IDENTIFICATION OF MAMMARY GLAND INVOLUTION DIFFERENTIALLY EXPRESSED GENES IN BREAST CANCER TISSUES*

Timothy Erickson¹, Brian D. Maciejko¹, Chris Gee¹, Victoria Lam¹, Denise Ng¹, Anne Blanchard², Yvonne Myal², Micah A. Chrenek¹*, Paul Wong¹†.

¹Department of Biological Sciences, University of Alberta, Edmonton, Canada T6G 2E9
²Department of Pathology, University of Manitoba, Canada

+Co-first authors
*Co-last authors
†Corresponding Author email pwwong@ualberta.ca

Research Support: This research was supported by grants from NSERC (PW) and (YM); The Alberta Cancer Board (PW). Scholarship support was provided by NSERC (TE, CG, VL, DN, MAC), AHFMR (TE, CG, DN), the Alberta Cancer Board (VL, PW), the Walter H. Johns Graduate Fellowship program (MAC), the Province of Alberta Graduate Fellowship program (MAC), and the University of Alberta (MAC, BDM)

Received December 1, 2003, Accepted January 29, 2004

The female mouse mammary gland (MG) undergoes a complex postnatal developmental program that involves a cyclic progression of the tissue from the virgin state, through pregnancy, lactation, involution, and back to a virgin-like state. This tissue progression is defined by regulated cycles of proliferation and apoptosis. Cells that evade apoptosis may become precursors to cancerous cells. Genes that are expressed during MG involution and in cancerous states may represent those genes that help predispose cells to evade proper regulation during involution. PCR products from a day 1 involution (I1) MG cDNA library were arrayed and screened with cDNA from I1 and I3 mouse MG. Clones representing putative differentially expressed genes were subjected to sequence and BLAST analyses. Northern and in silico analyses were performed on seven clones to profile tissue expression patterns and identify their human orthologues. The majority of the genes analyzed represented those involved in iron metabolism or fatty acid metabolism/regulation. Results of Northern analysis showed that all seven genes were differentially expressed during postnatal mouse MG development, thus confirming the effectiveness of our screen. In silico and literature mining analyses of the mouse and human orthologues revealed altered expression patterns in normal versus cancerous tissues for all seven genes, and that the PPAR family of transcription factors may play a role in regulating expression of these genes. We propose that by employing a combination of experimental and bioinformatic methods, examination of gene expression during normal mouse MG development can lead to the identification of human genes that may play a role in the initiation or propagation of breast cancer.

Key Words: apoptosis, gene expression, involution, mammary gland

Introduction

The mammary gland (MG) undergoes drastic post-natal developmental changes to first endow the virgin, non-lactating gland of females with the cellular apparatus required for the synthesis of milk, and then to dismantle that apparatus when milk synthesis is no longer required. This process of cell proliferation followed by apoptotic remodelling of the gland is repeated with each successive pregnancy. The cyclical, post-natal developmental program of the MG makes this tissue amenable to the study of non-embryonic, developmental genetics. However, the reasons for studying the MG surpasses the need to uncover the molecular and genetic mechanisms by which this tissue grows and regresses. Breast cancer is the second most deadly cancer amongst North American women.¹²
Current opinion favors the view that misregulation of the cell cycle, cell differentiation and programmed cell death are prime reasons for the growth of pre-oncogenic neoplasias. By defining the normal progressive molecular environment during MG development (cellular quiescence, proliferation, apoptosis, quiescence), it is hoped that key regulatory factors can be identified and used to define points where the induction of cell proliferation can become de-regulated and the signals for apoptosis evaded. Information concerning any or all of these normal developmental factors may provide insight into the pathogenesis of breast cancer. Sufficient similarities exist between mice and humans with regards to the pathogenesis, histology, cytology and genetics of MG cancer that researchers have been able to gather experimental data from the mouse MG model and make direct human comparisons.

Under control of hormones released during pregnancy, the milk-producing mammary epithelial cells (MECs) proliferate and supersede adipocytes as the prevailing cell type in the lactating MG. However, sudden weaning of the pups causes the gland’s secretory epithelium to go through two distinct stages of apoptosis that result in the rapid remodelling of the gland. Stage one (1-3 days post-weaning; I1-3) is reversible back to a lactating state and involves MG-derived signals for the targeted cell death of individual MEC. Stage one of MG involution is defined by changes in Ap-1, Nf1, Nfcb, Oct-1, p53, Stat3, and Stat5 transcription factor levels with a concomitant upregulation of apoptosis-related genes such as clusterin (Clu, Sgp-2), Fas antigen, Bcl-x, and Bax. A few morphologically isolated MECs undergo apoptosis, but the structural integrity of the MG ducts is maintained. Stage two MG involution (3-10 days post-weaning; I3-10) is marked by the up-regulation of matrix metalloproteinases, irreversible proteolytic degradation of the MG basement membrane, a dramatic increase in the incidence of MEC apoptotic death, and complete remodelling of the gland back to a virgin-like state. As the gland approaches a virgin-like physiology by day 9 of involution, adipocytes have again become the predominant cell type.

