Effects of oestrogen on gene expression in epithelium and stroma of normal human breast tissue

C L Wilson, A H Sims, A Howell, C J Miller and R B Clarke

Abstract

Oestrogen (E) is essential for normal and cancer development in the breast, while anti-oestrogens have been shown to reduce the risk of the disease. However, little is known about the effect of E on gene expression in the normal human breast, particularly when the epithelium and stroma are intact. Previous expression profiles of the response to E have been performed on tumour cell lines, in the absence of stroma. We investigated gene expression in normal human breast tissue transplanted into 9–10-week-old female athymic nude (Balb/c nu/nu) mice. After 2 weeks, when epithelial proliferation is minimal, one-third of the mice were treated with 17β-oestradiol (E2) to give human luteal-phase levels in the mouse, which we have previously shown to induce maximal epithelial cell proliferation. RNA was isolated from treated and untreated mice, labelled and hybridized to Affymetrix HG-U133A (human) GeneChips. Gene expression levels were generated using BioConductor implementations of the RMA and MAS5 algorithms. E2 treatment was found to represent the largest source of variation in gene expression and cross-species hybridization of mouse RNA from xenograft samples was demonstrated to be negligible. Known E2-responsive genes (such as TFF1 and AREG), and genes thought to be involved in breast cancer metastasis (including mammoglobin, KRT19 and AGR2), were upregulated in response to E treatment. Genes known to be co-expressed with E receptor α in breast cancer cell lines and tumours were both upregulated (XBP-1 and GREB1) and downregulated (RARRES1 and GATA3). In addition, genes that are normally expressed in the myoepithelium and extracellular matrix that maintain the tissue microenvironment were also differentially expressed. This suggests that the response to oestrogen in normal breast is highly dependent upon epithelial–stromal/myoepithelial interactions which maintain the tissue microenvironment during epithelial cell proliferation.

Introduction

Ovarian steroids, acting through nuclear receptors, are important for the development of the normal breast and breast cancer. An early menarche and a late menopause have been shown to increase breast cancer risk, whereas an early menopause is protective; these data suggest that breast cancer risk is related to a cumulative exposure to ovarian hormones. Oestrogen (E) is the focus of breast cancer therapies because tumours are often dependent on this steroid for growth (Clarke et al. 2004).

Oestrogen is central to breast cancer development and interventions that alter exposure, such as oophorectomy or selective E receptor modulators (SERMs) and aromatase inhibitors has been shown to significantly reduce breast cancer risk (reviewed in Howell et al. 2005). Despite this, little is known about the effect of E on the normal human breast, particularly when the epithelium and stroma are intact.

Improved understanding of the molecular and cellular biology of the breast is important for future risk-prediction and -reduction strategies. Direct effects...
on the human breast are difficult to study because of problems of access, and breast heterogeneity. Genetic experiments on the rodent breast, especially during puberty, have highlighted the importance not only of the epithelium but interactions between these cells and other cell types within the breast, such as fibroblast, adipocytes and macrophages (Howell et al. 2005, Wiseman & Werb 2002). The stroma represents the major component of the adult human breast; during early reproductive life approximately 20% of breast tissue comprises epithelium, 20% fat and 60% connective tissue; the amounts of epithelium and connective tissue decline with age (Hutson et al. 1985). It has been suggested that despite the lack of E receptors in human stromal cells, the stroma plays an important role in the control of hormone-mediated epithelial cell proliferation (Anderson et al. 1998), through paracrine mechanisms (Shekhar et al. 2001).

Tumour cell lines or mouse models are extensively relied upon in cancer studies; however, there are significant differences between the cellular composition and behaviour of cultured cells in an in vitro environment and those of normal human tissue in vivo. It is likely that these differences will have a big impact on intracellular signalling and proliferation in response to environmental stimuli. Harvell et al. (2006) recently demonstrated that 17β-oestradiol (E2) regulates different genes in human breast tumour xenografts compared with the identical cells in culture. To date, a number of gene-expression profiling studies have investigated the response to E stimulation in tumour cell lines (Coser et al. 2003, Frasor et al. 2003, 2004, Inoue et al. 2002, Vendrell et al. 2004), in the absence of stroma, which does not accurately portray the breast tissue microenvironment. Some of these studies (Frasor et al. 2003, 2004, Inoue et al. 2002, Seth et al. 2002) employed high doses (10 nmol/l) of E2, equivalent to those observed during pregnancy and considerably higher than the fluctuating levels normally seen during the menstrual cycle. Normal breast tissue cell lines, such as MCF10As, are inappropriate for studying the response to E stimulation as they have few or no classical E receptors (ERα/ESR1). Mouse models are also unsatisfactory for studying the effects of E on normal breast as mouse mammary fat pad stromal cells are ESR1-positive, whereas their human counterparts are ESR1-negative (Anderson et al. 1998). Although expression profiling of tumour cell line xenografts has been previously performed (Armes & Venter 2002, Creighton et al. 2003, 2005), we report the first attempt to characterize the effects of E on normal human breast tissue using xenografts.

