Integrative analysis of gene expression patterns predicts specific modulations of defined cell functions by estrogen and tamoxifen in MCF7 breast cancer cells.

Franck Gadal, Anna Starzec, Christophe Bozic, Céline Pillot-Brochet, Sophie Malinge, V. Ozanne, J. Vicenzi, Laurent Buffat, Gérard Y. Perret, François Iris, et al.

Integrative analysis of gene expression patterns predicts specific modulations of defined cell functions by estrogen and tamoxifen in MCF7 breast cancer cells

F Gadal[1], A Starzec[2], C Bozic[3], C Pillot-Brochet[3], S Malinge[3], V Ozanne[3], J Vicenzi[3], L Buffat[3], G Perret[2], F Iris[4] and **M Crepin[1]**

[1]Unité Inserm U 553, Hôpital St Louis, 75010 Paris, France and Université Paris 13, France

[2]Laboratoire de Pharmacologie Clinique et Expérimentale, Faculté de Médecine, Université Paris 13, 93017 Bobigny cedex, France

[3]Valigen, Tour Neptune, 92086 Paris La Défense, France

[4]Ecole Centrale de Paris, Grande voie des vignes 92295 Chatenay Malabry, France and Bio-Modeling Systems, 26 Rue St Lambert, Paris 15, France

(Requests for offprints should be addressed to F Gadal; Email: fgadal@club-internet.fr)

Abstract

To explore the mechanisms whereby estrogen and antiestrogen (tamoxifen (TAM)) can regulate breast cancer cell growth, we investigated gene expression changes in MCF7 cells treated with 17β-estradiol (E_2) and/or with 4-OH-TAM. The patterns of differential expression were determined by the ValiGen Gene IDentification (VGID) process, a subtractive hybridization approach combined with microarray validation screening. Their possible biologic consequences were evaluated by integrative data analysis. Over 1000 cDNA inserts were isolated and subsequently cloned, sequenced and analyzed against nucleotide and protein databases (NT/NR/EST) with BLAST software. We revealed that E_2 induced differential expression of 279 known and 28 unknown sequences, whereas TAM affected the expression of 286 known and 14 unknown sequences. Integrative data analysis singled out a set of 32 differentially expressed genes apparently involved in broad cellular mechanisms. The presence of E_2 modulated the expression patterns of 23 genes involved in anchors and junction remodeling; extracellular matrix (ECM) degradation; cell cycle progression, including G_1/S check point and S-phase regulation; and synthesis of genotoxic metabolites. In tumor cells, these four mechanisms are associated with the acquisition of a motile and invasive phenotype. TAM partly reversed the E_2-induced differential expression patterns and consequently restored most of the biologic functions deregulated by E_2, except the mechanisms associated with cell cycle progression. Furthermore, we found that TAM affects the expression of nine additional genes associated with cytoskeletal remodeling, DNA repair, active estrogen receptor formation and growth factor synthesis, and mitogenic pathways. These modulatory effects of E_2 and TAM upon the gene expression patterns identified here could explain some of the mechanisms associated with the acquisition of a more aggressive phenotype by breast cancer cells, such as E_2-independent growth and TAM resistance.

Journal of Molecular Endocrinology (2005) **34,** 61–75

Introduction

Estrogens play several major roles in mammalian physiology, including the control of the reproductive tract and development of secondary sex organs. They regulate the estrous cycle, control lactation in the mammary gland and affect bone, liver and cardiovascular systems (Sutherland *et al.* 1988, Couse & Korach 1999). In diseases such as breast cancer, estrogens act as mitogenic factors and are important in tumor initiation and progression *in vitro* and *in vivo* (Lippman *et al.* 1975, Soule & McGrath 1980, Key & Pike 1988, Colditz 1998).

Studies on estrogen receptor (ER)-positive breast cancer cell lines indicate that estrogens and antiestrogens act on cell populations in early to mid-G_1 phase (Sutherland *et al.* 1983, Leung & Potter 1987). It is commonly accepted that tamoxifen (TAM) binds to the ER and inhibits ER-mediated gene transcription. Although antiestrogens are widely employed in the treatment of hormone-responsive breast cancer to induce cell growth arrest, the mechanisms by which TAM or its derivatives regulate gene expression are not well understood. Understanding of these mechanisms could allow us to explain, at least in part, the development of cellular resistance to TAM treatment and consequent therapeutic escape (Gibson *et al.* 1990, Katzenellenbogen 1991, Osborne *et al.* 1995, El Etreby & Liang 1998).

To explore, at gene expression level, how estrogen and antiestrogen could regulate cell growth and tumor progression in breast cancer, we investigated the

DOI: 10.1677/jme.1.01631
0952–5041/05/034–061
Online version via http://www.endocrinology-journals.org

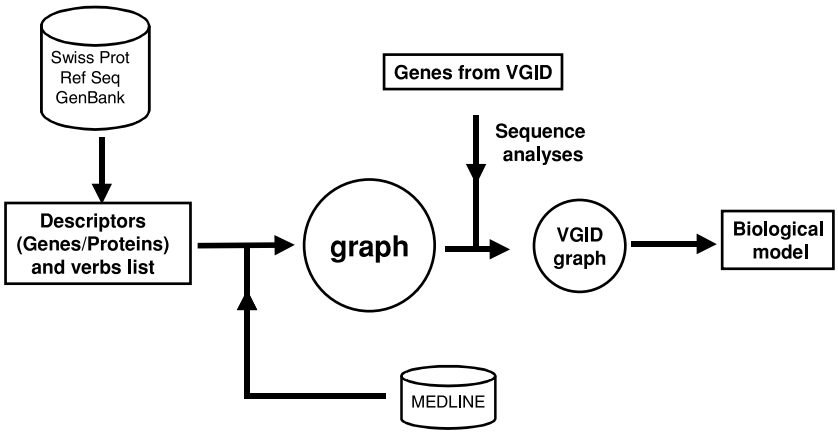

Figure 1 Global strategy to identify the biologic paths associated with differentially expressed genes. Using unified lists of descriptors and verbal forms, the contents of literature databases were parsed into a massive relational graph. Differential expression data, generated with VGID methodology and independently verified by microarrays, were then injected into the graph. The resulting subgraph was then mined to reveal the functional relationships linking the injected data with previously reported cellular and physiologic mechanisms. The resulting 'interaction maps' were then merged, by an iterative negative selection procedure, into a biologic model, which was then subjected to direct experimental evaluation.

patterns of gene expression in MCF7 cells treated with TAM and/or with 17β-estradiol (E_2). The MCF7 cells, isolated from a pleural effusion of metastatic human breast adenocarcima, are ER and progesterone receptor (PR) positive. Their growth is estrogen dependent and sensitive to TAM (Soule *et al.* 1973, Sommers *et al.* 1990).

