Changes in Gene Expression in Established Human Mammary Tumor Cell Lines When Compared with Normal Breast and Breast Tumor Tissue

Len Hall, Avril Henney, David N. L. Ralphs, David G. Herries, and Roger K. Craig

Cancer Research Campaign Endocrine Tumour Molecular Biology Group, Courtauld Institute of Biochemistry [L. H., A. H., R. K. C.], and Department of Surgical Studies [D. N. L. R.], The Middlesex Hospital Medical School, Mortimer Street, London W1P 7PN, and Department of Biochemistry, University of Leeds, Leeds LS2 9JT, [D. G. H.], United Kingdom

ABSTRACT

RNA complexity analyses of total cellular polyadenylate-containing RNA isolated from lactating human breast tissue, human breast tumor tissue, and a mixture of established cell lines of mammary origin demonstrate extensive homology between the tissue RNA populations but suggest a decrease in the complexity of cell line nuclear RNA populations, with the exception of an early-passage MCF-7 cell line. Cell-free protein synthesis and two-dimensional gel electrophoresis also show quantitative and qualitative differences in gene expression between human mammary tumor tissue and reduction mammoplasty or established mammary cell lines of early and late passage number.

The results (a) demonstrate a major role for transcriptional and posttranscriptional mechanisms in the regulation of gene expression in the human mammary gland and (b) show that studies on mammary gland gene expression using established cell lines of mammary origin reflect only in part gene expression in normal lactating human breast or breast tumor tissues.

INTRODUCTION

The biological behavior of mammary carcinoma varies greatly between individual patients. Consequently there is a need for reliable markers which indicate both the long term prognosis and the response to therapy. At the present time the best prognostic indicators are the extent and involvement of axillary lymph nodes (1). Knowledge of steroid-hormone receptor status of tumors is valuable for the prediction of response to endocrine therapy in advanced cancer of the breast and may have prognostic significance (2). Although the application of the techniques of molecular biology should provide the means to dissect mammary cancer at the molecular level, with the subsequent identification of gene products of diagnostic and prognostic significance, the application of molecular techniques has been limited thus far to the identification of estrogen-responsive gene products of unknown prognostic value from several cell lines of human mammary origin (3, 4). Recombinant techniques have been used to clone and characterize cDNA7 sequences encoding proteins of potential prognostic and diagnostic value. One cDNA sequence containing the coding sequence of human α-lactalbumin proved valuable as a hybridization probe only in a negative sense, indicating that α-lactalbumin was not a valid marker for estrogen responsiveness in human mammary tumors (see Refs. 5 and 6). The potential value of a second cDNA sequence, an estrogen-regulated gene product of the MCF-7 cell line, remains to be established (7–9).

Here we have compared gene expression in human lactating mammary tissue, human mammary tumor tissues, human mammary primary cell cultures, and established human cell lines of mammary origin. The results show: (a) using RNA complexity analysis a marked reduction in the overall number of transcribed poly(A)-containing RNA species in established cell lines, the decrease being particularly marked in the nuclear compartment; and (b) using cell-free protein synthesis and two-dimensional gel electrophoresis, quantitative and qualitative differences in the abundant and moderately abundant poly(A)-containing RNA species when comparing poly(A)-containing RNA from tumor and reduction mammoplasty, or an established mammary cell line of early and late passage number. Our observations reveal significant differences at the transcriptional and posttranscriptional level of gene expression between cells present in mammary tumors and cell lines of mammary origin in monolayer culture.

MATERIALS AND METHODS

Materials. [5-3H]dCTP (~20 Ci/mmole) and 1,5-[35S]methionine (~1000 Ci/mmole) were obtained from Amersham International, Pte, Amersham, United Kingdom. Tissue culture materials were from Gibco Bio-Cult Ltd., Paisley, Scotland, except for bovine insulin which was from Sigma Chemical Co., Poole, United Kingdom. Avian myeloblastosis virus reverse transcriptase was provided by Dr. J. W. Beard, Life Sciences Inc., St. Petersburg, FL. Hydroxyapatite was from Bio-Rad Laboratories, Ltd., Watford, United Kingdom. All other chemicals and solvents were obtained from sources described previously (10, 11).

Source of Human Mammary Tissue Samples. The tissue we describe as "normal" lactating breast (hN1) was obtained from a nursing mother, 15 days postpartum, and consisted of a large (2-g) retention cyst with virtually normal histology, showing lactating mammary tissue surrounded by areas of severe inflammation and abscess formation but no evidence of malignancy. Mammary tumor samples (hT1, hT2, hT9, hT33) were obtained from postmenopausal patients, none of whom had undergone either hormone therapy or chemotherapy prior to surgery. All tissue samples were deep-frozen in liquid nitrogen within minutes of excision and then stored at −70°C until required.

Human Mammary Cell Lines. The MCF-7 mammary cell lines (passages 12 and 257) were obtained from Dr. C. M. McGrath, Michigan Cancer Foundation; the Hs578T mammary cell line (late passage) was from Dr. A. Hackett, University of California; and the ZR75-1 mammary cell line and reduction mammoplasty cells were from Dr. J. Easty, Ludwig Institute, Sutton, London, United Kingdom. MCF-7 and Hs578T cell lines were grown in Eagle's minimal essential medium in Earle's balanced salt solution, nonessential amino acids, 1% (v/v) glucose, 11 mm sodium pyruvate, 2 mm L-glutamine, 10% (v/v) newborn calf serum, penicillin (100 units/ml), streptomycin (100 units/ml), and bovine insulin (1 μg/ml). In the case of the ZR75-1 cell line 50% of the Eagle's minimal essential medium was substituted by RPMI-1640.