In this study, we transplanted normal human breast tissue into athymic nude mice in order to obviate the effects of variations in the stage of the menstrual cycle between patients, which is known to affect the proliferative activity of the normal breast (McManus & Welsch 1984, Potten et al. 1988) and has previously been shown to provide a baseline for analysis of the effects of E2 stimulation (Clarke et al. 1997, Laidlaw et al. 1995). Subcutaneous implantation of whole pieces of normal breast tissue into the athymic nude mice preserves the normal tissue architecture and, presumably, normal epithelial–mesenchymal interactions. We have investigated genes whose E regulation is persistent and thus would reflect the physiological exposure to endogenous E, whereas previous studies have generally focused on short-term effects (Frasor et al. 2003, Inoue et al. 2002, Xu et al. 2005). Administration of a 2 mg E2 slow-release pellet for 7 days produces similar serum concentrations of E2 to those published for the luteal phase of the menstrual cycle and which we have previously shown to be optimal for inducing proliferation of normal breast epithelium (Laidlaw et al. 1995). Here we describe the first attempts to study the gene expression profile in epithelium and stroma of normal human breast tissue in response to E stimulation at the level of the luteal phase of the menstrual cycle.

Materials and methods

Patient samples

Following approval of the local research ethics committee and with patient consent, normal human breast tissue was obtained from six pre-menopausal women, either at least 1 cm away from benign lesions or from women undergoing reduction mammoplasty (median age, 36.5 years; range, 30–42 years; three parous, three nulliparous; median menarche age, 13.5 years; range, 11–15 years). The tissue was confirmed to contain no abnormalities by histology.

Implantation of breast tissue and E2 administration

All of the animals used were 9–10-week-old intact female athymic nude (Balb/c nu/nu) mice (Clarke et al. 1997; Harlan, Bicester, Oxon, UK). All surgical procedures were performed as described previously (Laidlaw et al. 1995) and were carried
out under the Animals (Scientific Procedures) Act of 1986 (UK). Intact mice were used since they have low median levels of -E2 (94 pmol/l) with short-lasting (<24 h) peak serum levels during oestrous that are equivalent to the human follicular phase. As we have previously demonstrated, human breast tissue implanted into intact mice shows little variation in E-dependent proliferation and progesterone receptor expression (Clarke et al. 1997, Laidlaw et al. 1995).

The normal breast tissue samples were divided into approximately 2 x 2 x 1 mm pieces and implanted subcutaneously by making two small incisions across the midline dorsal skin as previously described (Clarke et al. 1997). Eight pieces of breast tissue were implanted into six mice for each breast tissue sample. Two weeks after implantation of breast tissue, exogenous E2 was administered to two of the mice by inserting a 2 mg subcutaneous silastic pellet at the base of the tail away from the site of tissue implantation. At 1 week post-treatment or 3 weeks with no treatment, the xenografts were retrieved. In order to compensate for possible issues of breast tissue heterogeneity, we compared 12 pools of RNA extracted from untreated xenografts from six women with six pools of RNA from E2-treated tissue from the same women (see Fig. 1). Portions of breast tissue both prior to and following implantation were fixed in formalin (3.7% formaldehyde in PBS) for 1 h before transfer into paraffin for sectioning and histological examination of the tissue.