The patterns of differential gene expression associated with E_2-induced tumor progression and with TAM antitumor action were analyzed by the ValiGen Gene IDentification (VGID) process. This process is based on a subtractive hybridization method combined with microarray and integrative data analyses that have previously been described (Gadal *et al.* 2003). The subtractive hybridization method does not require specific probes or primers to isolate differentially expressed mRNAs, thus allowing identification of isoforms and unknown transcripts without having to predetermine the sequences to be investigated. Microarrays provide a rapid and high throughput screening method for validating the differentially expressed status of sequences isolated by the subtractive hybridization process. Differential expression data and extensive published information were then integrated into a detailed biologic model. First, we generated a relational graph encompassing all known functional interactions between genes, proteins and small molecules recorded in the scientific literature pertaining to mammalian cellular and physiologic mechanisms. Then, the entire differential expression data set was injected into the graph in order to extract a subgraph functionally correlating the data sets. This enabled us to construct physiologic/signaling interaction maps directly correlated with gene expression patterns and protein interactions (Fig. 1). For more details, see Gadal *et al.* 2003 and our web page: http://perso.club-internet.fr/fgadal. The resulting model points out some of the mechanisms through which E_2 and TAM could affect the initiation, growth and progression of breast cancer.

Thus, by combining subtractive hybridization, cDNA microarrays and data integration, we were able to highlight some of the interconnected mechanisms, such as cytoskeleton remodeling, matrix degradation, cell cycle progression, synthesis of genotoxic metabolites, DNA repair and estrogen receptor constitution/activity, through which estrogen and antiestrogen regulate tumor progression.

Materials and methods

Cells and cell culture

The MCF7 breast cancer cell line, obtained from ATCC (Rockville, MD, USA), was cultured routinely in Glutamax I Dulbecco's modified Eagle's medium (DMEM) (Gibco Invitogen) supplemented with 10% fetal calf serum (FCS) (Gibco Invitogen), 50 μg/ml streptomycin (Gibco Invitogen) and 50 UI/ml penicillin (Gibco Invitogen) at 37 ° C in 5% CO_2 atmosphere. The experiments were carried out in phenol red-free DMEM (Gibco Invitogen) supplemented with 10% charcoal-treated FCS, 50 μg/ml streptomycin, 50 UI/ml penicillin and 2 mM L-glutamin (Gibco Invitogen). The cells, divided into three batches, were incubated at 37 °C in 5% CO_2 atmosphere for 8 days. Then, one batch of

MCF7 cells was stimulated with 1×10^{-8} M E_2 (Sigma Aldrich) for 15 h, whereas the second batch was maintained in E_2-free medium. To explore E_2-associated differential expression, the VGID procedure was performed with cDNAs generated from batches 1 and 2. To explore TAM-associated differential expression, the third batch of MCF7 cells was incubated with 1×10^{-8} M E_2 and 1×10^{-7} M TAM (Sigma Aldrich) for 15 h, and the VGID procedure was performed with cDNAs generated from batches 2 and 3.

cDNA synthesis

Direct mRNA capture from lysed cells was performed with the Dynabeads mRNA direct kit (Dynal, Compiégne, France) according to the manufacturer's protocol. First-strand cDNA synthesis was carried out as follows: 6 µl aliquots of first-strand synthesis mixure (10 mM dNTP each (Promega), 0·1 M DTT and 5 first-strand buffer (both from Life Technologies, Cergy Pontoise, France)), 500 ng Dynabeads-bound denatured mRNA and 200 units Superscript II (Life Technologies) were separately preincubated at 37 °C for 1 min in a water bath, and then mixed and incubated in a final volume of 30 µl at 37 °C for 1 h. The first-strand samples were then transferred to 170 µl diluted second-strand synthesis mixure (1 final *E. coli* DNA ligase buffer (New England Bio-Labs, Saint Quentin Yveline, France), 900 mM KCI (Sigma Aldrich), 20 mg/ml glycogen (Boehringer Mannheim, Mannheim, Germany), 10 mM dNTP each, 10 U/µl *E. coli* DNA ligase (New England Bio-Labs), 2 U/µl RNase H and 10 U/µl *E. coli* DNA polymerase I, all purchased from Promega) and incubated for 2 h at 16 °C in a PerkinElmer (Courtaboeuf, France) 9700 thermal cycler. The cDNA populations, still bound to Dynabeads, were then digested with Sau3A (4 U/µl; New England Bio-Labs) for 2 h at 37 °C to generate the fragments of 256 base pairs in statistical length. The reaction was stopped by heating for 20 min at 65 °C. The digested cDNAs were then ligated to adapters, and PCR was performed with adapter-specific primers in a PerkinElmer 9700 thermal cycler, using 12 cycles of 95 °C for 30 s, 55 °C for 45 s and 72 °C for 1 min, followed by 72 °C for 7 min. The PCR products were purified by the Qiaquick PCR purification kit (Qiagen) and used for VGID and/or microarray experiments (see below).

