue for DNA adducts and found 32P-postlabeled aromatic adduct spots on chromatograms (4, 5), as well as putative malondialdehyde-DNA adducts (6). In these studies, it seemed that DNA obtained from whole tissue was not specific to epithelial cells, particularly ductal cells, and thus was not necessarily reflective of cells from which breast cancers arise. An additional problem with using tissue derived from surgery for breast cancer or from reduction mammoplasty is that these tissues may be difficult to procure, and preparation (separation from adipose tissue) is labor intensive. Clearly, a source of an enriched population of DECs that is readily and abundantly accessible would greatly facilitate studies of breast carcinogenesis.

An excellent source of luminal epithelial cells for analysis may be human breast milk. These cells reflect what actually occurs in the breast and are the cells from which most breast cancers arise (7–9). They are also more easily accessible, readily available, and noninvasive than those from surgically

Received 5/6/97; revised 9/5/97; accepted 9/30/97.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 To whom requests for reprints should be addressed, at Division of Molecular Epidemiology, National Center for Toxicological Research, 3900 NCTR Road, Jefferson, AR 72079.

2 The abbreviations used are: DEC, ductal epithelial cell; K-H buffer, Krebs-Henseleit bicarbonate buffer; NAT2, N-acetyltransferase 2.
excised tumor tissue and reduction mammoplasty. It has been established that human milk contains numerous cells, including macrophages, lymphocytes, neutrophils, and epithelial cells (10). After 8 days postpartum, the majority of epithelial cells are secretory cells and are believed to be shed from the luminal alveoli (11). One light and electron microscopic analysis of human breast milk revealed that after 4 weeks postpartum, 56% of the cells counted were epithelial cells, and this proportion increased to 71% at 2–4 months and 84% at 4–6 months (12). Because the majority of breast cancers arise from luminal epithelial cells, analysis of these exfoliated cells, rather than basal cells, would be ideal. Analysis of DNA from luminal epithelial cells derived from human breast milk for DNA adducts and evaluation of adduct levels in relation to exposures and genetic polymorphisms could greatly elucidate breast cancer etiology. Use of these epithelial cells could also be of value for other studies of breast carcinogenesis. In this pilot study, we sought to develop and refine methodology for separation of luminal epithelial cells from human breast milk and to extract and quantify DNA from these cells for the purpose of assessing the feasibility of use of these cells in studies of DNA adducts, carcinogen metabolism, and gene-environment interactions. We were also interested in evaluating possible predictors of the yield of DNA from individual samples to optimize future yields.

**Materials and Methods**

**Study Population and Data Collection**

Pregnant women attending breast-feeding classes were informed of the study by the nurse lactation specialist at Millard Fillmore Suburban Hospital in western New York. Those who indicated that they were interested in participating were contacted 2–4 weeks after their due date by a nurse interviewer who discussed the study with them. The protocol for this study was approved by the Institutional Review Boards of the School of Medicine and Biomedical Sciences at the State University of New York at Buffalo, Millard Fillmore Suburban Hospital, and the Food and Drug Administration. Women who agreed to participate were mailed a consent form, a sterile specimen cup with detailed instructions for obtaining the breast milk specimen, and a questionnaire for information on demographics and possible exposures to putative breast carcinogens, such as aromatic and heterocyclic amines, and polycyclic aromatic hydrocarbons. We were interested in possible predictors of cell yield and DNA quantity and hypothesized that if women collected specimens over a 24-h period, with several samples from each breast, it would result in a greater volume of exfoliated DEC s. Women were asked to collect and refrigerate the milk (about 60 ml, if possible) at an agreed-on time, to be collected by the nurse interviewer. At the time of milk collection, another questionnaire was administered that included queries regarding breast-feeding habits and practices, such as possible breast infection or inflammation, frequency of nursing, method of obtaining milk (pump or hand expression), how many specific samples were collected over a 24-h period, and if the baby was nursed directly before the specimen was collected. These specimens were transported on ice to the laboratory in the Department of Social and Preventive Medicine, State University of New York at Buffalo, for processing.