**Array processing and analysis**

Tissue samples were ground to a fine powder under liquid nitrogen and RNA was isolated using Trizol (Ambion) according to the manufacturer’s instructions. RNA was purified using Qiagen RNeasy columns (Qiagen, Valencia, CA, USA) and quantified using a Nanodrop spectrophotometer (Labtech). The quality and amount of starting RNA were confirmed with an Agilent Bioanalyzer 2100 (Agilent) prior to labelling and hybridization to Affymetrix HG-U133A GeneChips (using standard protocols available at http://bioinformatics.pcr.man.ac.uk/mbcf/protocols.shtml). After scanning, array images were assessed by eye to confirm scanner alignment and the absence of significant bubbles or scratches on the chip surface. Ratios (3’/5’) for GAPDH and β-actin were confirmed to be within acceptable limits (0.86–1.38 and 0.70–1.24, respectively), and BioB spike controls were found to be present on all chips, with BioC, BioD and CreX also present in increasing intensity. When scaled to a target intensity of 100 (using Microarray Analysis Suite™ (MAS) version 5.0 array analysis software; Affymetrix) scaling factors for all arrays were within acceptable limits (0.93–2.60), as were background, Q value and mean intensities. The raw spot readings were imported into BioConductor (http://www.bioconductor.org) using R (Ihaka & Gentleman 1996) and normalized using either MAS or RMA algorithms (Irizarry et al. 2003) as implemented in the Simpleaffy package (Wilson & Miller 2005). The MAS data were filtered to remove probesets that were not called present on either four out of the six treated chips or eight out of the 12 untreated chips. Due to the low number of samples, a rank product analysis (Breitling et al. 2004) method was used to determine the significance of differentially expressed genes using the RankProd BioConductor module within R. Using the rank product analysis method, a false discovery rate
of 10% was employed and fold changes were calculated by comparing the mean expression level of probesets from the two untreated samples with those of the E2-treated samples. Unsupervised principle components analysis was performed using the MaxD View software available from the University of Manchester Bioinformatics Department (http://bioinf. man.ac.uk/microarray/maxd/). The microarray data are Minimum Information About a Microarray Experiment (MIAME) compliant and accessible via MIAME VICE (http://bioinformatics. picr.man. ac.uk/mvice/index.jsp).

Possible cross-species hybridization

Recovery of the human tissue from the mice introduces the potential for skewing of the data due to contamination of the human RNA with traces of mouse RNA. To investigate potential cross-species hybridization of mouse sequences with probesets designed to detect human sequences, varying concentrations of mouse RNA (100, 50 or 20% of the normal protocol amounts, in duplicate) were hybridized to six Affymetrix human U95A GeneChips (these experiments were performed prior to later versions of GeneChips becoming available).

Morphometry

Before and after implantation, the normal human breast tissue was examined by morphometry. The cellular percentage composition of the tissue samples was determined using an 11 x 11 grid at x200 magnification and by counting the squares filled with epithelial cells, stromal cells (including fibroblasts and endothelium) and fat.

Immunohistochemical quantification

Proliferative activity was assessed by immunohistochemistry using the mouse monoclonal antibody MIB-1 (Coulter) raised against the Ki67 proliferation-associated antigen. Expression of PgR and pS2 was determined by immunohistochemistry using a rat monoclonal anti-PR antibody (clone KD68; Abbott Laboratories) and a rabbit anti-pS2 polyclonal (NCL-pS2; Novocastra) respectively. Microwave antigen-retrieval methods and dilutions were as described previously (Clarke et al. 1997). Antibody binding was detected indirectly using the appropriate biotinylated second antibodies, a peroxidase-conjugated avidin–biotin complex (ABC Elite; Vector Laboratories) and diaminobenzidine as the chromogen. Quantitation of immunostaining was carried out with a light microscope and was restricted to the epithelial cells of the terminal-duct lobulo-alveolar units. Areas to be counted were selected out of focus at low power and then complete high-power fields were scored. At least 1000 epithelial cells were scored per sample and the number of labelled cells was expressed as a percentage of the total cells counted. The intensity of staining was not assessed.

Results

Differential expression in response to E

In order to find statistically significant changes in gene expression with our relatively small data set we utilized a rank products method which reliably and consistently outperforms a non-parametric t test, even with noisy data (Breitling et al. 2004). Analysis of gene expression in normal human breast tissue identified that a greater number of probesets are significantly differentially expressed in response to luteal phase levels of E2 when data are analysed using RMA (365; Irizarry et al. 2003), compared with MAS (93; Sasik et al. 2002). Comparing these two lists identified a total of 61 probesets (Table 1) commonly differentially expressed with both methods. A consensus approach that only considers probesets identified as changing by a number of different normalization methods has the advantage of providing a smaller but consistent gene list at the expense of missing some differentially expressed genes (Choe et al. 2005). However, the proportion of up- and downregulated genes was different between the two methods; 79% (289/365) of probesets were downregulated in response to E using RMA, compared with 52% (48/93) of probesets downregulated using MAS.