Adapters and oligonucleotides

Sequences of adapters (CyberGene, Saint Malo, France) used in VGID procedure were as follows: BamHI adapter, 5′CTT AGA ACG AGA CGG ATC CT3′ and 3′TT GAA TCT TGC TCT GCC TAG GAC TAGp5′; BglII-bio adapter, bio 5′CCA GCT AAC ACC TAG ATC TC3′ and 3′TT GGT CGA TTG TGG ATC TAG AGC TAGp5′. The fragments of cDNA ligated to these adapters were amplified with the corresponding oligonucleotides (CyberGene): BamHI primer, 5′AA CTT AGA ACG AGA CGG ATC CTG ATC3′ and BglII-bio primer, bio 5′AA CCA GCT AAC ACC TAG ATC TCG ATC3′. The inserts were amplified with the primers pVG17 (–74), 5′GCA AGG CGA TTA AGT TGG GTA3′ and pVG17 reverse (–88), 5′CTT CCG GCT CGT ATG TTG TGT3′.

VGID gene identification technology (patent no. 6221585)

As previously described (Gadal *et al.* 2003), VGID directly isolates overexpressed and underexpressed cDNA associated with the transition from a defined phenotypic state to another state within a congenic system. For the first denaturation–renaturation step, 300 ng BamHI-ligated-amplified tester were mixed with 1200 ng BglII-bio-ligated-amplified driver. The mixture was ethanol-precipitated, resuspended in 4 µl HEPES 0·5 M-EDTA 0·2 mM) and overlaid with 20 µl mineral oil, denatured 5 min at 98 °C and finally chilled on ice. The salt concentration was adjusted to 0·5 M with 1 µl of 2·5 M NaCl. After 5-min denaturation at 98 °C, the sample was allowed to anneal for 20 h at 65 °C. After hybridization, the oil was removed. The sample volume was adjusted to 100 µl with 100 mM NaCl and 10 mM Tris HCl, pH 8, and mixed with 210 µl streptavidin magnet beads (10 mg/ml, Boehringer Mannheim) to recover tester single- and double-strand DNA, as described by the manufacturer. This step was repeated to remove all biotinylated cDNA. To recover the tester, double-strand cDNA, the supernatant, which contained unbiotinylated cDNA, was incubated with 1·5 µg single-strand binding protein (SSB, Promega) for 30 min at room temperature. The sample was loaded onto a Millipore Micropure EZ membrane (which retained the proteins) and centrifuged for 1 min at 20 800 ***g*** at room temperature. The flow through was combined with 1200 ng driver, and the next round of hybridization was set up as described above. A total of three hybridization rounds were performed. The cDNA recovered after the last round, was amplified by PCR, as described above, using 25 cycles and the BamHI primer. The sequences under- or overexpressed were cloned into a modified pUC19 vector (Gibco Invitogen).

Cloning

The vector used for the cloning step was a derivative of the pUC19 vector in which polylinker EcoRI– HindIII was replaced by the 5′AATTCGGATCCA3′ pVG17 sequence. To avoid cloning chimeric structures, the vector/insert DNA ratio used was 3:1. Vector (50 ng) was ligated to the recovered cDNA and amplified after

subtraction in the presence of 1 U ligase (Boehringer Mannheim) in a final volume of 10 µl overnight at 16 °C. The ligation mixture was purified by the GeneClean procedure (BIO 101), according to the manufacturer's recommendation, and the 2 µl-aliquot was used to transform 50 µl MAX Efficiency DH10B cells (Gibco Invitogen). Transformation was performed by electroporation at 1800 V with the EC100 Electroporator (E-C Apparatus Corporation, Petersburg, FL, USA). Immediately after electroporation, 500 µl SOC medium were added. Cells were shaken for 1 h at 37 °C, and an aliquot was spread on LB medium plate containing 100 µg/ml ampicillin.

Sequencing

Isolated colonies were selected in 150 µl LB with 100 µg/ml ampicillin in order to make glycerol stocks. A volume of 4 µl of this solution was cycle-sequenced in a PerkinElmer thermal cycler with the AmpliTaq Gold DNA polymerase kit (Perkin Elmer), using 10 pmol pVG17(-74) and pVG17 rev(-88) primers in a final volume of 50 µl. A pre-PCR step, performed at 95 °C for 13 min, allowed the activation of the polymerase and the bacteria lysis. Each of the 40 PCR cycles included three segments: 95 °C/30 s, 54 °C/1 min and 72 °C/2 min 30 s. The PCR products were sequenced without purification with the BigDye Terminator Cycle Sequencing kit and ABI Prism 3700 automatic DNA sequencer (PerkinElmer Applied Biosystem).

Sequence analysis

The clone sequences were analyzed by in-house programs written in Perl. After removal of the vector sequence, the repeats were masked and compared with public databases with the BLASTn and BLASTx programs. The significance of the similarities was checked at both the nucleic and protein level. DNA sequences were considered for further analysis when the level of similarity with known sequences was greater than 98%. At the protein level, significant similarity threshold was fixed as at least 40. The sequence data were then clustered with Fasta software (Infobiogen, Evry, France). Unidentified genes (sequences failing to match anything in the databases, including expressed sequence tags (ESTs)) could not help us to construct the relational graph. Thus, these sequences were not introduced in the analysis.

Microarray construction

All the clones obtained from the subtraction process were studied with microarray technology. Each clone was PCR-amplified in a final volume of 100 µl, using 20 pmol of both primers pVG17 (–74) and pVG17 rev (–88), 10 nmol dNTPs and 1·5 U DyNazyme EXT. After 13-min enzyme activation at 95 °C, 40 cycles were carried out (95 °C/30 s, 56 °C/1 min and 72 °C/1 min 30 s). A final incubation was performed for 7 min at 72 °C. The PCR products were purified with the Multiscreen PCR Kit (Millipore, Saint Quentin Yveline, France), according to the manufacturer's instructions, and concentrated under Vacuum SpeedVac (E-C Apparatus). The PCR products, resuspended in 3 SSC, with 11 yeast genes used as internal control, were spotted in triplicate onto GAPS Amino Silane Coated Slides (Corning, Schiphol-Rijk, The Netherlands) with the GMS 417 arrayer (Genetic MicroSystem, Affymetrix, High Wycombe, Herts, UK). The slides were UV cross-linked at 300 mJ, prehybridized in 50% formamide, 0·1% SDS, 1% bovine serum albumin (BSA) and 5 SSC buffer at 42 °C for 1 h, washed first in water and then in 95% ethanol, and finally vacuum dried.