The one-day post-weaning transformation from a molecular environment favoring MEC survival to one of apoptotic susceptibility is one of the points at which abnormal execution of the MG developmental program could lead to oncogenesis. We sought to identify genes whose mRNA expression is differentially regulated at the I1 time point in order to elucidate the genetic environment during this critical transitional period between lactation and irreversible involution. To do this, we have performed a differential screen of a cDNA library derived from mouse MG I1 RNA. In our initial screen, we identified over 100 clones representative of potential differentially expressed genes and we have initiated the process of sequencing these genes. From our preliminary sequence analysis we found that a number of the isolated MG I1 clones represented genes that could be grouped into two distinct functional categories. Four of these genes (fatty acid binding protein 3 (Fabp3), fatty acid bind protein 4 (Fabp4), lipocalin 2 (Lcn2), and orosomucoid 1 (Orm1)) encode members of the calycin lipid binding protein superfamily. This protein superfamily includes three families related by structural rather than sequence homology: 1) Fatty acid binding proteins (FABPs); 2) Lipocalins; and 3) Avidins. Our screen also identified three genes (transferrin (Trf), lactotransferrin (Ltf), and ferritin heavy chain (Fth)) whose protein products are involved in iron metabolism. The Trf, Ltf and Fth proteins operate in concert to maintain an adequate supply of bioavailable iron while sequestering free iron to limit bacterial growth and prevent iron-induced free radical formation.

In this communication, we analyse the expression patterns of mouse genes Fabp3, Fabp4, Lcn2, Orm1, Trf, Ltf, Fth and their human orthologues. Northern analysis of these seven genes confirmed that each is differentially expressed over the course of mouse MG development. The results of bioinformatic analysis revealed that each gene has a human orthologue and that both the mouse and human genes have been isolated from cancer tissue derived cDNA libraries and exhibit altered expression patterns in normal versus cancerous tissues. Taken together, these data suggest that the study of gene expression in normal mouse MG development can identify mouse genes whose human orthologues are expressed in cancerous cells. Using the methods described here, the mouse MG can serve as a valuable model system in which investigators may conduct preliminary experiments to identify genes that may serve as markers for certain cancers and / or that are directly involved in the process of breast carcinogenesis.

*Published online at www.tibetsjournal.org.
Materials and Methods

MG tissues
Female CD1 mice were obtained from and housed at the Central Animal Care Services facility, Faculty of Medicine, University of Manitoba (Winnipeg, MB, Canada). All protocols used are in accordance with the interdisciplinary principles and guidelines for the use of animals in research, testing, and education issued by the New York Academy of Sciences' Ad Hoc Committee on Animal Research. Tissues were isolated from female mice at various time points over the course of the MG development cycle representative of the first parturition: virgin 15 weeks (V15), virgin 20 weeks (V20), pregnancy 2 days (P2, day 0 being the day when the vaginal plug was evident), pregnancy 13 days (P13), lactation 3 days (L3, where day 0 was the day of parturition), lactation 10 days (L10), involution 2 hours/0.08 days (I0.08), involution 1-10 days (I1-10). MG involution was induced by removal of the pups (weaning) at lactation day 10, and all involution time points are measured with reference to this event. Tissue samples were shipped on dry ice to the Wong lab and stored at −70°C until needed.

RNA extraction and cDNA synthesis
Total RNA was isolated using Trizol solution and in accordance with the protocol supplied by the manufacturer (Invitrogen, Carlsbad, CA). cDNA was synthesized using either a cDNA Synthesis Kit (Stratagene, La Jolla, CA) or by use of Powerscript-Reverse Transcriptase (BD Biosciences Clontech, Palo Alto, CA) and the protocols supplied by the manufacturers.

cDNA library synthesis and making of a macroarray
A cDNA library was created from cDNA derived from RNA isolated from mouse MG excised from dams 1 day after weaning (I1) using the UniZAP-XR cDNA Library Synthesis Kit (Stratagene, La Jolla, CA). The unamplified cDNA library was plated and isolated plaques were recovered and stored individually in SM buffer. An aliquot of each phage clone was PCR amplified using standard T3 and T7 primers that flanked the cloning site. The presence of PCR products was confirmed by gel electrophoresis. Those PCR products shown to consist of a single band greater than 500 base pairs in length were selected for inclusion on custom-made macroarrays. Multiple copies of the clone array were made by spotting the PCR products onto nylon membranes in triplicate using a 96-pin arraying device (V&P Scientific, San Diego, CA).