The list of significantly differentially expressed genes in response to E2 in normal human breast tissue (Table 1) can be assigned into three major groups. Firstly, a number of established E2-responsive genes indicated in Table 1 have been identified in previous studies of breast cancer cell lines (Coser et al. 2003, Frasor et al. 2003, 2004, Inoue et al. 2002, Vendrell et al. 2004), including Trefoil factor 1 (TFF1; formerly pS2), amphiregulin (AREG), myosin-binding protein C (MYBPC1) and transcription factor b-Myc (MYB). TFF1 has previously been shown to stimulate cell migration of breast cancer cells (Crosier et al. 2001, Prest et al. 2002) and a paracrine role for TFF1 has been postulated in gastrointestinal epithelial cells in
Table 1  E2-response probesets identified as differentially expressed in normal human breast tissue. FC, fold change; pfp is the estimated false discovery rate value for the list up to each gene if that gene was the cutoff point. Italics represent a mouse sequence present on the human GeneChip, which should be discounted (see Table 2).

<table>
<thead>
<tr>
<th>Gene Description</th>
<th>Probeset ID</th>
<th>FC MAS</th>
<th>pfp MAS</th>
<th>FC RMA</th>
<th>pfp RMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFF1 (Trefoil factor 1 (pS2))</td>
<td>205009_at</td>
<td>12.2</td>
<td>0.000</td>
<td>10.9</td>
<td>0.000</td>
</tr>
<tr>
<td>MYBPC1 (Myosin-binding protein C, slow type)</td>
<td>214087_s_at</td>
<td>4.4</td>
<td>0.000</td>
<td>2.9</td>
<td>0.000</td>
</tr>
<tr>
<td>AREG (Secretoglobin, family 1D, member 2 (lipophilin B))</td>
<td>206799_at</td>
<td>3.6</td>
<td>0.000</td>
<td>2.3</td>
<td>0.000</td>
</tr>
<tr>
<td>TFF3 (Trefoil factor 3 (intestinal))</td>
<td>204623_s_at</td>
<td>3.3</td>
<td>0.000</td>
<td>2.6</td>
<td>0.000</td>
</tr>
<tr>
<td>SERPINA1 (Serine (or cysteine) proteinase inhibitor, clade A)</td>
<td>202833_s_at</td>
<td>2.5</td>
<td>0.000</td>
<td>2.0</td>
<td>0.000</td>
</tr>
<tr>
<td>C1orf34 (Chromosome 1 open reading frame 34 (DEME-6))</td>
<td>210652_s_at</td>
<td>1.9</td>
<td>0.000</td>
<td>1.6</td>
<td>0.000</td>
</tr>
<tr>
<td>PIP (Prolactin-induced protein)</td>
<td>206509_at</td>
<td>1.9</td>
<td>0.000</td>
<td>1.6</td>
<td>0.000</td>
</tr>
<tr>
<td>AGR2 (Anterior gradient 2 homologue)</td>
<td>204798_at</td>
<td>1.6</td>
<td>0.000</td>
<td>1.6</td>
<td>0.000</td>
</tr>
<tr>
<td>C1orf34 (Chromosome 1 open reading frame 34 (DEME-6))</td>
<td>210652_s_at</td>
<td>1.9</td>
<td>0.000</td>
<td>1.6</td>
<td>0.000</td>
</tr>
<tr>
<td>GREB1 (GREB1 protein)</td>
<td>205862_s_at</td>
<td>2.8</td>
<td>0.000</td>
<td>1.9</td>
<td>0.000</td>
</tr>
<tr>
<td>SERPINA1 (Serine (or cysteine) proteinase inhibitor, clade A)</td>
<td>211429_s_at</td>
<td>2.3</td>
<td>0.000</td>
<td>2.5</td>
<td>0.000</td>
</tr>
<tr>
<td>SCGB1D2 (Secretoglobin, family 1D, member 2 (lipophilin B))</td>
<td>206799_at</td>
<td>3.6</td>
<td>0.000</td>
<td>2.3</td>
<td>0.000</td>
</tr>
<tr>
<td>SCGB2A2 (Secretoglobin, family 2A, member 2 (mammaglobin))</td>
<td>206378_at</td>
<td>2.9</td>
<td>0.000</td>
<td>4.2</td>
<td>0.000</td>
</tr>
<tr>
<td>SCGB1D2 (Secretoglobin, family 1D, member 2 (lipophilin B))</td>
<td>206378_at</td>
<td>2.9</td>
<td>0.000</td>
<td>4.2</td>
<td>0.000</td>
</tr>
</tbody>
</table>