Microarray hybridization, scanning and data acquisition

cDNA (500 ng) from MCF7 cells treated or not with TAM or E_2 were labeled by random priming with incorporation of Cyanine5-dUTP for the tester DNA and Cyanine3-dUTP for the driver samples respectively. Then, the samples were mixed, concentrated by evaporation under vacuum and resuspended in prehybridization buffer, as described above, with Denhardt's solution replacing BSA. The two-labeled cDNA mixtures (MCF7 cDNA mixed with MCF7 + E_2 one and MCF7 + E_2 cDNA mixed with MCF7 + E_2 + TAM one) were hybridized with the arrayed slides overnight at 42 °C. The slides were then washed for 5 min with 1 SSC-0·1% SDS, 3 min with 1 SSC, 3 min with 0·1 SSC and 1 min with water, and finally 95% ethanol-dried and scanned (GenePix 4000A; AXON, Union City, CA, USA). Accurate differential measurements (final fluorescence ratios) were expressed as the average of nine independent assays where each sequence was arrayed in triplicate. Visualization, quantification and gene expression analysis were performed with GENEPIX 3·0 software (AXON). The data were normalized by the autonormalization method of Yang *et al.* (2000).

Assay of NAD(P)H quinone oxidoreductase and NADH-menadione oxidoreductase activity

MCF7 cells were cultured, as described above, with E_2. The MCF7 cells were harvested with a rubber policeman, centrifuged at 1000 ***g*** for 5 min at 4 °C and homogenized in cold buffer (100 mM potassium phosphate, pH 7·0, and 2 mM EDTA). After centrifugation at 10 000 ***g*** for 15 min at 4 °C, the supernatant containing cytosolic and microsomal enzymes was isolated. NAD(P)H quinone oxidoreductase activity was

determined with 2,4-dichlorophenol indophenol (DCPIP) (Sigma). Samples (10 μl) were mixed with 80 μl of 1 mM NAD(P)H (Sigma Aldrich) in a microtiter plate and incubated for 15 min at 37 °C. Then, 4 mM DCPIP (10 μl) was added to each well, and the absorbance was measured at 600 nm every 5 min. To assess NADH-menadione oxidoreductase activity, the samples (10 μl) were mixed with 70 μl of 10 μM menadione (Sigma Aldrich) and 10 μl of 40 μM cytochrome C (Sigma Aldrich). After determining the baseline at 550 nm, 10 μl of 0·5 mM NADH (Sigma Aldrich) was added, and the absorbance was measured every minute. Changes in absorbance were obtained by plotting the absorbance values as a function of time to determine the slope (rate) of the linear portion of the curve. Then, changes in absorbance corresponding to the activity of the enzymes were expressed as percentages.

Results

Differential gene expression patterns in MCF7 cells treated with E_2 or E_2 and TAM

To determine the sequences overexpressed in response to E_2 in MCF7 cells, the sequences contained in the E_2-treated MCF7 cDNA library (tester) were subjected to competitive hybridization against an excess amount of cDNA from untreated MCF7 cells (driver). This was followed by selective trapping of all driver material, including sequences common to both tester and driver. In order to obtain E_2-underexpressed sequences, the cDNA obtained from E_2-treated MCF7 cells was ligated to BglII-bio adapter to become the driver library, whereas the cDNA from untreated MCF7 cells was ligated to BamHI adapter to form the tester library. Competitive hybridizations were then performed.

In the TAM experiments, the E_2 + TAM-treated MCF7 cDNA library (tester) was subjected to competitive hybridization against an excess amount of cDNA from E_2-treated MCF7 cells (driver). Sequences underexpressed in response to TAM treatment were obtained by inversing the adapter-dependent tester:driver identification tagging, as described above.

All competitive hybridizations were performed with a driver/tester ratio of 4, which allowed optimal recovery of sequences differentially expressed (data not shown). Three iterative rounds of subtraction and selective trapping resulted in high enrichment and efficient isolation of tester-specific sequences, as visualized by migration on agarose gel (Fig. 2). The electrophoretic profiles were composed of smears, ranging from 100 to 1250 pb, as well as characteristic bands corresponding to cDNAs underrepresented (lanes 1 and 3) and overrepresented (lanes 2 and 4) in MCF7 cells treated with E_2 alone (lanes 1 and 2) or in combination with TAM (lanes 3 and 4). E_2 treatment resulted in the identification of three different major cDNAs bands, observed in both lane 1 (underrepresented cDNAs) and lane 2 (overrepresented cDNAs) while treatment with E_2 + TAM led to the identification of nine differentially expressed sequences, five of which were underexpressed (lane 3). All these differentially expressed sequences were characterized by distinct lengths, suggesting that they reflect different gene expression patterns, which could be specifically associated with differences in the physiologic states of the cells after estrogen or antiestrogen treatments.

All the cDNAs recovered after subtractive hybridization were then cloned in a modified Puc vector. To avoid preferential cloning of smaller cDNAs over larger fragments, the VGID outputs were electrophoretically separated on agarose gels, which were then divided into four parts to separate the cDNAs into size populations. The first group corresponded to cDNAs larger than 700 pb. The sizes of the second and third cDNA groups were 500–700 pb and 300–500 pb respectively. The last section was composed of cDNA fragments less than 300 pb in length. Overall, the subtractive hybridization processes yielded 1005 cDNA inserts, ranging in size from 100 to 1250 bp. Sequence analysis (Table 1) revealed that treatment with E_2 resulted in the differential expression of 279 known and 28 unknown sequences, whereas treatment with E_2 + TAM resulted in the differential expression of 286 known and 14 unknown sequences. Redundant identification of known sequences was relatively low, corresponding to 28–40% of the datasets, depending upon the VGID experiments (Table 1). This strategy allowed us to identify a wide panel of genes involved in different cell functions affected by E_2 and TAM treatments. Here, redundant identification of sequences/genes did indicate relative expression level, which could then be validated by microarray analysis. For example (Table 2), the pS2 sequence was represented by 50 clones, suggesting substantial overexpression of the gene, whereas the upstream factor 2 (USF2) sequence was found only once, arguing for moderate to low overexpression. Indeed, microarray analysis showed that expression of pS2 and USF2 was enhanced by 4·93- and 1·44-fold respectively.