**Breast Epithelial Cell Enrichment Procedure for Breast Milk Samples**

Human milk samples were collected and diluted 3:1 with K-H buffer containing 1% BSA (w/v). The milk cells were pelleted by centrifugation at 500 × g for 10 min at room temperature. The milk fat and supernatant were removed by aspiration; the cell pellet was resuspended with 10 ml of K-H buffer and washed two times. The cell pellet was resuspended in 2 ml of K-H buffer and transferred to a 20-ml glass scintillation vial, flushed with O2/CO2 mix (19:1), and incubated for no more than 2 h at 37°C. A simple adherence of monocytes to glass was used to enrich these samples for epithelial cells. The nonadherent cells were removed and washed two times with 10 ml of K-H buffer without BSA. Just before the final wash, the cells were counted on a Coulter counter and examined microscopically. Viability was determined by staining with trypan blue. The cell recoveries agreed with published estimates of total nucleated cells/ml breast milk, with a range of 5 × 10^6 to 1 × 10^7 cells/ml and an average of 5 × 10^6 cells/ml. The majority of these cells were macrophages and lymphocytes, and less than 20% of these cells were epithelial cells, as reported in the literature (13). After the glass adherence step, the individual cell pellets were enriched for epithelial cells (65 ± 22%). A representative slide of these cells is shown in Fig. 1. After the final wash, the cell pellet was stored at −80°C until shipment to the National Center for Toxicological Research, where the DNA isolation procedure was performed. As noted, these enriched epithelial cell pellets contain proteinaceous debris that subsequently hindered the purity of DNA recovered. Modifications noted below were made to reduce the protein contamination during the DNA isolation procedure. Numerous attempts to reduce protein precipitation in the epithelial cell enrichment procedure were unsuccessful.

**DNA Isolation and Quantitation**

DNA was isolated from the frozen cell pellets using a modification of the PureGene DNA isolation kit (Gentra Systems, Research Triangle Park, NC). Briefly, cell pellets were lysed using the cell lysis solution provided with the kit. Pellets were
vortexed vigorously for 2 min, followed by incubation with shaking in the presence of proteinase K at 55°C for 3 h. After complete homogenization of the cell pellet, the samples were treated with RNases A (150 μg/ml) and T1 (1,000 units/ml) for 30 min at 37°C, as described by Gupta et al. (14). The homogenate was extracted sequentially with Tris-EDTA buffer-saturated phenol, phenol:Sevag (chloroform-isoamyl alcohol, 24:1), and Sevag. DNA was precipitated in 0.1 volume of 5 M NaCl and 1 volume of chilled ethanol and collected by centrifugation (10,000 rpm for 10 min at 4°C). The DNA pellet was washed twice to remove residual salt.

Routine UV spectroscopy was performed for DNA quantification of samples with adequate DNA. For those samples that could not be quantified by UV spectrophotometry, normal 32P nucleotide analysis was performed (14, 15). Because the concentration of DNA in each of these samples was not known, range-finding experiments were performed using a known concentration of calf thymus DNA as a standard, as described previously (15). Initial samples (5 μl) were hydrolyzed to the 3’ nucleoside monophosphates with micrococcal nuclease and spleen phosphodiesterase according to standard protocols for 3 h at 37°C. Fivefold serial dilutions of the hydrolyzates were postlabeled using 50 pmol of [32P]ATP (specific activity, 7000 Ci/mmol). The serial dilutions were performed to ensure ATP excess to accurately quantify the DNA, using 10 pmol of calf thymus DNA to generate a standard curve using 2-fold serial dilutions.

**DNA Purity Determination**

DNA concentration was measured spectrophotometrically to assess purity and concentration of nucleic acids. A260 measurement using 1 = 50 μg/μl is quantitative for relatively pure preparations in microgram quantities. 32P postlabeling of normal nucleotides was performed on those samples indicated (Table 1) that were below the detectable range using spectrophotometry. The incorporation of 32P, followed by TLC, was used to measure the contamination of normal nucleic acids as well as to determine the degree of RNA contamination. RNA-derived nucleotides migrate slower than the DNA counterparts, and RNA contamination can be visualized in the samples as four additional spots.

**Characterization of DNA Isolated from Enriched Breast Epithelial Cells**

**Normal Nucleotide Analysis by 32P Postlabeling**. To determine the purity of DNA obtained from the enriched breast epithelial cells using the PureGene kit, normal nucleotide analysis was performed with excess ATP on several samples using the 32P postlabeling method, as described previously (15). Samples of DNA (10 pmol) were labeled with 40 pmol of [32P]ATP (7000 Ci/mmol, 280 μCi). The excess ATP was then removed by incubation with apyrase. An aliquot of each sample was applied to polyethyleneimine-cellulose F, predeveloped in water to the origin, and then developed to the top of one chromatogram in 0.17 M sodium phosphate (pH 6.8). The chromatogram was then dried and analyzed using a Hewlett-Packard InstantImager (Meriden, CT). Representative data are shown in Fig. 2a.