www.endocrinology-journals.org
inflammatory bowel disease (Rodrigues et al. 2001, Wright et al. 1993). Amphiregulin is a secreted growth factor related to epithelial growth factor (EGF) and transforming growth factor α (TGFα). TFF1 or AREG are likely candidates for stimulating the paracrine secretion of growth factors which induce proliferation of ER-positive cells in response to systemic E. Indeed, Sternlicht et al. (2005) recently used AREG/C255 = C255 mice to suggest that AREG may have an essential role in the epithelial/stromal crosstalk that drives mammary development (Sternlicht et al. 2005). Some genes that are thought to be co-expressed with ESR1 were downregulated in E-treated xenografts, including RARRES1 and GATA3 (Lacroix & Leclercq 2004). In contrast, other genes that are thought to be co-expressed with ESR1 in breast cancer cell lines and tumours were upregulated, including XBP-1 (Lacroix & Leclercq 2004), GREB1 (Lin et al. 2004), elongation factor 1α2 (EEF1A2) and the little-studied C1orf34 (DEME-6), expression of which has been previously detected in primary breast carcinomas, but not in normal breast tissue (Kuang et al. 1998).

A second group of genes with increased expression due to E treatment (Table 1) were known diagnostic or prognostic markers, including lipophilin B (SCGB1D2), mammaglobin (SCGB2A2), prolactin-inducible peptide (PIP) and cytokeratin 19 (KRT19), which were recently used for a multi-gene reverse transcriptase PCR assay to detect circulating epithelial cells in the blood of patients with breast cancer (Ring et al. 2005). KRT19 is also a putative stem cell marker in the breast (Clarke et al. 2005, Gudjonsson et al. 2002). Anterior gradient 2 homologue (AGR2) was significantly upregulated in response to E2; the presence of detectable AGR2 mRNA in breast carcinoma cells is known to significantly correlate with carcinoma in preference to normal tissue, and ERα-positive in preference to ERα-negative carcinomas, suggesting that the metastasis-related properties of AGR2 may contribute, in some way, to the malignant progression of some ER-positive breast cancers (Liu et al. 2005).

A third category of differentially expressed genes was identified that appear to be involved with breast tissue microenvironment maintenance (Table 1). Serpins A1 and A3 were highly upregulated; serpin A1 encodes an anti-trypsin proteinase inhibitor which plays a key role in the control of homeostasis by neutralizing the deleterious effects of neutrophil elastase. Inhibition of serine proteinases

### Table 1

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>Probeset ID</th>
<th>FC MAS</th>
<th>pfp MAS</th>
<th>FC RMA</th>
<th>pfp RMA</th>
</tr>
</thead>
<tbody>
<tr>
<td>SERPINH1 †</td>
<td>Serine (or cysteine) proteinase inhibitor, clade H (hsp 47)</td>
<td>207714_s_at</td>
<td>0.5</td>
<td>0.073</td>
<td>0.8</td>
<td>0.061</td>
</tr>
<tr>
<td>RGS5 †</td>
<td>Regulator of G-protein signalling 5</td>
<td>209071_s_at</td>
<td>0.5</td>
<td>0.074</td>
<td>0.7</td>
<td>0.002</td>
</tr>
<tr>
<td>IL8 †</td>
<td>Interleukin 8 (CXCL8)</td>
<td>202859_x_at</td>
<td>0.4</td>
<td>0.003</td>
<td>0.7</td>
<td>0.001</td>
</tr>
<tr>
<td>IGHA1</td>
<td>Immunoglobulin heavy constant α1</td>
<td>211430_s_at</td>
<td>0.4</td>
<td>0.005</td>
<td>0.8</td>
<td>0.000</td>
</tr>
<tr>
<td>BGN †</td>
<td>Biglycan</td>
<td>201262_s_at</td>
<td>0.3</td>
<td>0.000</td>
<td>0.7</td>
<td>0.001</td>
</tr>
</tbody>
</table>

† Known E2-responsive genes from previous studies of breast cancer cell lines; ‡ Known prognostic or metastatic genes; † Putative microenvironment maintenance genes; see text.