To confirm the differential expression status of the sequences isolated by VGID and obtain quantitative information, the cDNA insert of all the clones, known as well as unknown nonredundant sequences (607 sequences), were PCR amplified and arrayed on slides. Because of the existing overlap between the two data sets, the 147 genes subject to differential regulation in response to both E_2 and TAM were represented twice on the same microarray. The microarrays thus constructed, representing a total of 460 different probes, were then independently hybridized to each of the different mRNA populations corresponding to our experimental conditions. Thus, when verifying the

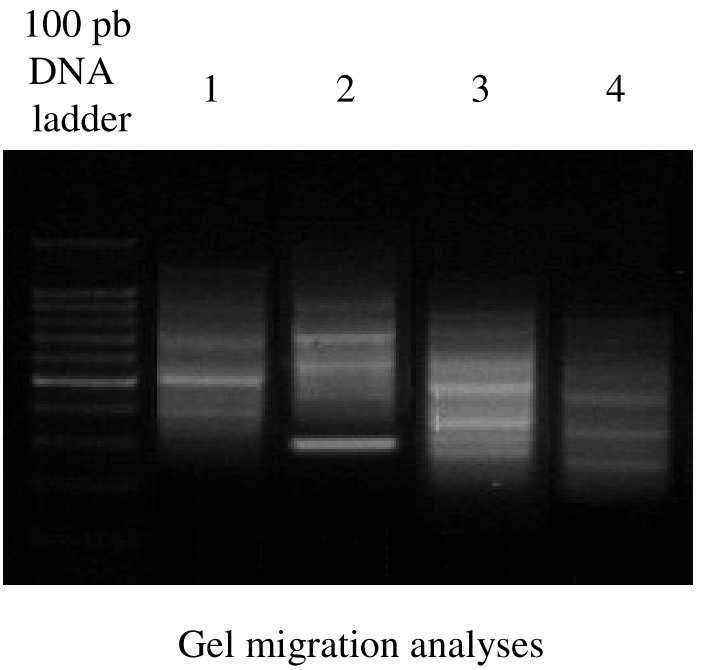

Figure 2 Isolation of transcripts differentially expressed in MCF7 cells. Gel migration analyses (1·5% agarose) were performed at each step of the VGID process. Lanes 2 and 4 correspond respectively to cDNA samples of sequences overrepresented in the MCF7 + E_2 and MCF7 + E_2 + TAM libraries. For identification of the transcripts underexpressed in MCF7 + E_2 and MCF7 + E_2 + TAM, the process was repeated, using the MCF7 and the MCF7 + E_2 cDNA libraries as tester respectively. Thus, lanes 1 and 3 represent sequences underrepresented in the MCF7 + E_2 and the MCF7 + E_2 + TAM cDNA libraries after differential trapping.

differential responses obtained after estrogen treatment, for instance, the microarrays utilized contained not only all the genes identified as estrogen responsive after VGID analysis, but also all those identified as TAM responsive. Multiple randomly selected cDNAs were also spotted on the same slides to serve as internal controls. Post-hybridization analyses were based on the autonormalization method of Yang *et al.* (2000). For example, the cDNA obtained from E_2-treated MCF7 cells was labeled with Cy-5 fluorochrome (red); that from untreated MCF7 cells with Cy-3 fluorochrome (green). After hybridization (Fig. 2), red and green fluorescence indicated greater relative expression in the presence and in the absence of E_2 respectively. Yellow fluorescence corresponded to equivalent expression levels. Microarray analyses confirmed both the patterns and levels of differential expression of the majority (~90%) of the genes identified as differentially expressed through

Table 1 Distribution of the MCF7 differentially expressed sequences to response in E_2 and TAM treatments. cDNA transcripts obtained by subtractive hybridizations were cloned into a modified Puc vector and sequenced. Gene identification was performed with BLAST software against nucleotide and protein banks.

	VGID subtractive hybridization			
	Under-expressed with E_2	Over-expressed with E_2	Under-expressed with TAM	Over-expressed with TAM
Number of clones sequenced	280	253	236	236
Number of exploitable sequences	263	233	214	230
Number of known sequences	237 (142 non-redundant)	219 (137 non-redundant)	207 (148 non-redundant)	221 (138 non-redundant)
Number of unknown sequences	23 (20 non-redundant)	8 (8 non-redundant)	7 (5 non-redundant)	9 (non-redundant)

Table 2 Function and expression level of E_2-inducible genes. The levels of expression indicate over-expression (upward arrows) and under-expression (downward arrows). These expression patterns were qualitatively and quantitatively verified by microarray expression analyses. Relative expression levels are indicated by arrows and x-fold. The last column indicates the genes whose expression is reversed by TAM treatment in MCF7 cells. In the 'TAM inducible gene expression switching' column. "–" indicates that no switching effects were experimentally observed. Data for all VGID-identified genes can be accessed at the following website address: http://perso.club-internet.fr/fgadal