Primary cultures of human reduction mammoplasty cells were grown for 2 weeks in RPMI 1640 containing 10% (v/v) fetal calf serum, cholera toxin (50 μg/ml), and bovine insulin (1 μg/ml). Frozen human

Received 8/2/85; revised 3/21/86; 7/10/86; accepted 8/5/86.

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.

The Cancer Research Campaign supported the bulk of this work.

1. The Cancer Research Campaign Endocrine Tumour Molecular Biology Group, Courtauld Institute of Biochemistry, University of Bristol, Bristol BS8 1TD, United Kingdom.

2. Recipient of a Studentship from the Medical Research Council.

3. Present address: Department of Biology, University of York, York YO1 5DD, United Kingdom.

4. To whom requests for reprints should be addressed.

5. The abbreviations used are: cDNA, complementary DNA; poly(A), polyadenylate.

5786
fibroblasts of mammary origin (hF1B-1; passage 4) were a gift from Dr. J. Easty.

Preparation of Total Poly(A)-containing RNA. Total cellular poly(A)-containing RNA was isolated from frozen human lactating mammary tissue or mammary carcinoma as described previously (5). For cell lines a modified procedure was adopted. About 15 Petri dishes (9 cm in diameter) of each cell type were grown until semiconfluent. The medium was then removed and cells were washed quickly in situ with phosphate-buffered saline. Lysis was brought about by the addition to each dish of 1 ml of 20 mM Tris-HCl, pH 7.6, containing 1 mM EDTA, 2% (w/v) sodium dodecyl sulfate, and proteinase K (250 μg/ml). After incubation at 37°C for 2 h, the lysates were pooled, extracted with phenol-chloroform, and processed as described for frozen tissue. Yields of poly(A)-containing RNA were typically 0.5–1% for mammary tissues and 1–2% for established cell lines.

Preparation of Nuclear and Postnuclear Poly(A)-containing RNA. Nuclear and postnuclear supernatant fractions of human cell lines were prepared in a manner similar to that described by Thomas et al. (12). Up to 60 dishes (9 cm in diameter) of cells were grown until semiconfluent and then processed in batches of 15 dishes in the following manner. Medium was aspirated and the plates were then placed on ice water. Cells on each plate were rapidly washed with 10 ml of ice-cold hypotonic buffer (5 mM Tris-HCl, pH 7.4, containing 1.5 mM KCl and 2.5 mM MgCl2) and then lysed on ice with 1.5 ml of hypotonic buffer containing 6 mM spermidine-HCl, 1% (v/v) Triton X-100, and 1.6% (w/v) sodium deoxycholate for 5 min. Lysates from 15 dishes were pooled after gentle pipetting to release nuclei from the plastic and centrifuged at 2500 × g for 5 min (0°C). The postnuclear supernatant was made 0.5% (w/v) with respect to sodium dodecyl sulfate and 100 μg/ml with respect to proteinase K and incubated at 45°C for 30 min. The sedimented nuclei were resuspended in 20 ml of hypotonic buffer containing 6 mM spermidine-HCl and 1% (v/v) Triton X-100, centrifuged as described above, and then resuspended in 5 ml of 20 mM Tris-HCl, pH 7.4, containing 1 mM EDTA. Sodium dodecyl sulfate and proteinase K were added to the lysates to concentrations of 2% (w/v) and 250 μg/ml, respectively and incubated at 45°C for 30 min. Nuclear and postnuclear preparations were then precipitated with ethanol, reprecipitated to remove traces of detergent, and incubated with iodoacetylated DNase (20 μg/ml) (13) for 20 min on ice in 10 mM Tris-HCl, pH 7.4, containing 10 mM NaCl and 10 mM magnesium acetate. Proteinase K and sodium dodecyl sulfate were then added to each and the mixtures were incubated as described above, then extracted with phenol-chloroform, and concentrated by ethanol precipitation. The poly(A)-containing RNA was then isolated from each preparation by two cycles of affinity chromatography on oligodeoxythymidylate-cellulose, as described previously (10).

Nuclear and postnuclear RNA preparations were obtained from frozen mammary tissue samples by a similar procedure. In this case the frozen tissue was first ground in liquid nitrogen using a mortar and pestle and then plunged into the lysis buffer and homogenized gently by hand to ensure efficient lysis. The lysate was then processed as described above.

Synthesis of cDNA and Fractionation into Abundance Groups. High specific radioactivity [3H]cDNA (1.4 × 10⁶ cpm/μg) complementary to poly(A)-containing RNA isolated from normal lactating mammary tissue, breast tumors, and cell lines was prepared as described by Craig et al. (11). Abundant and scarce-frequency cDNA populations, transcribed from MCF-7 nucleus poly(A)-containing RNA, were prepared by fractionation on hydroxyapatite as described by Bathurst et al. (14). A tumor tissue-specific cDNA probe (representing sequences not found in late-passage cell lines) was prepared in the same manner.

Hybridization Procedures. RNA excess hybridizations were performed at 68°C in 50 mM 4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid, pH 7.4, containing 0.6 mM NaCl, 1 mM EDTA, and 0.2% (w/v) sodium dodecyl sulfate. All other procedures were as described by Craig et al. (11), except that glass capillaries were boiled in 1 M HCl before the siliconizing treatment. Hybridizations were carried out for periods from 10 s and 116 h at RNA concentrations ranging from 2.5 μg/ml to 2.5 mg/ml. Whenever possible high RNA concentrations were used in preference to long hybridization times.