**32P Postlabeling**. 32P postlabeling of the enriched breast epithelial cell DNA was performed essentially as described earlier. The preliminary results presented here were performed using ATP-deficient conditions with no enrichment procedures to eliminate normal nucleotides. Briefly, 1 μg of DNA was digested with micrococcal nuclease and spleen phosphodiesterase as described: 0.4 μg of the DNA hydrolyzate was reacted with polynucleotide kinase (United States Biochemical; 4 units) in the presence of 160 μCi of [32P]ATP (2.3 mCi) for 45 min at 37°C. After apyrase treatment to remove unreacted ATP, samples were spotted onto polyethyleneimine-cellulose thin layer plates. TLC was performed using a solvent system designed specifically to examine the mobility of the spots (16). D1 was performed using 1.0 M sodium phosphate (pH 6.0) in the opposite direction of D3. The plate was cut several centimeters behind the origin and imaged using the InstantImager to detect the mobility of samples in D1. It was noted that the breast epithelial DNA samples tested had components that migrated in D1. The plates were developed using the following solvent system to further study the mobility of the spots detected: D3, 3.6 M lithium formate and 6.8 M urea (pH 3.5); D4, 1.0 M lithium chloride, 0.5 M Tris-HCl, and 7.45 M urea (pH 8.0); and D5, 1.7 M sodium phosphate (pH 6.0). An example of these preliminary studies is shown in Fig. 2b.

| Table 1 | Amounts and quality of DNA recovered from exfoliated ductal epithelial cells from human breast milk |
|-------|---|---|
| Sample no. | Concentration (total μg recovered) | Quality* (OD_{260}/280) |
| 1 | 555 | 1.7/no RNA |
| 2 | 494 | 1.7/no RNA |
| 3 | 300 | 1.6/no RNA |
| 4 | 293.4 | 1.6/no RNA |
| 5 | 258 | 1.6/no RNA |
| 6 | 238 | 1.7/no RNA |
| 7 | 214 | 1.7/no RNA |
| 8 | 200 | 1.7/no RNA |
| 9 | 160 | 1.7/no RNA |
| 10 | 144 | 1.8/no RNA |
| 11 | 142 | 1.7/no RNA |
| 12 | 88 | 1.9/no RNA |
| 13 | 88.1 | 2.0/no RNA |
| 14 | 17.3 | 2.1/no RNA |
| 15 | 13.1 | 1.7/no RNA |
| 16 | 9.8 | 1.8/minor RNA contamination |
| 17 | 8.2 | 2.2/RNA contamination |
| 18 | 7.5 | 1.9/no RNA |
| 19 | 6 | 2.0/no RNA |
| 20 | 4.7 | 1.8/no RNA |
| 21 | 4.3 | 2.3/minor RNA contamination |
| 22 | 3.8 | 2.4/phenol contamination |
| 23 | 3.4 | 1.9/no RNA |
| 24 | 3 | 2.1/minor RNA contamination |
| 25 | 2.4 | 1.5/minor RNA contamination |
| 26 | 1.8 | 1.8/minor RNA contamination |
| 27 | 1.7 | 1.6/no RNA |
| 29 > 30 > 31 > 32 | <1 | Too low to determine by OD_{260}/280 (no detectable RNA) |

* DNA quality was determined for each of the samples by combining the yield, OD_{260}/280 ratio, and normal nucleotide analysis. These studies were performed to optimize the DNA isolation procedure to ensure high-quality DNA, particularly for postlabeling reactions in which both RNA and protein contamination may contribute to erroneous spots on TLC plates. Ratios of 1.8–2.0 indicate highly purified preparations of DNA or RNA; absorption of protein at 280 nm will lower this ratio. Notations indicate the OD_{260}/280 ratio and extent of RNA contamination determined by normal nucleotide analysis.

"Normal nucleotide analysis (NN) by TLC, described in "Materials and Methods," was performed on each individual sample to determine the extent of RNA contamination. The presence of greater than four spots by TLC was considered contamination with the slower-moving ribonucleotides.

Minor RNA contamination was considered less than 5% of total incorporated 32P in the ribonucleotides; 95% incorporated in the deoxyribonucleotides.
Exfoliated DECs from Human Breast Milk

**Materials and Methods.**

Exfoliated DECs from Human Breast Milk in demonstrating the utility of these samples for the detection of enriched breast epithelial cells isolated from human milk. 12P postlabeling was performed as described in "Materials and Methods" on four DNA samples isolated from enriched breast epithelial cell DNA obtained from human milk. Lane 0 obtained from enriched breast epithelial cell DNA obtained from human milk. A representative chromatograph is shown in Fig. 2b.