	E_2-responsive genes	Accession number	VGID identification status	Number of clone corresponding to identified genes	Microarray status (X fold expressed)	TAM-inducible gene expression switching
Broad functions						
Cytoskeletal	Keratin 8	U76549	↗	8	↘1·44±0·02	↗2·03±0·07K
remodelling	Keratin 18	M26326	↗	17	↘1·72±0·01	↗1·43±0·03
C	Claudin 4	NM_001305	↗	3	↘1·74±0·01	↗1·56±0·05
J	Junction plakoglobin	AAH00441	↗	2	↘3·47±0·07	–
A	Actin gamma	NM_001614	↗	1	↘1·29±0·05	↗1·58±0·02
A	Arp 2/3 complex, subunit 1B (41 kDa)	AF006084	↗	1	↘1·60±0·03	–
B	Beta catenin	X89448	↗	1	↘1·63±0·04	–
	Hsp27	X54079	↗	2	↗1·90±0·10	↘1·53±0·01
ECM remodelling	USF2	Y07661	↗	1	↘1·44±0·04	–
	HGFA 12	AB006534	↗	1	↘1·32±0·03	↗1·31±0·06
	pS2 protein	X00474	↗	50	↗4·93±0·03	↘1·27±0·02
	Cathepsin D	X05344	↗	14	↗3·16±0·16	↘1·10±0·05
Cell-cycle progression	TGF-β	M60316	↗	4	↘4·09±0·03	–
	e-myc oncogene	V00568	↗	1	↗1·57±0·02	–
	Nm23-H2	L16785	↗	2	↗2·00±0·02	–
	Cks1	NM_001826	↗	3	↗2·42±0·05	–
	CDK2	X61622	↗	1	↗1·36±0·05	–
	B-myb	X13293	↗	1	↗2·28±0·02	–
	ASK	NM_006716	↗	1	↗2·06±0·09	–
Cetechols/quinone	NAD(P)H menadione oxidoreductase	J03934	↗	2	↘1·30±0·05	↗1·28±0·01
synthesis regulation	UDP-glucuronosyltransferase	M57951	↗	1	↘1·20±0·10	–
	Thiol-specific antioxidant	Z22548	↗	2	↘1·56±0·02	↗1·18±0·04
	NADH dehydorgenase ubiquinone	NM_005006	↗	4	↗1·91±0·09	↘1·33±0·01

Table 3 Function and expression level of TAM-responsive genes. The levels of expression indicate over-expression (upward arrows) and under-expression (downward arrows). These expression patterns were qualitatively and quantitatively verified by microarray expression analyses. In the 'microarrays status' column '0' indicates failure to detect differential expression. Relative expression levels are indicated by arrows and x-fold.

	TAM-responsive genes	Accession number	VGID™ identification status	Number of clone corresponding to identified genes	Microarray status (X fold expressed)
Broad functions					
Cytoskeletal remodelling	Profilin	J03191	↗	1	↗1·32±0·05
Cetechols/quinone synthesis regulation	HSD17 B4	NM_000414	↗	1	↗1·24±0·02
DNA repair	DDB-2	U18300	↘	1	↘1·27±0·04
	XR-CC1	L34079	↘	1	↘1·44±0·01
Active ER constitution	Hsp90	X15183	↘	1	↘1·37±0·06
	Immunophlin	M88279	↘	1	0
	REA	AF150962	↗	1	↗1·22±0·02
Mitogenic pathway	ERK1	X60188	↗	1	↗1·49±0·01
Growth factor	TGF-α	AH003012	↘	1	↘2·14±0·01

utilization of VGID, except for immunophilin (Table 3) (entire data set available at http://perso.club-internet.fr/fgadal).

Integrative analysis of the information embodied in the literature, combined with the VGID and microarray results, singled out a set of 32 differentially expressed genes coding for proteins associated with broad cellular functions (Tables 2 and 3). The expression status (under- or overexpressed), as indicated by the VGID outputs for 31 of these identified genes, was corroborated by microarray analysis. In our dataset, eight genes were found to be differentially expressed by a factor superior to twofold, nine by a factor of 1·50–2·00 and 14 by a factor inferior to 1·50. These results indicate that the VGID procedure is sensitive enough to isolate transcripts the expression level of which is only slightly modified.

E_2 modulates biologic functions in MCF7

Integrative analysis indicated that, in MCF7 cells, E_2 treatment seemed to affect (Table 2):

(i) Anchors and junctions remodeling. Seven of the eight genes identified as playing significant roles in these mechanisms were repressed, including β catenin, junction plakoglobin, γ-actin-1, claudin 4, the Arp2/3 complex involved in adherent junctions, and keratins 8 and 18, implicated with junction plakoglobin in desmosomes constitution. Only Hsp27, which plays a role in adherent junctions, was found to be upregulated.

(ii) Extracellular matrix (ECM) degradation. In ECM degradation, underexpression of hepatocyte growth factor activator inhibitor 2 (HAI-2) and USF2, together with overexpression of cathepsin D and pS2, argued for an important role of E_2 in the activation of proinvasive mechanisms through matrice remodeling.

(iii) Cell-cycle progression. Overexpressed c-myc, nm 23-H2 nucleotide diphosphate kinase, cyclin kinase subunit 1 (Cks1), CDK2, B-myb, activator of S phase kinase (ASK) and underexpressed TGFβ are involved in cell-cycle progression, in particular with the G_1/S checkpoint and in S-phase progression (Fig. 3).

(iv) Synthesis of genotoxic metabolites. Genes involved in detoxification mechanisms, such as UDP glucuronosyltransferase, NAD(P)H menadione oxidoreductase and thiol-specific antioxidant, were found to be underexpressed, whereas NAD(P)H dehydrogenase ubiquinone was overexpressed in MCF7 cells after E_2 treatment, suggesting a decrease in the effciency of at least one important detoxification pathway.

E_2 decreases the activity of both NAD(P)H menadione oxidoreductase and NAD(P)H quinone oxidoreductase

To verify our conclusion that E_2-treatment induced modifications in the activities of the oxidant-dependent and genotoxic quinone-dependent pathways (Fig. 4), we measured the activities of two key enzymes, NAD(P)H menadione oxidoreductase and NAD(P)H quinone oxidoreductase. In E_2-treated cells, we observed that activities of both enzymes were reduced by $11 \pm 2\%$ and $30 \pm 6\%$ ($P<0·05$) respectively, as compared with untreated MCF7 cells.