Numerical Analysis of Hybridization Kinetics. Hybridization curves resolved into a number of first-order components using a computer program developed for the least-squares analysis of hybridization kinetics (15). The correction coefficient of fit for all homologous hybridization curves was better than 0.998. Attempts to fit fewer components than those shown always produced significantly poorer correlation coefficients. The nucleotide complexity of each transition was calculated, based on the hybridization kinetics of a purified globin mRNA standard as described by Craig et al. (11). To estimate the number of different RNA species in each transition, the number average RNA size of each poly(A)-containing RNA preparation was determined by sucrose gradient centrifugation under denaturing conditions, followed by [3H]-polyuridylyl assays of gradient fractions, as described previously (11). In all cases a size distribution was obtained with a number average size of around 1350 nucleotides.

Cell-free Protein Synthesis and Product Analysis. Cell-free protein synthesis was performed in a nucleoside-treated reticulocyte cell-free system as described by Pelham and Jackson (16), while two-dimensional gel electrophoretic analysis of the in vitro synthesized [35S]methionine-labeled polypeptides was as described by O’Farrell (17). Fluorography was performed as described by Bonner and Laskey (18), except that acetic acid replaced dimethyl sulfoxide (19). Protein molecular weight markers electrophoresed in the second dimension were phosphorylase b (M, 92,000), bovine serum albumin (M, 69,000), ovalbumin (M, 46,000), carbonic anhydrase (M, 30,000), and lysozyme (M, 14,000).

RESULTS

Comparative RNA Complexity Analysis of Total Cellular, Nuclear, and Cytoplasmic Poly(A)-containing RNA Populations Isolated from Human Mammary Tissues, and Human Cell Lines of Mammary Origin. Comparative analysis of the nucleotide base sequence complexity of total cellular poly(A)-containing RNA isolated from lactating human mammary tissue, human mammary tumor tissue, and established human cell lines of mammary origin showed differences in terms of overall complexity. Examination of the hybridization kinetics of high specific activity [3H]cDNA prepared from poly(A)-containing RNA isolated from the lactating human mammary gland by hybridization against an excess of its homologous RNA produced a curve which could be resolved into four first-order components, hybridizing over six logarithmic decades. Calculation of the base sequence complexity within each component by comparison with the kinetics of a highly purified rabbit globin poly(A)-containing RNA standard (Fig. 1A), demonstrated (on the basis of a number average molecular weight of 1350 nucleotides) that the most rapidly hybridizing or very abundant class contained 2–3 sequences, the abundant class about 40 sequences, the intermediate class about 1800 sequences, and the scarce class about 30,000 sequences (see Table 1). Analysis of the kinetics of hybridization of poly(A)-containing RNA populations isolated from two mammary tumors (hT2 and hT33) in the same manner produced curves which could be resolved into three components (Fig. 1A; Table 1). In each instance the very abundant poly(A)-containing RNA class, which we have previously shown to be representative of the human milk protein mRNA species (5, 6, 20), was absent, but the overall complexity and distribution of sequences within the three overall abundance classes were very similar to those of the abundant, intermediate, and scarce-frequency classes identified in poly(A)-containing RNA isolated from lactating mammary tissue.

In contrast to the results obtained for tissue poly(A)-containing RNA populations, similar analyses of total poly(A)-contain-
Fig. 1. Homologous and heterologous RNA complexity analyses of total cellular, nuclear, and cytoplasmic poly(A)-containing RNA isolated from human mammary tissue and cell lines. RNA excess hybridization procedure and data analysis were as described in "Materials and Methods." A, hybridization of total cellular poly(A)-containing RNA derived from human lactating tissue (hN1) and human mammary tumor (hT2) tissue (hir) with their homologous cDNAs. Hybridization of highly purified rabbit globin mRNA to its cDNA (h) provided a kinetic standard. B, hybridization of total cellular poly(A)-containing RNA derived from early- (hir) and late- (hT3) human mammary tumor (hT2) tissue (h) with their homologous cDNAs. C, heterologous hybridization between human mammary tumor (hT2) tissue (hir) and human lactating tissue (hN1) cDNA. D, heterologous hybridizations between human mammary fibroblast alone (hir) or fibroblast plus MCF-7 cell line (hir) total poly(A)-containing RNA and human lactating tissue (hN1) cDNA. Arrows, corrected Rofw values for each first-order component.

Table 1 Complexity of RNA populations derived from human lactating mammary tissue, mammary tumors, and mammary cell lines