Screening the macroarrays for highly abundant milk protein cDNAs
MG tissue at I1 is known to contain a very high proportion of milk protein mRNA transcripts from whey acidic protein (WAP) and the casein family (α, β, γ, δ, ε, κ). To screen the macroarray for clones of these highly abundant milk protein transcripts, phagemids known to contain cDNA inserts representing milk protein transcripts were isolated, the cloned inserts were excised with a Xho1 and EcoRI digestion, and gel purified using standard protocols. The purified inserts were then random-primer labelled with α-32P-dCTP. Probes were purified using QIAquick PCR purification (Qiagen, Hilden, Germany).

The arrays were pre-hybridized with Hybrisol II solution (Serologicals Corporation, Norcross, GA) at 68°C for 1.5 hours. Hybridization in the presence of the denatured probes and Hybrisol II took place overnight at 68°C. The membranes were washed 3 times for 0.5 hour each with 2X SSC / 0.1% SDS, then twice for 15 minutes each with 0.1X SSC / 0.1% SDS. Clones that showed hybridization with the milk protein cDNA probes were identified by autoradiography. Although many of these clones were still included on the macroarrays for the differential cross-screen, they were excluded from further sequence and Northern analyses.

Macroarray screen for differentially expressed genes in the involuting mouse MG
To identify genes whose expression pattern changes over the course of MG development, we probed two identical macroarray membranes individually with a random-primer labelled amplified cDNA probe derived from 10 day lactating (L10) or 1 day involuting (I1) mouse MG tissues. The same labelling and hybridization protocols described above were used. Minimum activity used per hybridization was 1.6 x 10⁶ CPM/ml of hybridization buffer.

*Published online at www.tibetsjournal.org.
Sequencing, sequence analysis, and bioinformatics

Clones corresponding to differentially expressed genes were identified by autoradiography. Sequencing was done using a Dy enamic ET Terminator Cycle Sequencing kit (Amersham-Pharmacia Biotech, Uppsala, Sweden) as per the manufacturer’s protocol. To determine the identity of the I1 mouse MG cDNA clones and their human orthologues, sequences were submitted to a BLAST search against the GenBank database at NCBI (www.ncbi.nih.gov).

The cDNA libraries from which the seven genes of interest have been isolated were identified and enumerated using data available through UniGene at NCBI. Comparison of putative expression levels for each of the seven genes in cancerous versus normal tissue types was done using data gathered from NCI-CGAP (http://cgap.nci.nih.gov/).

Northern Analysis

Total MG RNA (5-10 µg) from animals representing our developmental profile as well as a 0.24-9.5 kb RNA size ladder (Invitrogen, Carlsbad, CA) were electrophoresed on a 1% agarose / 18% formaldehyde / 10% MOPS gels. Prior to electrophoresis, the RNA was denatured in the presence of ethidium bromide, allowing visualization of the 18S and 28S rRNA bands by UV light after electrophoresis. The RNA on the gel was passively transferred to BioDyne-B nylon membrane (Pall Corporation, Port Washington, NY) using 10X SSC as a transfer medium. Transferred RNA was UV crosslinked to the nylon membrane using a Stratalinker (Stratgene, La Jolla, CA). Gel purified PCR products derived from a single clone as described above were used as templates for random primer radiolabeling of probes used for Northern analysis. The hybridization and wash conditions were identical to those described above for the array analysis. The autoradiograms were digitally scanned at 600dpi and then analysed by densitometry using the histogram function in Adobe Photoshop v7. The autoradiography results of the Northern hybridization were normalized to the 18S band of each blot’s corresponding RNA gel picture prior to transfer. The process quantifies the net intensity of each band on the Northern autoradiograph relative to one another while correcting for variation in RNA loading between each lane.

Results and Discussion

Identification of genes involved in iron and lipid metabolism

The results of the milk protein gene screen identified 77 of the 248 clones (31%) included on the array as corresponding to either WAP or a member of the casein family. These clones were excluded from further analysis. The remaining 171 I1 MG clones were analysed in a differential cross screen using cDNA probes derived from L10 and I1 MG tissues. By means of a visual inspection of the macroarray autoradiographs, 55 putatively differential clones were chosen for sequencing. Through BLAST queries against the non-redundant GenBank database at NCBI, we found that eight of our 55 sequenced clones (14.5%) represented genes that belonged to two groups: [1] genes involved in iron transport and metabolism (one Ltf and four different Fth clones identified); [2] genes coding for members of the calycin lipid-binding protein superfamily (Fabp3, Fabp4 and Lcn2 clones each identified once). Trf and Orm1 clones had been identified in an I1 mouse MG cDNA library by previous screens of the library and were included in our analysis because they belonged in the iron metabolism and calycin gene categories respectively. The seven mouse genes analysed in this study and their human orthologues are listed in Table 1.