TAM antagonizes only some effects of E_2 upon gene expression in MCF7 cells

Our integrative data analysis indicated that TAM treatment reverses E_2-induced changes in the expression of only 11 genes and apparently restores only some of the biologic functions affected by E_2 (Table 2).

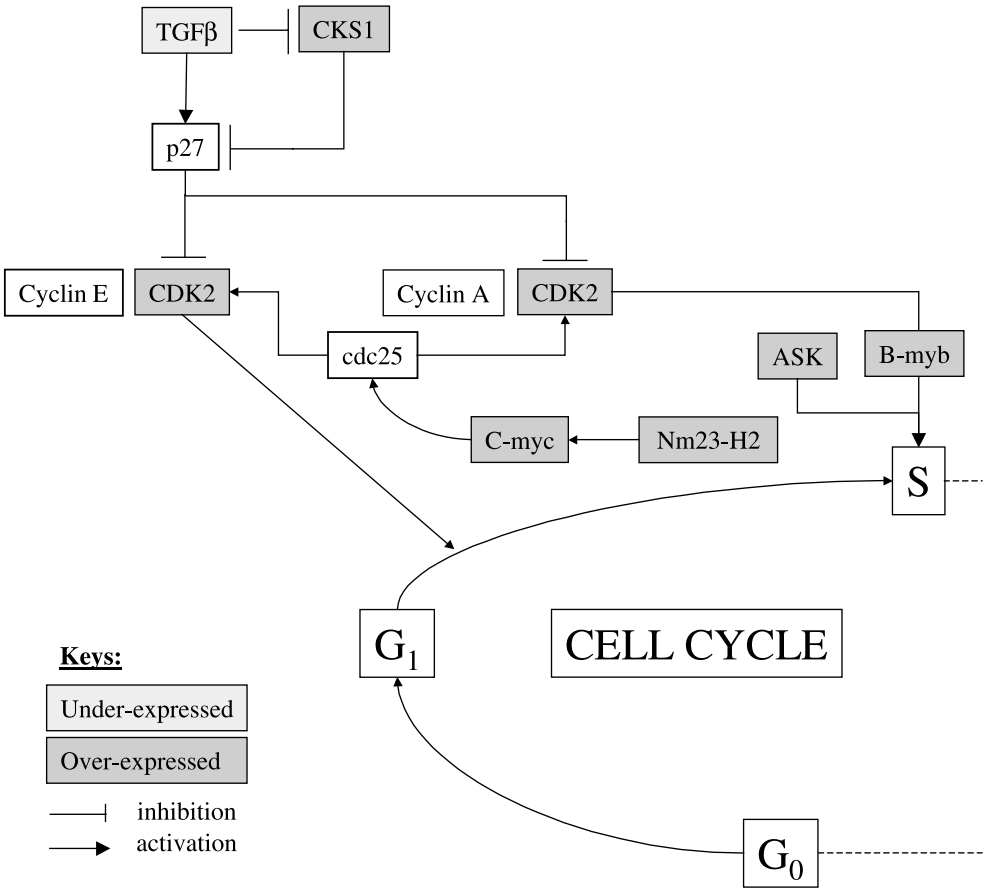

Figure 3 Effects of E_2 on the differential expression of genes regulating cell-cycle progression. Boxes with plain color represent genes found to be differentially expressed by the VGID process and confirmed by cDNA microarray analysis.

Increases in the expression of claudin 4, γ-actin-1, keratin 8, keratin 18 and decrease in the expression of Hsp27 suggest that TAM treatment could restore anchoring and junction structures.

Switching in HGFA I2, cathepsin D and pS2 expression suggests that TAM could restore protection of ECM integrity.

Expression switching of NAD(P)H menadione oxidoreductase, thiol-specific antioxidant, and NADH dehydrogenase ubiquinone argues for a possible restoration of detoxification mechanisms, including a decrease in E_2-induced quinones synthesis, which could prevent catechols/quinone genotoxic effects.

Interestingly, we did not find any change in the expression of genes involved in cell-cycle progression previously affected by E_2. Furthermore, we found (Table 3) that TAM treatment specifically decreases the expression of the DDB-2 and XR-CC1 genes involved in DNA repair. TAM treatment also inhibits the expression of Hsp90 and immunophilin while increasing that of REA. These three gene products are known to play a role in the formation of active ER. Concurrently, TAM treatment induces the overexpression of profilin, which is associated with actin dynamics. Finally, TAM treatment induces underexpression of TGFα and overexpression of ERK1 and hydroxysteroid (17-beta) dehydrogenase 4 (HSD17β4).

Discussion

In the present study, we used the analytic approach we had previously described (Gadal *et al.*) to identify the genes differentially expressed in MCF7 breast cancer cells stimulated with E_2 or treated with TAM. With this approach, the capture of biotinylated cDNA and single-strand cDNA is very effective, thereby avoiding the construction of normalized cDNA libraries and increasing significantly the number of differentially regulated genes correctly identified. The differential expression pattern of over 90% of the sequences thus isolated was confirmed by microarray analyses. These