<table>
<thead>
<tr>
<th>Source of RNA</th>
<th>Total cellular polyadenylated RNA</th>
<th>Postnuclear polyadenylated RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammary tumor tissue (hT2)</td>
<td>2 ± 1*</td>
<td>47 ± 9</td>
</tr>
<tr>
<td>Mammary tumor tissue (hT3)</td>
<td>25 ± 6</td>
<td>25 ± 6</td>
</tr>
<tr>
<td>Mammary tumor tissue (hT33)</td>
<td>18 ± 3</td>
<td>24 ± 5</td>
</tr>
<tr>
<td>MCF-7 (early passage)</td>
<td>47 ± 9</td>
<td>47 ± 9</td>
</tr>
<tr>
<td>MCF-7 (late passage)</td>
<td>25 ± 6</td>
<td>25 ± 6</td>
</tr>
<tr>
<td>Hs578T</td>
<td>29 ± 8</td>
<td>29 ± 8</td>
</tr>
<tr>
<td>hFIB-1</td>
<td>24 ± 5</td>
<td>24 ± 5</td>
</tr>
<tr>
<td>Reduction mammoplasty</td>
<td>48 ± 9</td>
<td>48 ± 9</td>
</tr>
<tr>
<td>Lactating tissue (hN1)</td>
<td>2 ± 1*</td>
<td>47 ± 9</td>
</tr>
<tr>
<td>Mammary tumor tissue (hT1)</td>
<td>17 ± 3</td>
<td>21 ± 6</td>
</tr>
<tr>
<td>Mammary tumor tissue (hT2)</td>
<td>25 ± 6</td>
<td>29 ± 8</td>
</tr>
<tr>
<td>MCF-7 (early passage)</td>
<td>47 ± 9</td>
<td>47 ± 9</td>
</tr>
<tr>
<td>MCF-7 (late passage)</td>
<td>25 ± 6</td>
<td>25 ± 6</td>
</tr>
<tr>
<td>Hs578T</td>
<td>29 ± 8</td>
<td>29 ± 8</td>
</tr>
<tr>
<td>hFIB-1</td>
<td>24 ± 5</td>
<td>24 ± 5</td>
</tr>
<tr>
<td>Reduction mammoplasty</td>
<td>48 ± 9</td>
<td>48 ± 9</td>
</tr>
</tbody>
</table>

*Mean ± SD.

Calculations of the relative number of RNA species present in different abundance groups reflect the RNA complexity (number of nucleotides) divided by the number average molecular weight (1350 nucleotides); see "Materials and Methods." The standard derivatives of these values are the square roots of their variances. The significance of differences between pairs of values (see text) may be determined by calculating

\[
\text{Value 1} - \text{Value 2} \div \sqrt{\text{Variance}_1 + \text{Variance}_2}
\]

where variance is the square of standard deviation. The P value may then be determined by reference to probability distribution tables.

No. of different RNA species

<table>
<thead>
<tr>
<th>Total cellular polyadenylated RNA</th>
<th>Postnuclear polyadenylated RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very abundant</td>
<td>Abundant</td>
</tr>
<tr>
<td>Mammary tumor tissue (hT2)</td>
<td>2 ± 1*</td>
</tr>
<tr>
<td>Mammary tumor tissue (hT3)</td>
<td>25 ± 6</td>
</tr>
<tr>
<td>Mammary tumor tissue (hT33)</td>
<td>18 ± 3</td>
</tr>
<tr>
<td>MCF-7 (early passage)</td>
<td>47 ± 9</td>
</tr>
<tr>
<td>MCF-7 (late passage)</td>
<td>25 ± 6</td>
</tr>
<tr>
<td>Hs578T</td>
<td>29 ± 8</td>
</tr>
<tr>
<td>hFIB-1</td>
<td>24 ± 5</td>
</tr>
<tr>
<td>Reduction mammoplasty</td>
<td>48 ± 9</td>
</tr>
<tr>
<td>Lactating tissue (hN1)</td>
<td>2 ± 1*</td>
</tr>
<tr>
<td>Mammary tumor tissue (hT1)</td>
<td>17 ± 3</td>
</tr>
<tr>
<td>Mammary tumor tissue (hT2)</td>
<td>25 ± 6</td>
</tr>
<tr>
<td>MCF-7 (early passage)</td>
<td>47 ± 9</td>
</tr>
<tr>
<td>MCF-7 (late passage)</td>
<td>25 ± 6</td>
</tr>
<tr>
<td>Hs578T</td>
<td>29 ± 8</td>
</tr>
<tr>
<td>hFIB-1</td>
<td>24 ± 5</td>
</tr>
<tr>
<td>Reduction mammoplasty</td>
<td>48 ± 9</td>
</tr>
</tbody>
</table>

*Mean ± SD.

ing RNA from two well-established human mammary tumor cell lines (MCF-7 passage 275 and Hs578T) showed for each a three-component curve of much reduced overall base sequence complexity (Table 1; see also Fig. 1B). This was particularly apparent in the intermediate and scarce-frequency classes where the number of different poly(A)-containing RNA species appeared to be reduced by as much as 50%. However, analysis of very-early-passage MCF-7 cells (passage 17) revealed a distribution of sequence complexity indistinguishable from the tissue samples (Table 1; Fig. 1B), and consequently of greater overall complexity than that present in MCF-7 late-passage cells (Table 1; Fig. 1B). Analysis of total poly(A)-containing RNA from a human mammary fibroblast cell line (hF1B-1) in early passage (passage 4) and reduction mammoplasty epithelial cells in primary culture also revealed a three-component curve and a nucleotide base sequence complexity midway between the mammary tissues and late passage cell lines, the reduction in complexity being confined to the scarce frequency class (see Table 1).

In order to determine within the limitations of the experimental procedure whether poly(A)-containing RNA sequences expressed in breast tumor tissue were common to those expressed in normal lactating tissue and also whether sequences present in the cell lines were also present in normal tissue, we have performed a series of heterologous cross-hybridization experiments using cDNA prepared from the normal lactating tissue (hN1), total cellular poly(A)-containing RNA isolated from mammary tumor tissue (hT2), and late-passage MCF-7 cells. Due to a scarcity of poly(A)-containing RNA originating from normal lactating tissue, the reciprocal crosses were not possible. However, the results (Fig. 1C) demonstrate that the majority of sequences present in the lactating tissue cDNA could form stable hybrids with the breast tumor poly(A)-containing RNA since hybridization reached 82% (Fig. 1C), a value 5–10% lower than the corresponding homologous hybridization. This difference will reflect to a very great extent the
absence in tumor tissue of the very abundant class of poly(A)-containing RNA representing the milk protein RNA transcripts, as we have shown elsewhere that one of these, encoding α-lactalbumin, is absent from breast tumor tissue (5). Consequently, with the exception of these very abundant sequences, the normal and malignant breast tissue total cellular poly(A)-containing RNA populations exhibit a high degree of sequence homology. Heterologous hybridization between total cellular poly(A)-containing RNA and cDNA from two different mammary tumors also revealed extensive sequence homology, indistinguishable from the corresponding homologous hybridizations (data not shown).