Table 1. List of mouse genes analyzed in this study and their human orthologues.

<table>
<thead>
<tr>
<th>Mouse Gene Name</th>
<th>GenBank mRNA Accession Number</th>
<th>Human Orthologue mRNA Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fatty Acid Binding Protein 3 (Fabp3)</td>
<td>NM_010174</td>
<td>NM_004102</td>
</tr>
<tr>
<td>Fatty Acid Binding Protein 4 (Fabp4)</td>
<td>NM_024406</td>
<td>NM_001442</td>
</tr>
<tr>
<td>Ferritin Heavy Chain (Fth)</td>
<td>NM_010239</td>
<td>NM_002032</td>
</tr>
<tr>
<td>Lactotransferrin (Ltf)</td>
<td>NM_008522</td>
<td>NM_002343</td>
</tr>
<tr>
<td>Lipocalcin 2 (Lcn2)</td>
<td>XM_130171</td>
<td>NM_005564</td>
</tr>
<tr>
<td>Orosomucoid 1 (Orm1)</td>
<td>NM_008768</td>
<td>NM_000607</td>
</tr>
<tr>
<td>Transferrin (Trf)</td>
<td>NM_133977</td>
<td>NM_001063</td>
</tr>
</tbody>
</table>

*Published online at www.tibetsjournal.org.*
Northern analysis

To profile the expression pattern of each gene in the mouse MG, we performed Northern analysis using RNA isolated from various MG developmental time points. Inspection of the Northern autoradiographs reveals that each gene analyzed in this study is differentially regulated over the course of mouse MG development (Figures 1, 2, 3). Northern analysis of WAP (Figure 1) is included to demonstrate that our MG RNA samples correlate with previously published experimental samples of mouse mammary gland. 19

![Whey Acidic Protein](image)

**Figure 1.** Northern analysis of whey acid protein (WAP) over the course of mouse mammary gland development. Northern analysis shows that WAP expression is greatly induced in the lactating mouse MG and declines rapidly after day one of involution. By day five of involution, WAP mRNA is undetectable by Northern analysis. Our Northern analysis of WAP expression during mouse mammary development agrees with previously published findings19, thereby supporting the authenticity of our mouse MG RNA samples. The top portion of the panel is an image of an autoradiogram from a mouse MG developmental profile northern probed with a cDNA clone for WAP with the band size derived from comparison to the RNA ladder loaded on the gel. The middle portion of the panel is an image of the 18S r RNA band from the gel. The graph shows the Northern results normalized to their corresponding 18S rRNA bands to correct for variations in RNA loading. Normalization units are arbitrary and not intended for quantitative comparisons between different Northern autoradiograms. Legend: MG developmental time points: V15, V20= Virgin (weeks); P2, P13 = Pregnancy (days); L3, L10 = Lactation (days); I0.08 - I9 = Involution (days).

Calycin genes

The calycin superfamily of genes under analysis are Fabp3, Fabp4, Lcn2, and Orm1. Consistent with the variety of functions ascribed to the members of the calycin lipid-binding protein superfamily, Northern analyses of Fabp3, Fabp4, Lcn2 and Orm1 reveal a great variety of expression patterns over the course of mouse MG development (Figure 3A-D).

Fabps are cytosolic proteins of 14-15 kDa whose structure is characterized by a ten-stranded β-barrel, the interior of which binds hydrophobic ligands. 20 There are at least 10 distinct Fabp genes in mice, each with a unique tissue enriched expression profile. 21 Fabp4 (A-Fabp / AP2) shows enriched expression in differentiated adipocytes where it is believed to transport and target intracellular fatty acids via protein-phospholipid and protein-protein interactions. 22 Fabp3 (H-FABP / MDGI) is highly expressed in the heart and differentiated mammary epithelial cells (MECs) where it mediates the uptake, transportation and targeting of fatty acids. 23,24 The Fabp3 protein exhibits growth-inhibitory and prodifferentiation functions that may be independent of its fatty acid-binding properties. 25 In our analysis we observe that Fabp3 and Fabp4 expression patterns are inversely correlated to one another (Figure 3A and 3B). The virgin and virgin-like (I9-I0) states, in which the predominant cell type are adipocytes, are marked by high levels of Fabp4 expression, while Fabp3 mRNA predominates in the lactating gland, where the major cell type are MECs. These expression patterns agree with the known cell-type specific expression of these genes and serve to reinforce the fact that MG development is characterized by a wholesale