### Table 2

Possible cross-species hybridization of mouse RNA to human GeneChips. Probesets detected when mouse RNA was hybridized to human chips that were also identified as significantly differentially expressed in response to E (Table 2). H, human; M, mouse.

<table>
<thead>
<tr>
<th>Affymetrix GeneChip</th>
<th>Gene ID</th>
<th>Nucleotide similarity, H/M</th>
<th>U95A versus mouse</th>
<th>U133 versus mouse</th>
<th>U95 versus human</th>
<th>U133 versus human</th>
<th>Transcript identified in E2 experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>U133</td>
<td>U95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>215906_at</td>
<td>31565_at (S65921)</td>
<td>†</td>
<td>16/16</td>
<td>10/11</td>
<td>4/16</td>
<td>3/11</td>
<td>Mouse</td>
</tr>
<tr>
<td>211204_at</td>
<td>33644_at</td>
<td>ME1</td>
<td>165/1832 (90%)</td>
<td>16/16</td>
<td>11/11</td>
<td>16/16</td>
<td>11/11</td>
</tr>
<tr>
<td>204540_at</td>
<td>35174_i_at</td>
<td>EEF1A2</td>
<td>1241/1392 (89%)</td>
<td>0/16</td>
<td>0/11</td>
<td>4/16</td>
<td>11/11</td>
</tr>
</tbody>
</table>

† The full target consensus sequence is derived from the mouse nucleotide sequence; ‡ Many non-specific nucleotides in the consensus/target sequence, which includes the poly-A tail and 3’ untranslated sequence.
may explain the action of serpin A1 in the suppression of cell growth in breast cancer cells, because it prevents the proteolytic release of membrane-bound TGFβ (Kang et al. 2005). This model would predict that serpin A1 can act as a tumour suppressor in inhibiting the growth of breast cancer cells (Yavelow et al. 1997). Genes involved with the generation of vascular stroma, which is thought to be essential for solid tumour growth, were downregulated. These genes include collagen (COL6A1), perlecan (HSPG2, heparan sulphate proteoglycan), versican (CSPG2, chondroitin sulphate proteoglycan 2), biglycan (BGN) and fibronectin (FN1). These genes encode components of the extracellular matrix and basement membrane, are involved with stabilization of other molecules that are associated with fibroblasts and are overexpressed in breast cancer (Brown et al. 1999). A number of chemokines (CCL2, interleukin-8, CXCL9, CXCL11, CCL19) were also down-regulated by E2 in normal breast tissue. In mammary tissues, SDF1 staining is primarily seen in stromal cells and weakly in epithelial cells (Kang et al. 2005). CCL2 and interleukin-8 have previously been suggested to have promalignant activity (Ben-Baruch 2003).

Tissue heterogeneity

Unsupervised principle component analysis identified that there were greater differences in the overall gene expression profile between the E2-treated and untreated tissues, than between the two untreated samples (data not shown). Up to seven different tissue samples from before and after implantation were examined by morphometry for cellular content to look at the variability in breast tissue heterogeneity. The median breast tissue composition from individual samples was 87% connective tissue and stroma (interquartile range (IQ) range, 80–90%), 11% epithelium (IQ range, 8–14%) and 2% fat (IQ range, 0–5%). There were smaller differences in cellular composition after combining the xenograft samples for each woman (median, 86% stroma (IQ range, 83–87%); 11% epithelium (IQ range, 10–13%) and 3% fat (IQ range, 2–5%), as was performed to generate the RNA pools for gene-expression analysis. There was no significant difference between the median cellular compositions, before and after implantation or between treated and untreated samples (Mann–Whitney U test). Tissue blocks were not available for every single xenograft, but, based on the samples we did look at, we feel we can be confident that each pool of RNA extracted from the multiple xenografts is derived from comparable proportions of stroma, epithelium and fat. The proportions of different cell types in the xenografts is in the range previously published for pre-menopausal women (Hutson et al. 1985).