Consistent with results obtained from nucleotide complexity analysis (Table 1), total poly(A)-containing RNA from late-passage MCF-7 (Fig. 1C) or Hs578T cell lines (data not shown) hybridized to only 60–65% of those sequences present in cDNA synthesized from the lactating tissue poly(A)-containing RNA population. Thus a large proportion of the poly(A)-containing RNA sequences present in the normal lactating tissue, and by inference tumor tissue, appear not to be expressed in the equivalent poly(A)-containing RNA populations in late-passage cell lines. Similar analyses using poly(A)-containing RNA from the mammary fibroblast cell line (hF1B-1) resulted in hybridization to 72% of the lactating tissue cDNA, a value increased by only 1–2% when an equal mixture of late-passage MCF-7 and hF1B-1 poly(A)-containing RNA was used (Fig. 1D). This result, in keeping with the complexity of the early-passage MCF-7 cell line, suggests that the greater complexity of the mammary tissue poly(A)-containing RNA population cannot be explained simply by the presence of different cell types within mammary tissues, each with qualitatively different patterns of expression thereby accounting for the increased RNA complexity observed. Additional heterologous cross-hybridizations between poly(A)-containing RNA and cDNA from late-passage MCF-7 cells and Hs578T cells revealed extensive homology; in this instance hybridizations were only 5% less than the corresponding homologous hybridizations (data not shown). Thus two independently established mammary tumor cell lines show reduced complexity when compared with tissue poly(A)-containing RNA populations and express in qualitative terms similar poly(A)-containing RNA populations.

Nuclear and Cytoplasmic Distribution of Poly(A)-containing RNA in Mammary Tumors and Cell Lines of Mammary Origin. Poly(A)-containing RNA was isolated from nuclear and postnuclear preparations of a breast tumor (hT33) and several cell lines of mammary origin (late-passage MCF-7, Hs578T, ZR75-1), cDNA was prepared from each, and the kinetics of hybridization of each was determined against their homologous RNA populations. Analysis of the kinetics of hybridization of postnuclear preparations from the late passage MCF-7 cell line and tumor tissue (hT33) revealed three-component hybridization curves (Fig. 2). In each instance the first two transitions were similar in RNA complexity to those observed in the corresponding analyses of total cellular poly(A)-containing RNA (see Table 1). The postnuclear scarce-frequency class, however, showed a significantly reduced complexity (P < 0.001) when compared with the total cellular population for both tissue (8,800 against 26,400) sequences and cell line (7,000 against 15,700) sequences. Similar analyses of nuclear poly(A)-containing RNA populations when hybridized against their homologous cDNA preparations exhibited, as expected, markedly different kinetics of hybridization to either total or postnuclear poly(A)-containing RNA preparations. In this instance, two component curves were obtained by analysis of tumor tissue and MCF-7 cell line nuclear poly(A)-containing RNA populations (Fig. 2). The abundant class in tissue and cell lines had similar complexities representative of 40–50 poly(A)-containing RNA species (Table 1), but the scarce class of nuclear sequences differed markedly in complexity. Thus there were about 20,000 nuclear scarce-frequency poly(A)-containing RNA species present in the late-passage MCF-7 cell line compared with 28,000 in the equivalent tumor population. Analysis of nuclear and postnuclear poly(A)-containing RNA populations obtained from two other late-passage mammary tumor cell lines, Hs578T and ZR75-1, produced a distribution of nuclear and postnuclear poly(A)-containing RNA sequences very similar to that observed for the late-passage MCF-7 cells (see Table 1).

To examine in greater detail the relative distribution of the cell line nuclear poly(A)-containing RNA abundance classes within the postnuclear poly(A)-containing RNA population, cDNA was synthesized from late-passage MCF-7 nuclear poly(A)-containing RNA, then fractionated into the two nuclear abundance groups by hybridization to the homologous RNA population to a Rot of 5.6 mol · liter⁻¹ · s⁻¹ followed by separation of the single-stranded scarce-frequency cDNA from the hybridized abundant cDNA species on hydroxyapatite. Hybridization of the individual purified cDNA populations to MCF-7 nuclear poly(A)-containing RNA demonstrated (Fig. 3) that both hybridized with the expected kinetics, with little evidence of cross-contamination between the abundance groups. Examination of the distribution of each nuclear cDNA population within the postnuclear poly(A)-containing RNA population by heterologous cross-hybridization produced contrasting results. The abundant nuclear cDNA hybridized essentially to completion over 2.5 logarithmic decades, with kinetics similar to that of the abundant postnuclear poly(A)-containing RNA population. The scarce-frequency nuclear population, however, did not hybridize to completion, hybridization occurring over 3.5 logarithmic decades and reaching a value of about 60% relative to the corresponding homologous hybridization at high Rot values (see Fig. 3). Although hybridization occurred predominantly with the expected kinetics of scarce-frequency postnuclear poly(A)-containing RNA species, low yet significant hybridization...
zation also occurred with kinetics similar to the intermediate postnuclear poly(A)-containing RNA population. However, in spite of this observation, the results clearly demonstrate that as much as 40% of the scarce nuclear poly(A)-containing RNA species were specifically confined to the nucleus and that the remainder constituted predominantly the scarce frequency postnuclear poly(A)-containing RNA class. A series of heterologous hybridizations using the MCF-7 abundant- and scarce nuclear cDNA probes driven with nuclear and postnuclear poly(A)-containing RNA isolated from Hs578T and ZR75-1 cell lines were also performed, producing kinetics of hybridization essentially identical to those described for the late-passage MCF-7 cells (results not shown).