*Published online at www.tibetsjournal.org.*
turnover in cell populations. Both Fabp3 and Fabp4 have been implicated in intracellular lipid signaling pathways leading to changes in gene expression. Although the normally cytosolic FABPs have been identified in the nucleus, they are hypothesized to affect transcription mainly by delivering (either directly or indirectly) hydrophobic ligands to the PPAR (peroxisome proliferator activated receptor) family of transcription factors. 22

Lipocalin 2 (Lcn2 / NGAL / 24p3) and orosomucoid 1 (Orm1 / α1-Acid glycoprotein / AGP) are secreted proteins with a signature lipocalin eight-stranded β-barrel structure. The expression of Lcn2 and Orm1 is upregulated at day one of involution (Figures 3C and 3D). Orm1 is a heavy glycosylated protein whose levels rise in response to injury or inflammation. 27 Although Orm1 has been shown to bind hydrophobic compounds, its anti-inflammatory properties depend largely upon its pattern of glycosylation and may be independent of the peptide altogether. 28 Where Orm1 expression steadily declines after 11 to levels resembling those of the virgin gland, the level of Lcn2 mRNA remains high in the regressed gland. It has been suggested that elevated post-involution expression of Lcn2 contributes to the protective, anti-cancer effect of the first full-term pregnancy. 29 Lcn2 has been implicated in a variety of activities including the induction of apoptosis, 30, 31 and the transport of siderophore-bound iron into the cell. 32 Lcn2 is also an immunosuppressive anti-inflammatory agent that, together with lactotransferrin, is a component of neutrophil granules. 33 The human LCN2 protein NGAL associates with a neutrophil gelatinase metalloproteinase MMP-9. 34 Gelatinases have been implicated in extracellular matrix (ECM) degradation and tissue remodeling, as well as being responsible for tissue damage during inflammation. 35 Together, Orm1 and Lcn2 may act as anti-inflammatory agents to protect tissue from damage in the wake of a possible innate immune response.

Iron metabolism genes
The iron metabolism genes we identified in our screen are Fth, Ltf, and Trf. Iron is both an absolute metabolic requirement and a dangerous toxin for most cellular organisms. Thus, mechanisms have evolved to provide for a cell’s iron requirements while minimizing toxicity. Transferrin is a secreted protein required to transport extracellular iron to the cytosol for use in normal cellular metabolism. 15 Lactotransferrin, a homologue of transferrin, plays a role in detoxification by sequestering extracellular iron, thereby serving the dual purpose of preventing tissue damage in an iron rich environment and removing free iron that could be utilized in bacterial metabolism. 36, 37 In addition, other antimicrobial, enzymatic, 16 growth regulatory, 38 and transcriptional regulatory 39 properties of lactotransferrin have been described. Once internalized, both transferrin and lactotransferrin transfer their bound iron to ferritin where it is safely stored until required by the cell.

The expression patterns observed for Trf, Ltf, and Fth suggests that iron metabolism requirements of the MG are developmentally influenced (Figure 2A-C). Trf is most highly expressed during lactation and early involution (Figure 2A). Ltf expression is induced in the involuting gland (Figure 2B). Fth, the major subunit of the multimeric ferritin intracellular iron-storage protein, is expressed throughout MG development (Figure 2C). However, Fth mRNA levels vary drastically over the course of MG development with peak expression observed at involution day three (13). Although all three iron metabolism genes show unique expression patterns, they are co-expressed during the first stage of MG involution signifying an emphasis on iron metabolism at this developmental stage. The induction of Lcn2 in early involution and its proposed role in iron metabolism supports this notion (Figure 3C).