Immunohistochemistry

Proliferation of epithelial cells before and after treatment with E2 was measured by immunohistochemistry. The percentage of epithelial cells expressing Ki67 following E2 treatment (median, 3%; IQ range, 2–6%) was significantly higher (P < 0.0001; Mann–Whitney U test; Fig. 2A) than in untreated xenografts (median, 1%; IQ range, 0.5–2%), confirming that the current study was in accordance with the previously validated xenograft model (Clarke et al. 1997, Laidlaw et al. 1995). Immunohistochemical analysis also demonstrated that, as previously shown (Clarke et al. 1997), progesterone receptor (PgR) levels in epithelial cells were significantly (P < 0.0001 Mann–Whitney U test; Fig. 2B), increased in the E2-treated xenograft tissue (median, 19%; IQ range, 10–31%) compared with the untreated tissue (median, 6%; IQ range, 2–13%). Gene-expression levels of PgR were significantly increased following addition of E2; however, the representative probeset (208305_at) was not called present on a sufficient number of GeneChips, thus eliminating it from the analysis using our strict criteria. The reason for probesets not being called present in many of the untreated samples is likely to be a direct result of very low levels of PgR transcripts. Neither protein nor transcript levels of E receptor α (ESR1) were observed to be significantly different in the E2-treated xenografts. Only one of the previous studies on the effect of E on breast cancer cell lines reported a significant reduction in ESR1 transcript levels (Inoue et al. 2002). In the current study, genes known to be co-expressed with ESR1 in breast cancer were both upregulated (XBP-1 and GREB1) and downregulated (RARRES1 and GATA3) in E2-treated normal tissue. The protein level of TFF1 (pS2) was seen to be significantly increased, in agreement with its transcript levels (Fig. 2C).

Cross-hybridization of mouse RNA to human GeneChips

The potential for skewing the data due to contamination of human RNA with traces of mouse RNA...
Figure 2 Immunohistochemistry analysis of E2-treated and untreated normal breast tissue. Proliferation of epithelial cells following treatment with E2. (A) Proliferation using Ki67 staining; (B) progesterone receptor staining; (C) TFF1 (pS2) staining.
was investigated by hybridizing mouse RNA to human GeneChips. Of the 12,265 full-length human genes represented on the human U95 GeneChips, fewer than 6% of the probesets (685) were called present on three or more arrays. Affymetrix array comparison spreadsheets (http://www.affymetrix.com) were used to reconcile the 685 probesets from the U95 array with their counterparts on the U133A array (containing 22,823 probesets). Matches were found for 559 of the 685 probesets, which represents just 2.5% of the total number of probesets on the U133A GeneChips, compared with an average 55% (range, 43–61%) of probesets called present on the 18 GeneChips used to investigate the response to E2. Just three probesets (215906_at, 211204_at and 204540_at) were identified as having the potential to be affected by cross-species hybridization and significantly differentially expressed in response to E2 (see Table 2 and Fig. 3). Probeset 215906_at was designed against a consensus sequence thought to be homologous to mouse anti-colorectal carcinoma light-chain mRNA (S65921); however, the mouse transcript only partially overlaps with relatively short human cDNAs (AW605031, AW605037, AW605049 and AW934907) with the rest of the sequence made up of the mouse transcript. Consequently, only three out of the 11 probes exactly match the known human cDNAs, whereas 10 out of 11 probes match the mouse transcript, suggesting that the probeset more reliably detects the mouse transcript (Table 1). All probes from 204540_at exactly match human EEF1A2; hybridization of mouse RNA can be explained due to four out of 16 probes on the U95A GeneChip matching the orthologous human sequence for this gene. However, none of the probes on the U133A GeneChip match mouse EEF1A2, so it is highly likely that this probeset hybridizes to human transcripts (Table 2). All of the 211204_at probes exactly match human and mouse ME1 sequences. It was therefore not possible to determine whether the resulting signal for this probeset is due to the presence of human or mouse RNA (Table 2).

**Discussion**

While steroid receptor expression and epithelial cell proliferation occur in separate cells in the normal breast (Anderson et al. 1998), proliferating breast
tumour epithelial cells often express ESR1. This distinction in the biology between normal and tumour epithelium is apparent at an early point in breast tumourigenesis and it can be detected in premalignant lesions (Shokei et al. 1999). The differences in gene expression that we have observed in response to E2 in normal human breast tissue compared with the response previously described in tumours or cell lines may reflect these physiological changes.