The results presented above implicate posttranscriptional events (see "Discussion") and provide further evidence that the disparity in total complexity between tissue and late-passage cell line poly(A)-containing RNA populations resides in the nuclear fraction, since subcellular fractionation did not reveal major qualitative disparities between the tumor or cell line postnuclear poly(A)-containing RNA populations. In order to establish that a significant subset of the nuclear polyadenylated RNA sequences expressed in tumor tissue were not expressed in the corresponding late-passage cell line populations, a series of heterologous cross-hybridizations were performed between cDNA populations synthesized from tumor (hT33) postnuclear and nuclear poly(A)-containing RNA, and late-passage MCF-7 postnuclear and nuclear poly(A)-containing RNA populations. This demonstrated (Fig. 4A) that the kinetics and degree of hybridization of MCF-7 postnuclear poly(A)-containing RNA to the equivalent tumor cDNA preparation were strikingly similar to those of the tumor cDNA with its homologous postnuclear poly(A)-containing RNA population but that the analogous nuclear cross-hybridization resulted in only 70% protection of the tumor nuclear cDNA population by the MCF-7 nuclear poly(A)-containing RNA population, when compared with the tumor nuclear homologous hybridization (see also Fig. 2B). To further establish whether the differences reflect the absence of tumor scarce-frequency nuclear sequences from the late-passage MCF-7 nuclear poly(A)-containing RNA population, a tumor tissue-specific cDNA population was generated by hybridization of late-passage MCF-7 total cellular poly(A)-containing RNA with cDNA derived from breast tumor (hT33) total poly(A)-containing RNA to a Rd of 316 mol-liter⁻¹·s⁻¹, followed by recovery of the "tumor-specific" unhybridized single-stranded cDNA on hydroxyapatite. Hybridization analyses with the tumor tissue-specific cDNA probe revealed (Fig. 4B) negligible hybridization to the late-passage MCF-7 total cellular poly(A)-containing RNA, slight hybridization to tumor tissue postnuclear poly(A)-containing RNA, but substantial hybridization to tumor tissue nuclear poly(A)-containing RNA, with the expected kinetics of the scarce-frequency class of nuclear sequences. These results demonstrate that a subset of the least abundant "tumor tissue" and by analogy "normal tissue" nuclear poly(A)-containing RNA sequences (see Table 1 and Fig. 1C) are not expressed (or expressed at a very much reduced level) in mammary tumor cell lines after prolonged culture in vitro.

Differential Expression of Poly(A)-containing RNA Sequences in Mammary Tissues and Cell Lines as Determined by mRNA-directed Cell-free Translation Systems. Quantitative and qualitative changes in individual gene expression were investigated in tissue and cell line poly(A)-containing RNA preparations using a mRNA-directed reticulocyte lysate cell-free translation system and the distribution of in vitro synthesized [35S]methionine-labeled proteins examined by two-dimensional gel electrophoresis and fluorography. The results revealed contrasting distributions of [35S]-labeled proteins when translation products directed by tumor tissue and reduction mammaryplast, or MCF-7 early- and late-passage poly(A)-containing RNA preparations were compared. In all instances (Fig. 5) a wide spectrum of [35S]-labeled proteins were synthesized ranging in molecular weight from 14,000 to in excess of 100,000, with the majority of the proteins focusing within a pi range of 4.5-7.3. Comparison of tumor (hT9) and reduction mammaryplast cell-free products revealed a very different pattern of predominant [35S]-labeled proteins (Fig. 5, A and B). Some of the differences observed between the two populations were quantitative in nature, but in general significant seemingly qualitative differences were also apparent, within the abundant and less abundant sequence.
CHANGES IN HUMAN MAMMARY GENE EXPRESSION

Fig. 5. Differential gene expression in human mammary tumor tissue, reduction mammoplasty, and cell lines of mammary origin. Total poly(A)-containing RNA was translated in a nuclease-treated reticulocyte-lysate cell-free system in the presence of \(^{35}\)S-methionine, and the individual translation products (20,000–30,000 cpm) were identified by two-dimensional polyacrylamide gel electrophoresis and fluorography. Arrows, position of the major endogenous reticulocyte protein; molecular weight markers were as described in “Materials and Methods.” A, mammary tumor hT9; B, reduction mammoplasty; C, MCF-7 cell line early passage; D, MCF-7 cell line late passage.

The advent of recombinant DNA technology and the subsequent use of sequence-specific DNA hybridization probes has in recent years eclipsed the value of RNA complexity analyses using liquid hybridization techniques. To a large extent this has been accelerated by an increased awareness of the limitations of complexity analysis, particularly in terms of quantitation. Hybridizations in solution rarely go to completion and cDNA probes, if not prepared under optimum conditions and carefully characterized, tend to be more representative of the 3' ends of gene expression.

classes within the two poly(A)-containing RNA populations examined.