Literature mining reveals an association between PPARs and the seven genes studied
A survey of the literature reveals a link between the PPAR (peroxisome proliferator activated receptor) family of transcription factors and the expression of six of the genes studied here (Fabp3; 40 Fabp4; 22 Ltf; 41 Trf; 42 Lcn2; 21 Orm1 45). Although no experimental evidence has yet been found linking PPARs with Fth expression, Jang et al. 44 have found evidence suggesting that PPARγ may modulate expression of the ferritin light chain (FTL) gene. Ferritin functions as a multimer of heavy and light chains, and therefore alterations at the genetic level of the ferritin light chain through PPAR interactions may indirectly have an effect on the biological activity of ferritin. PPARs may therefore play an important regulatory role during mouse MG involution and the expression of genes involved in response to oxidative and inflammation-induced stress. As previously mentioned, FABPs may actively direct intracellular lipid signalling by delivering PPAR-activating ligands, thereby indirectly affecting gene transcription. A more detailed analysis of FABP-mediated signalling and the expression of these cytoprotective genes in the context of PPAR regulation are therefore warranted.
Figure 2. Northern analysis of iron metabolism genes *transferrin*, *lactotransferrin* and *ferritin heavy chain* during mouse MG development. Using total RNA isolated from mouse mammary glands at various developmental time points, we performed Northern analyses to show that the expression of iron metabolism genes [A] *transferrin* (*Trf*), [B] *lactotransferrin* (*Ltf*), and [C] *ferritin heavy chain* (*Fth*) are differentially regulated over the course of mouse MG development. All three genes are highly expressed during early involution. The top portion of each panel is an image of the autoradiogram with band sizes indicated derived from comparison to the RNA ladder on the gel. The bottom portion of each panel is an image of the 18S rRNA band from the gel that is used as a loading control. Accompanying each autorad picture is a graph representing the Northern results normalized to the 18S rRNA band from the original RNA gel used for the Northern transfer. Normalization units are arbitrary and not intended for quantitative comparisons between different Northern autoradiograms. Only select time points are shown to facilitate qualitative comparisons with the calycin Northern data. Legend: V15, 20 = Virgin (weeks); P2, 13 = Pregnancy (days); L3, L10 = Lactation (days); I0.08 – I9 = Involution (days).

In silico analysis of gene expression in cancerous tissues and cell lines

Having established that each of the seven genes are differentially regulated during mouse MG development by Northern analysis, we turned to *in silico* analyses to both qualitatively and quantitatively determine the possible expression of these genes in cancerous cells. Using the UniGene tissue expression resource at NCBI, we searched for ESTs representing the genes under study here, derived from the analysis of cancerous tissue cDNA libraries. Ten cancerous mouse tissue cDNA libraries were identified from which at least one of our seven genes was isolated as a cDNA clone (Table 2). *Trf*, *Ltf*, *Fth* and *Lcn2* clones were isolated from the majority of mouse cancer cDNA libraries identified, suggesting that expression of these iron-metabolism genes is commonly associated with a cancerous phenotype. The fact that clones representing *Fabp3*, *Fabp4* and *Orm1* were isolated from a more restricted subset of cancer cDNA libraries suggests that the expression of these three genes may be unique for specific forms of cancer.

Table 2. Identification of ESTs representing our seven mouse genes in cancer cDNA libraries. To determine if any of our seven mouse genes are expressed in cancerous tissues, we examined the tissue expression data from NCBI's UniGene database. Areas marked with an "X" indicate that at least one EST representing a particular gene has been identified in a cDNA library made from a cancerous mouse tissue. These data show that ESTs representing *Lcn2*, *Fth*, *Ltf* and *Trf* are isolated from a greater variety of cancerous tissues than those representing *Fabp3*, *Fabp4* or *Orm1*.
In order to extend our analysis to human cancer, we repeated the above analysis for the human orthologues of the seven genes. Due to the large variety of human cancer cDNA libraries, Table 3 includes only those tissues or cell lines from which two or more of our seven genes have been isolated as cDNA clones. In agreement with the mouse data in Table 2, the general trend remains that \(\text{LCN2, FTH1, LTF, and TF}\) are expressed in a greater diversity of human cancers than are \(\text{FABP3, FABP4, or ORM1}\). However, whereas \(\text{Lcn2, Trf, Ltf, and Fih}\) are commonly co-expressed in mouse cancers, their human orthologues do not show this same expression pattern. \(\text{FTH1}\) clones have been identified in almost all cDNA sources listed in Table 3 (except the hepatocellular carcinoma cell line), indicating that, like mouse \(\text{Fih}\), \(\text{FTH1}\) expression is a common feature of most cancers. The relatively rare isolation of \(\text{FABP3, FABP4, or ORM1}\) clones again suggests that their expression may be confined to a small subgroup of cancers.

In order to quantitatively compare the expression of each of our seven genes in normal and cancerous mouse and human tissue, we analyzed the Cancer Genome Anatomy Project (CGAP) database (http://cgap.nci.nih.gov/). CGAP collects and analyses expression data from cDNA libraries made from a variety of normal and cancerous tissues. To quantify expression levels, we graphed the raw EST data as the frequency at which a specific gene is isolated. We analyzed the expression data for our seven genes in normal all tissue, cancer all tissue, normal MG, and cancer MG for both the mouse and human orthologues. The term “all tissue” describes the EST data from all individual tissue cDNA libraries that has been pooled to show the general expression trend in normal and cancerous states. The analysis of all tissues or all cancer states was performed to get a feel on the trends in expression based on the existing database, levels reflect the pools of tissues and samples that have been sampled with respect to expression. It is expected that within any specific tissue or tissue type cancers that variations in the trends may be observed. The results are shown in Figure 4A-G. Analysis of the data reveals that a cancer phenotype correlates with altered expression patterns for all of the genes analyzed. Consistent with the UniGene data, mouse \(\text{Fabp3, Fabp4, andOrm1}\) form a group based on the observation that they are generally downregulated in cancer given the current database. Also consistent with the UniGene data, mouse \(\text{Trf, Ltf, Fih, and Lcn2}\) all appear to be transcriptionally upregulated in cancer and may point to the importance of iron metabolism in a murine cancer phenotype.