**Statistical Analysis**

Data were collected on a number of variables that could affect DNA yield. These included quantity of milk specimen, number of individual samples in each specimen, the method of breast milk collection (hand expression or pump), usual frequency of breast-feeding, if the infant was nursed immediately before the specimen was obtained, and number of weeks postpartum. Continuous variables were stratified into tertiles based on the distribution of the data, and associations between all possible predictors of DNA yield and quantity of DNA obtained were evaluated using $\chi^2$ analysis and Fisher’s exact test. Samples were collected from more than 100 women. However, refinement of the methodology for epithelial cell separation and DNA extraction required numerous specimens. These results are based on 38 breast milk samples.

**Results**

There was wide interindividual variability in total DNA recovered from the enriched breast epithelial cells obtained from human breast milk. Twenty-seven specimens had more than 1 µg, with values ranging from 1.7–555 µg (Table 1). Of the total specimens, 87% contained enough DNA to perform adduct analyses and genotyping assays.

When the quantity of DNA was examined in relation to possible predictors of cell yield, it was found that the quantity of DNA was not significantly associated with the volume of milk collected. Although women were asked to supply about 60 ml of breast milk, amounts ranged from 15–220 ml, although the median value was 60 ml, and the mean was 65 ml. The number of specimens collected over the 24-h period or if the infant was nursed immediately before milk collection also did not predict DNA yield. Of the variables evaluated, only the number of weeks postpartum was significantly associated with the quantity of DNA recovered ($P < 0.01$). As shown in Fig. 3, the optimal time for milk collection was between 6 and 8 weeks postpartum.

**Discussion**

This pilot study demonstrates that exfoliated ductal breast epithelial cells, those from which most breast cancers arise, may be separated from human breast milk for DNA extraction and that quantities are sufficient to provide a source for molecular epidemiological studies of gene-environment interactions. Use of tissue from surgery for tumor excision or for reduction mammoplasty is hampered by limited availability of the tissue which may indicate etiological agents in breast cancer. They

**PCR Amplification.** PCR amplification was performed in the presence of NAT2-specific primers 5'-CTCTGATCGATTCAACTCTGC and 5'-GGAACAAATTTGCTTCG using the method described previously by Bell et al. (17). This analysis was performed to demonstrate the ability of the enriched breast epithelial DNAs to be used in the PCR amplification reaction for genotype studies. A representative amplification is shown in Fig. 2c.

**Fig. 2.** a, autoradiogram of normal nucleotides of a representative sample from enriched breast epithelial cells isolated from human milk. 32P postlabeling was performed using excess [32P]ATP to demonstrate the purity of the DNA preparations obtained using the PureGene kit with the modifications noted in “Materials and Methods.” P, inorganic phosphate; T, thymidine-3',5'-bisphosphate; C, deoxycytidine-3',5'-bisphosphate; A, deoxyadenosine-3',5'-bisphosphate; G, deoxyguanosine-3',5'-bisphosphate. A, chromatogram of 32P-postlabeled DNA sample obtained from enriched breast epithelial cells isolated from human milk. 32P postlabeling was performed as described in “Materials and Methods” on four DNA samples isolated from human breast epithelial DNAs. A representative chromatograph is shown demonstrating the utility of these samples for the detection of in vivo modified DNA. O, origin. c, representative PCR amplification of NAT2 using human genomic DNA obtained from enriched breast epithelial cell DNA obtained from human milk. A single 1092-bp band is amplified for NAT2 using the primer pair described in “Materials and Methods.” Lane 1, control DNA; Lane 2, sample DNA; Lane 3, 100-kb ladder.

**Results**

There was wide interindividual variability in total DNA recovered from the enriched breast epithelial cells obtained from human breast milk. Twenty-seven specimens had more than 1 µg, with values ranging from 1.7–555 µg (Table 1). Of the total specimens, 87% contained enough DNA to perform adduct analyses and genotyping assays.

When the quantity of DNA was examined in relation to possible predictors of cell yield, it was found that the quantity of DNA was not significantly associated with the volume of milk collected. Although women were asked to supply about 60 ml of breast milk, amounts ranged from 15–220 ml, although the median value was 60 ml, and the mean was 65 ml. The number of specimens collected over the 24-h period or if the infant was nursed immediately before milk collection also did not predict DNA yield. Of the variables evaluated, only the number of weeks postpartum was significantly associated with the quantity of DNA recovered ($P < 0.01$). As shown in Fig. 3, the optimal time for milk collection was between 6 and 8 weeks postpartum.