The comparison of MCF-7 early- and MCF-7 late-passage cell-free translation products also revealed differences in the major \(^{35}\)S-methionine-labeled protein products (Fig. 5, C and D). In particular a number of abundant proteins (M, 60,000, pl 6.0; M, 44,000, pl 5.3; M, 18,000, pl 5.6; M, 13,000, pl 6.0) encoded by MCF-7\(_{\text{early}}\) poly(A)-containing RNA were not directed in detectable amounts by the MCF-7\(_{\text{late}}\) poly(A)-containing RNA preparation. Conversely a number of predominant gene products (M, 64,000, pl 5.3; M, 47,000, pl 6.9; M, 44,000, pl 4.6; M, 43,000, pl 4.6; M, 24,000, pl 4.6) were directed by poly(A)-containing RNA from MCF-7\(_{\text{late}}\) as opposed to MCF-7\(_{\text{early}}\) cells. Differential gene expression was also apparent in early- and late-passage MCF-7 cell-free translation products within the putative actin gene cluster (M, 46,000, pl 5.6–5.8).

DISCUSSION

The advent of recombinant DNA technology and the subsequent use of sequence-specific DNA hybridization probes has in recent years eclipsed the value of RNA complexity analyses using liquid hybridization techniques. To a large extent this has been accelerated by an increased awareness of the limitations of complexity analysis, particularly in terms of quantitation. Hybridizations in solution rarely go to completion and cDNA probes, if not prepared under optimum conditions and carefully characterized, tend to be more representative of the 3' ends of gene expression.

Downloaded from cancerres.aacrjournals.org on April 15, 2017. © 1986 American Association for Cancer Research.
the RNA templates. Both of these technical drawbacks lead to uncertainty if accurate quantitation is required. In addition, problems can arise through numerical presentation of data. Thus whereas nucleotide complexity of RNA populations may be considered as absolute values when compared with a common standard (globin mRNA), calculations of the numbers of RNA species present in different RNA populations illustrate relative differences only. This is because the calculation of the number of different RNA species is based on the number average weight of each RNA population, a value which may differ from tissue to tissue and which is inevitably biased by the presence of a small number of relatively abundant RNA species. Taking into account such variables, it is hardly surprising that studies with cloned, sequence-specific probes, which tend not to suffer from these limitations, are proving more popular. However, it is rather naive to think that careful analyses with just a handful of these specific, cloned probes alone will allow us to make general statements about the overall control of the many thousands of RNA transcripts found in a mammalian cell at any given time. This led us to reappraise the value of RNA complexity analysis and apply the technique in such a way that any conclusions we made were not solely dependent on the accuracy of quantitation. Consequently, whenever a difference was found between cell types or cell compartments, in terms of calculated RNA complexity (or in this instance RNA species, since a number average molecular weight of 1350 nucleotides was used in all calculations), these differences were further confirmed by appropriate heterologous hybridizations, or hybridization with individually purified components. In this way most of the above uncertainties can be overcome or minimized.

We have previously used hybridization analysis to investigate the relative complexity of nuclear and cytoplasmic poly(A)-containing RNA populations in lactating guinea pig tissues. Such studies have demonstrated a major role for posttranscriptional events in the preferential accumulation of transcribed sequences in the cytoplasm, and the tissue-specific distribution of transcribed nuclear poly(A)-containing RNA sequences between nucleus and cytoplasm (14, 21). Our studies on nuclear and cytoplasmic poly(A)-containing RNA populations from human mammary tumor tissue and a number of established mammary cell lines of tumor origin extend these observations. In common with the animal model system, in normal lactating human breast tissue, in breast tumor tissues, and in established cell lines of mammary origin, posttranscriptional mechanisms operate which determine the preferential accumulation and relative abundance of transcribed sequences in the cytoplasm. Such studies are clearly dependent on the purity of nuclear and postnuclear preparations, and for this reason the isolation of mRNA from purified polysomes may seem preferable. However, while this is technically possible for the established cell lines, preparation of polysomes in reasonable yield from human tumor tissue of this type is notoriously difficult and could lead to selective losses of particular sequences. We therefore chose to adopt a method based on that of Thomas et al. (12) which results in a nuclear and postnuclear RNA preparation in good yield and which, on the basis of homologous and heterologous hybridization experiments, results in minimal cross-contamination. However, even slight contamination of the less complex postnuclear population with the more complex nuclear population could significantly affect the observed complexities of these two fractions, as given in Table 1, and such values should be taken as an indication only. Nevertheless, such uncertainty would not affect our overall conclusion that posttranscriptional control mechanisms play an important part in mammary cell yields and which, on the basis of homologous and heterologous hybridization experiments, results in minimal cross-contamination. However, even slight contamination of the less complex postnuclear population with the more complex nuclear population could significantly affect the observed complexities of these two fractions, as given in Table 1, and such values should be taken as an indication only. Nevertheless, such uncertainty would not affect our overall conclusion that posttranscriptional control mechanisms play an important part in mammary cell lines and tissues since (a) the complexity of profiles for nuclear and postnuclear populations are clearly very different, with a different number of components, and (b) the observation that as much as 40% of the scarce nuclear sequences are not expressed in the cytoplasm must be taken as a minimum estimate since nuclear contamination of the postnuclear fraction would reduce this value. In general, therefore, cross-contamination between nuclear and postnuclear fractions would tend to decrease the observed differences between these two cell compartments and hence underestimate the role played by posttranslational mechanisms.