There is a general consistency between the mouse and human CGAP gene expression patterns (Figure 4 A-G). As with their mouse orthologues, \(\text{FABP3, FABP4, and ORM1}\) are expressed at a lower level in cancerous tissue relative to the normal condition, while higher levels of \(\text{FTH1}\) expression is consistently observed in most human cancers. However, there are some notable differences between the mouse and human gene expression data. Expression of \(\text{Ltf}\) and \(\text{Trf}\) are convincingly upregulated in mouse cancer; however this trend is not observed for their human orthologues (Figure 4A and B). \(\text{TF}\) expression appears to be downregulated in human cancer, a trend opposite to that observed for mouse \(\text{Trf}\). The CGAP data reveal a slight tendency for transcriptional upregulation of \(\text{LTF}\) expression in all tissue human cancer, but for a decrease in human breast cancer. These gene expression differences between mice and humans were not unexpected given the considerable evolutionary distance between the two species. This therefore demonstrates that our trans-species expression analyses can be used to reveal both the similarities and the differences between mouse and human cancer phenotypes. Although our CGAP analysis confirmed that mouse and human cancer gene expression data for five of our seven genes lends support for the use of the mouse MG model system for studying human breast cancer.

A comparison of the UniGene cancer expression data (Tables 2 and 3) with the more quantitative CGAP data (Figure 4) reveals that the two data sets may both be indicative of general gene expression trends in cancerous cells. Genes that are commonly isolated from cancer cDNA libraries tend to be transcriptionally upregulated as part of a cancerous phenotype, whereas genes that are rarely identified in cancer cDNA libraries tend to be transcriptionally downregulated in cancer. For example, Table 2 illustrates that \(\text{Trf, Ltf, Fih, and Lcn2}\) have been identified in a greater assortment of mouse cancer cDNA libraries than have \(\text{Fabp3, Fabp4, and Orm1}\). Likewise, CGAP expression analysis reveals that the former grouping of genes are all upregulated in cancerous tissues while genes of the latter group are all downregulated (Figure 4). Therefore, the relative number and diversity of cancer cDNA libraries from which a gene has been isolated may be used as a first indicator of transcriptional trends in cancer.

*Published online at www.tibetsjournal.org.
Table 3. Identification of ESTs representing human orthologues from various human cancer cDNA libraries. To extend the expression analysis of our seven mouse genes to human cancer, we identified the orthologous human genes and examined the tissue expression data from NCBI’s UniGene database. Areas marked with an “X” indicate that at least one EST representing a particular gene has been identified in a cancerous tissue or cell line cDNA library. Because of the large number of cancers from which only one of the seven genes was isolated, we limited our analysis to those tissues in which two or more of our genes were represented. These data suggest that FTH1 expression is a very common feature of human cancers.

<table>
<thead>
<tr>
<th>Human cDNA Source</th>
<th>FABP3</th>
<th>FABP4</th>
<th>ORM1</th>
<th>LCN2</th>
<th>FTH1</th>
<th>LTF</th>
<th>TF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenocarcinoma</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma - endometrial; well differentiated</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma - moderately differentiated</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma - poorly differentiated with signet cell features</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carcinoma - serous papillary; high grade 2 pooled</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell line - adenocarcinoma</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell line - ductal carcinoma</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell line - epithelioid carcinoma</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell line - hepatocellular carcinoma</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chondrosarcoma</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chondrosarcoma - metastatic</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Choriocarcinoma</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glioblastoma</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lymphoma - follicular small and large cell</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myeloma</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neuroblastoma cells</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oligodendrogloma - anaplastic</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumour – colon RER+</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumour – germ cell; pooled</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumour – lung</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Tumour – nervous</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumour – ovarian</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumour – parathyroid</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumour – prostate; invasive</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumour – renal cell</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Conclusion**