**Discussion**

This pilot study demonstrates that exfoliated ductal breast epithelial cells, those from which most breast cancers arise, may be separated from human breast milk for DNA extraction and that quantities are sufficient to provide a source for molecular epidemiological studies of gene-environment interactions. Use of tissue from surgery for tumor excision or for reduction mammoplasty is hampered by limited availability of the tissue which may indicate etiological agents in breast cancer. They

**Fig. 2.** a, autoradiogram of normal nucleotides of a representative sample from enriched breast epithelial cells isolated from human milk. 32P postlabeling was performed using excess [32P]ATP to demonstrate the purity of the DNA preparations obtained using the PureGene kit with the modifications noted in “Materials and Methods.” P, inorganic phosphate; T, thymidine-3',5'-bisphosphate; C, deoxycytidine-3',5'-bisphosphate; A, deoxyadenosine-3',5'-bisphosphate; G, deoxyguanosine-3',5'-bisphosphate. A, chromatogram of 32P-postlabeled DNA sample obtained from enriched breast epithelial cells isolated from human milk. 32P postlabeling was performed as described in “Materials and Methods” on four DNA samples isolated from human breast epithelial DNAs. A representative chromatograph is shown demonstrating the utility of these samples for the detection of in vivo modified DNA. O, origin. c, representative PCR amplification of NAT2 using human genomic DNA obtained from enriched breast epithelial cell DNA obtained from human milk. A single 1092-bp band is amplified for NAT2 using the primer pair described in “Materials and Methods.” Lane 1, control DNA; Lane 2, sample DNA; Lane 3, 100-kb ladder.
may also be used to study associations between DNA adducts, environmental exposures, and polymorphisms in genes involved in metabolism of endogenous and exogenous compounds. Additionally, it should be possible to measure metabolites or markers of agents that may be breast carcinogens in the liquid portion of the milk and correlate those with polymorphisms and/or adducts in the epithelial cells. These cells may also be useful for the culturing and derivation of cell lines from healthy representative cells, as has been done in earlier studies (18–21).

Using the methodology previously described by Thompson and Smith (13), DECs were obtained from breast milk with >90% viability, routinely. However, it was noted that cell viability decreased rapidly with time in the absence of BSA or FCS during the enrichment procedure. The addition of BSA, however, increased protein contamination in the DNA isolation procedure and was subsequently removed from the final washing buffer. Additionally, it was noted that in a small percentage of samples (<10%), significant bacterial contamination was acquired during sample collection. These samples were not included in this data set. Future collection may require the addition of low doses of bacteriostatic agents in collection containers to reduce any contribution of DNA from bacterial contamination.

In addition to the modifications noted above for the isolation of epithelial cells from breast milk, it was concluded from these studies that the quality and quantity of DNA recovered in the isolation procedure should be determined by combining the strengths of both the absorbance ratio of A<sub>260</sub>/A<sub>280</sub> for protein contamination as well as the quantitative strength of normal nucleotide analysis. Furthermore, normal nucleotide analysis determines the extent of RNA contamination not determined spectrophotometrically, which can interfere with post-labeling studies of DNA adducts. By using both methodologies, we were able to modify the initial DNA isolation procedure to decrease both RNA and protein contamination levels in subsequent samples.

A possible drawback of the use of these cells is that cells in the lactating breast may differ from those in a non-nursing woman and that hormonal profiles may affect cell processes. It is also possible that factors related to milk production may affect the metabolism of carcinogens as well and that human mammary epithelial cells found in breast milk may, in fact, not represent normal processes within the breast. In that case, adduct levels in cells shed into breast milk may be higher or lower than those from the nonlactating breast. It does not seem likely, however, that lactation would result in the total absence of carcinogen-DNA adducts in breast epithelial cells if they are present in nonlactating cells, nor would it necessarily alter the relationship between genotypes and adduct outcome, because effects should be nondifferential by genotype. However, these are issues that should be kept in mind when interpreting results from studies using exfoliated DECs from human breast milk.

Acknowledgments

We thank Gloria Bruncato, R. N., who was instrumental to the success of this project, and Christine Grass for all of her efforts in enrolling participants and collecting specimens. We are particularly grateful to all of the generous mothers who donated not only their breast milk, but their time and energy, to participate in this study to advance our understanding of breast carcinogenesis. Thanks also to Dr. Maurizio Trevisan, who graciously extended resources in the Department of Social and Preventive Medicine of the State University of New York at Buffalo for the continuation of the study.

References

Exfoliated ductal epithelial cells in human breast milk: a source of target tissue DNA for molecular epidemiologic studies of breast cancer.

P A Thompson, F F Kadlubar, S M Vena, et al.


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/7/1/37