Over 60% of the probesets identified as E2-responsive in (Table 2) with known Gene Ontology cellular component classifications are assigned as extracellular, suggesting that a major element of the response to E2 in normal tissue is devoted to signalling. These proteins include growth factors, matrix metalloproteases and protease inhibitors. Allinen et al. (2004) showed that the most dramatic and consistent changes in breast cancer progression occur in myoepithelial cells and myofibroblasts and the majority of differentially expressed genes encode secreted and cell-surface proteins. The breast microenvironment is increasingly being studied in the context of its effect on tumourigenesis and metastasis. Inflammatory cells, cytokines and chemokines have been suggested to play a key role in breast carcinoma (Ben-Baruch 2003). The present study suggests that growth factors, proteases and chemokines are also important in the response to E in normal breast. Two recent studies have used different methods to determine breast-tissue cell-type-specific genes. Jones et al. (2004) combined an immunomagnetic sorting protocol selecting with the luminal epithelial marker (EMA) and myoepithelial membrane antigen (CD10) with cDNA microarray hybridizations, whereas Allinen et al. (2004) combined cell sorting with serial analysis of gene expression. The majority of genes identified in this study as being involved with proliferation and metastasis (TFF1, AGR2, KRT19, MYB1) appear to be most associated with the luminal epithelium in the cell-type-specific studies; by contrast, the genes identified as being involved with tissue maintenance are most specifically expressed in the myoepithelium and stroma (SEPINA1, BGN, COL6A1, FN1, CCL2). Normal mammary tissue has an intact epithelial basement membrane, as shown by a continuous linear staining for collagen, laminin, perlecan and fibronectin. This staining is widely lost in the invasive carcinomas (Brown et al. 1999). While putative proliferation and metastasis genes appear to demonstrate increased expression in both breast cancer and normal epithelium in response to E, the tissue-maintenance genes have reduced expression in normal tissue, but have been shown to have increased expression in DCIS myoepithelium compared with normal myoepithelium (Allinen et al. 2004). Similarly, SEPINA1 was upregulated in normal tissue but downregulated in breast cancer (Yavelow et al. 1997). These changes in the expression of tissue maintenance genes may account for differences in the response to E by normal breast tissue during menstrual cycles and its role in uncontrolled epithelial proliferation and carcinogenesis.

Although every effort was made to only isolate RNA from the recovered implanted human tissue, we also looked at the possibility for contaminating traces of mouse RNA to hybridize with probesets designed to recognize human sequences. Reassuringly, our analysis is consistent with previous cross-species hybridization studies (Creighton et al. 2003, 2005) showing there is a minimal contribution to the observed expression levels from contaminating RNA.

We report the first global gene expression study to examine the effects of E on the epithelium and stroma of normal human breast tissue in vivo. Previously, we showed that treatment of xenografts with 2 mg E2 for 1 week induces epithelial proliferation equivalent to that seen during the luteal phase of the menstrual cycle (Clarke et al. 1997). The majority of genes that we have identified as E2-responsive in normal human tissue appear to be either connected to proliferation/migration/invasion or are involved with maintaining tissue architecture. However, at this point it is unclear whether the differences in the gene expression profiles observed between normal breast tissue and tumour cell lines in response to E are due to the difference between tumour normal cells or between tissue and cell lines. As with the studies of E2 response in tumour cell lines (Coser et al. 2003, Frasor et al. 2003, 2004, Inoue et al. 2002, Vendrell et al. 2004), we acknowledge that our list of E2-responsive genes is not absolute, but dependent upon the specific approach (including regimens of E2 treatment, microarray platforms and analysis methods). However, we believe that the genes we have identified in this study are representative of the true response to E in normal tissue. Further studies will be required to characterize the true differences and similarities in the response to E2 in normal and tumour tissues, which may provide much-needed clues as to how E as a risk factor leads to breast cancer development.
Acknowledgments

We thank our colleagues in the University of Manchester Department of Surgery for breast tissue sample collection and are grateful for the excellent technical assistance of Stuart Pepper, and Yvonne Hey of the Cancer Research UK Affymetrix core facility and Caroline Chadwick of the Biological Resource Unit at the Paterson Institute for Cancer Research. R B C, C J M and C L W are supported by Cancer Research UK. A H S is supported by the Genetic Innovation network and Breakthrough Breast Cancer. The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

References


Harvell DM, Richer JK, Allred DC, Sartorius CA & Horwitz KB 2006 Estradiol regulates different genes in human breast tumor xenografts compared to the identical cells in culture. Endocrinology 147 700–713.