Mouse MG development is a powerful tool for the identification of genes involved in cell proliferation and programmed cell death. Analysis of our Northern data in light of MG developmental biology suggests that the transition from a molecular environment favoring cell survival (lactation) to one that is sensitive to apoptosis (involution) requires major changes in gene expression and shifts in cellular metabolism. The down-regulation of Fabp3 one day after weaning (Figure 3B) implies that there is a switch from lipid to glucose metabolism in the organ at this developmental time point. Therefore, there may be an association between glucose metabolism and the apoptotic process. Lowering fatty acid utilization in the involuting gland may provide the tissue some protection from free radical damage that can arise from non-enzymatic lipid peroxidation in the presence of an iron overload. Along these same lines, the up-regulation of iron metabolism genes Fth, Ltf, Trf and Lcn2 may also serve to protect the early involuting gland from iron-induced oxidative stress. Lcn2 and Ltf may cooperate with Orm1 to protect the gland from potential tissue damage by the inflammatory response during involution. The risk of free radical damage through lipid peroxidation and an inflammatory response has been shown to increase during tissue regression. In addition, all seven genes examined in this study may be partially regulated by the PPAR family of transcription factors. Further investigation into the role of PPARs in regulating MG responses to oxidative stress is therefore warranted. In at least one involuting tissue (thymus), evidence that oxidative stress antagonizes apoptosis has been found. It would therefore appear that even in tissues undergoing active cell death there is an undercurrent of events that provide levels of protection to the tissue as a whole.

*Published online at www.tibetsjournal.org.*
Figure 3. Northern analysis of calycin genes Fabp4, Fabp3, Lcn2, and Orm1. Using total RNA isolated from mouse mammary glands at various developmental time points, we performed Northern analyses to show that the expression of calycin genes [A] Fatty acid binding protein 4 (Fabp4), [B] Fatty acid binding protein 3 (Fabp3), [C] Lipocalin 2 (Lcn2), and [D] Orosomucoid 1 (Orm1) are differentially regulated over the course of mammary development. The top portion of each panel is an image of the autoradiogram with band sizes derived from comparison to the RNA ladder on the original gel. The bottom portion of each panel is an image of the 18S rRNA band from the gel which is used as a loading control. Accompanying each autorad picture is a graph representing the Northern results normalized to the 18S rRNA band from the original RNA gel used for the Northern transfer. Normalization units are arbitrary and not intended for quantitative comparisons between different autoradiograms. Legend: Vx = Virgin (weeks); Px = Pregnancy (days); Lx = Lactation (days); Ix = Involution (days).
Figure 4. Analysis of CGAP gene expression data in normal and cancerous mouse and human tissues. The graphs represent the proportion of gene-specific ESTs to the total number of ESTs identified according to the CGAP database as of August 6, 2003. Analysis of these data show altered expression patterns for all seven genes in normal versus cancerous tissues. In general, iron metabolism genes Trf [A], Ltf [B], Fth [C], and Lcn2 [D] are upregulated in cancer while lipid-binding protein genes Fabp3 [E], Fabp4 [F], and Orm1 [G] are downregulated. The absence of coloured bars on a graph means that no ESTs for that gene has been identified in that tissue. Legend: Tissue types: A – Human normal all tissue; B – Human cancer all tissue; C – Human normal mammary gland; D – Human cancer mammary gland; E – Mouse normal all tissue; F – Mouse cancer all tissue; G – Mouse normal mammary gland; H – Mouse cancer mammary gland. The red, green, blue, and yellow coloured bars represent CGAP gene expression data for human all tissue, human MG, mouse all tissue and mouse MG respectively.
To determine whether these genes had altered expression patterns in mouse and human cancer, we analyzed the data available through the NCBI UniGene resource and the CGAP database. We found that all seven genes were differently expressed in cancerous tissue relative to its normal counterpart. In five of seven cases, the changes in gene expression patterns were the same in both mice and humans, thereby lending support to the use of the mouse model system to characterize human cancer. In the two cases where the mouse and human expression data do not agree (Trf / TF and Ltf / LTF), the in silico analysis still yields meaningful data in that detailing the genotypic differences between murine and human cancers may lead to a better understanding of oncogenesis in general. Given that our in silico analyses for these seven known genes produced expression data that correspond to those in the literature, we feel that the study of gene expression in the normal mouse MG, followed by the appropriate bioinformatic analyses, can provide useful and accurate data concerning cancer expression patterns for uncharacterized genes. This methodology may be employed as a preliminary means to profile the human orthologues of many of the novel mouse genes identified in MG differential screen. By applying the in silico methods outlined here, human orthologues of uncharacterized mouse genes can be identified, and their putative expression levels in cancerous tissues quantified.

Acknowledgements

This research was supported by grants from NSERC (PW) and (YM). Scholarship support was provided by NSERC (TE, CG, VL, DN, MAC), AHFMRI (TE, CG, DN), the Alberta Cancer Board (VL, PW), the Walter H. Johns Graduate Fellowship program (MAC) and the Province of Alberta Graduate Fellowship program (MAC).

References


*Published online at www.tibetsjournal.org.*