The similarity between the total complexity of poly(A)-containing RNA in normal lactating breast and breast tumor tissue is in agreement with studies on a number of experimental animal model systems, in particular liver, and cell lines transformed by viruses or carcinogens using total, cytoplasmic, or nuclear polyadenylated RNA (22–25). These studies also demonstrated little variation in the overall complexity of poly(A)-containing RNA from normal and transformed tissues and cell lines, although, where analyzed, differences in the relative abundance of sequences or groups of sequences were observed (24, 25). Our preliminary studies using mRNA-directed cell-free protein synthesis and two-dimensional sodium dodecyl sulfate-polyacrylamide gel analysis of the labeled products support the view that differences in the relative abundance of some sequences exist between tumor tissue and “normal” reduction mammoplasty cells, as well as between early- and late-passage MCF-7 cells.

It is apparent that the complexity of total cellular poly(A)-containing RNA in an early-passage clonal cell line (MCF-7 norm) is within experimental error the same as the total cellular poly(A)-containing RNA present in normal lactating breast or tumor tissues, each containing many cell types, the former predominantly epithelial (see Ref. 20), but the latter a polyglot of cell types. This suggests that within mammary tissue a remarkably similar spectrum of poly(A)-containing sequences is transcribed in all cell types and that the greater complexity of total sequences in tissues, when compared with established cell lines, cannot be explained simply by the expression of different sets of genes in the different cell types. This conclusion is supported by additive experiments where no significant increase in overall complexity was observed using mixtures of poly(A)-containing RNA isolated from primary mammary fibroblast cultures and late-passage MCF-7 cells. Thus, within experimental limits, in human mammary tissue a similar spectrum of genes (excluding the major lactation-specific milk proteins; see Refs. 6 and 20) is expressed, irrespective of cell type and irrespective of whether the RNA is derived from normal or tumor tissue. Such observations argue that in mammary tissue, posttranscriptional mechanisms play a major role in the determination of cell phenotype, a conclusion in keeping with early comparative work using sea urchins, which showed that transcribed single-copy nuclear sequences are common to embryo and adult tissues but that cytoplasmic mRNA populations differ (26–28).

The apparent decrease in the number of transcribed nuclear polyadenylated sequence during long-term culture and, in the case of the early and late MCF-7 cell lines, an associated change in relative abundance of the major mRNA species, may also have important implications. Numerous comparisons have been made between the relative complexities of polysomal poly(A)-containing RNA species from cell lines and normal and tumor-derived tissues, generally resulting in reports of qualitative as well as significant quantitative differences, particularly in the
abundant mRNA species (29–32). Relatively few studies have compared nuclear and cytoplasmic poly(A)-containing RNA from normal and tumor tissue, as well as cell lines of similar origin. Fausto et al. (25) working with normal, regenerating, and neoplastic rat liver reported quantitative, as opposed to qualitative, changes in nuclear polyadenylated RNA sequences, while others, comparing normal rat liver with transplantable hepatoma preparations, reported the loss of a substantial proportion of nuclear sequences in the hepatoma when compared with normal liver (33–34).

It is our opinion that the difference in transcribed nuclear sequences, seen in the cell lines and tissues in question, may reflect conditions of cell culture and hence cell shape. Recent studies on the growth and differentiation of normal mammary epithelial cells in culture have demonstrated the importance of the extracellular matrix and the effect of cell shape on the maintenance of their differentiated state (35–38). The extracellular matrix has also been shown to be crucial for the sustained growth and three-dimensional organization of primary mammary epithelial cells (39), and for the reconstitution of branching tubule structures from tumor cell lines in culture (40). From our data it is apparent that even after short-term culture a significant loss of scarce-frequency transcribed sequences occurs (compare reduction mammoplasty or mammary fibroblasts vs. mammary epithelial cells in culture). These showed an equivalence of complexity, whereas a reduced complexity was observed in primary glial cell cultures, glioma, and neuroblastoma preparations (41). By analogy with our own data we would predict that this reduced complexity reflects the preferential loss of nuclear sequences.

In conclusion, our studies (a) raise interesting questions as to the role of posttranscriptional events in the regulation of gene expression and (b) show that studies on mammary gland gene expression using established cell lines may only partly reflect expression in breast tissues. Such observations should be borne in mind when using human mammary tumor cell lines grown in monolayer culture, as model systems for the study of human breast cancer, particularly when designing strategies for the identification and isolation of gene probes of potential diagnostic or prognostic value. A better approach might be to identify genes expressed only in tumors as opposed to normal breast tissue, when using defined probes return to established cell lines to determine whether or not they are expressed and ultimately what factors modulate their expression.

ACKNOWLEDGMENTS

We are indebted to Dr. Jerry Easty, Ludwig Institute of Cancer Research, Sutton, United Kingdom, for preparing and providing the human mammary fibroblast and reduction mammoplasty primary cell cultures.

REFERENCES

32. Medeiros, J., and Harel, J. Sequence homology between polyadenylated mRNAs from transplantable mouse tumors of different origin, as compared...
CHANGES IN HUMAN MAMMARY GENE EXPRESSION

Changes in Gene Expression in Established Human Mammary Tumor Cell Lines When Compared with Normal Breast and Breast Tumor Tissue

Len Hall, Avril Henney, David N. L. Ralphs, et al.


Updated version  Access the most recent version of this article at: [http://cancerres.aacrjournals.org/content/46/11/5786](http://cancerres.aacrjournals.org/content/46/11/5786)

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.