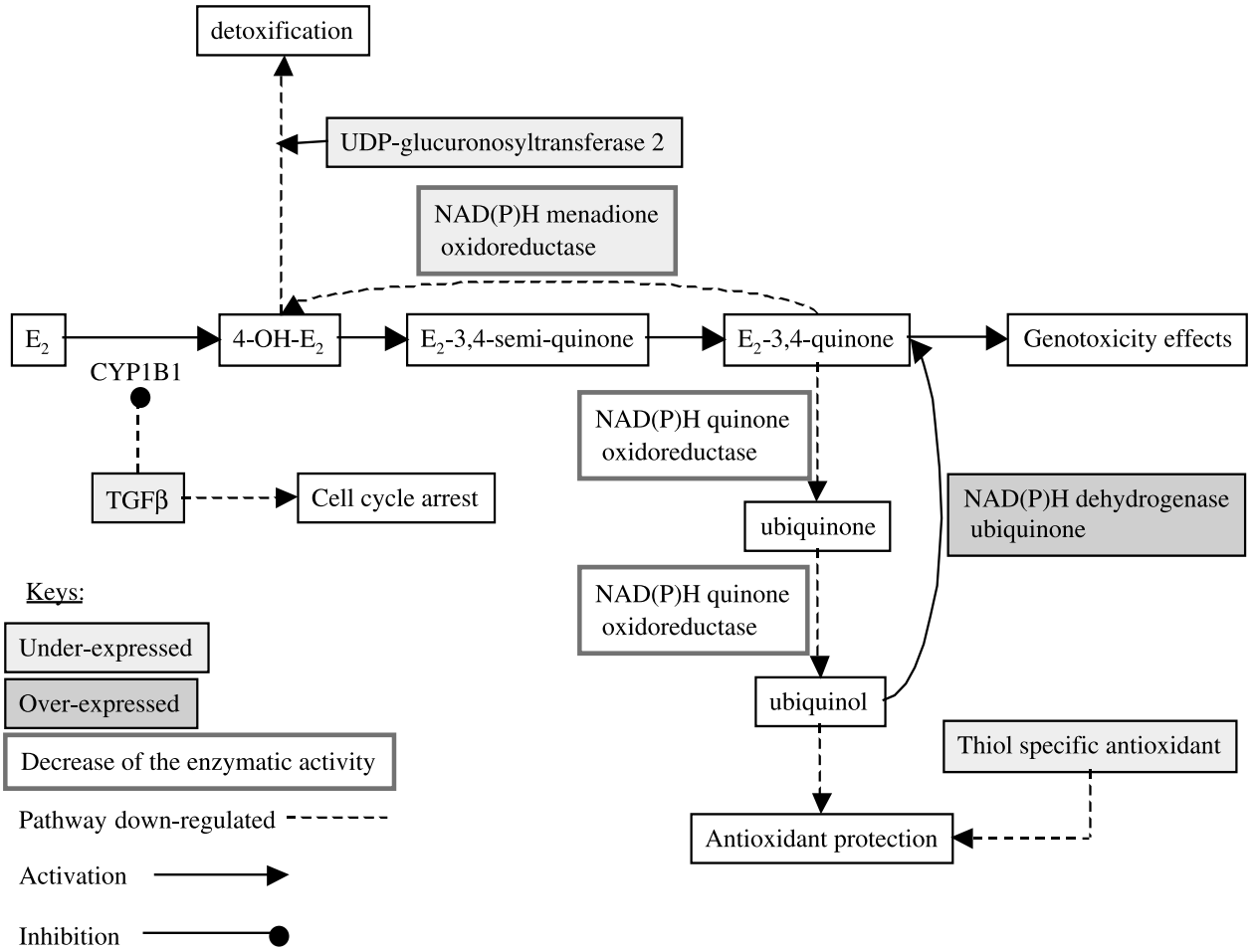

Figure 4 Evidence of E_2 genotoxic effects in MCF7 cells. Boxes with plain color represent genes found to be differentially expressed by the VGID process and confirmed by cDNA microarray analysis.

techniques have been significantly improved over the past 4 years, and it is now widely accepted that they can reliably be used to verify accurately and independently the relevance of differential expression data generated by new technologies, such as the VGID approach, for instance. For comparison, with Suppression Subtractive Hybridization (SSH) as the primary means to detect differential expression, only 70% of identified genes were confirmed by a microarrays approach similar to that described above (Kuang *et al.* 1998). Furthermore, and in opposition to high-density microarrays, the VGID strategy avoids the introduction of analytic bias resulting from the *a priori* selection of genes to be investigated. We then carried out computer-assisted data integration to achieve an understanding of the physiologic mechanisms potentially affected by differential expression of the identified genes. Initially, the sets of genes identified as estrogen or TAM responsive were completely separate because they were independently obtained from different RNA populations. Subsequent data analysis showed that some genes were differentially modulated under each of our two experimental conditions, thereby indicating partial overlap between the two datasets. The VGID data sets included genes previously shown to be regulated in response to E_2 or TAM, such as Hsp27 (Porter *et al.* 2001), pS2 (Inadera *et al.* 2000), cathepsin D (Inadera *et al.* 2000, Safe 2001), c-myc (Doisneau-Sixou *et al.* 2003), ERK1 (Rabenoelina *et al.* 2002) and β-catenin (Gunin *et al.* 2003), thereby substantiating the validity of our experimental approach. The fact that only immunophilin expression changes were not corroborated by microarray analyses may be explained by the observation that quantitative information obtained with small cDNA inserts as well as clones arising from less abundant mRNA species could be unreliable due to their propensity to generate weak signals falling within the 'noise' of microarray hybridization signals (Yang *et al.* 1999). Here, the transcript for

immunophilin could well belong to the low-abundance class since it was represented by only one isolated clone (Table 3). Despite this limitation, our approach allowed us to isolate and screen a large number of transcripts, pointing out physiologic mechanisms apparently affected by estrogenic or antiestrogenic treatments in MCF7 breast cancer cells.

MCF7 cells develop into xenogafts in ovariectomized athymic mice only when supplemented with E_2 (Soule & McGrath 1980), indicating that E_2 promotes primary tumor growth. Our results show (Table 2) that, in MCF7 cells, E_2 regulates the expression of seven genes involved in the transition from the G_1 to the S phase of the cell cycle, such as c-myc (Safe 2001, Doisneau-Sixou *et al.* 2003), CDK2 (Safe 2001, Doisneau-Sixou *et al.* 2003), b-myb (Saville & Watson 1998), ASK (Masai *et al.* 2000), CKS1 (Spruck *et al.* 2001), TGFβ (Polyak *et al.* 1994) and nm23-H2 (Ji *et al.* 1995). Five of these, CDK2, nm23-H2, CKS1, b-myb and ASK, are reported here, for the first time, to our knowledge, as being regulated by E_2 at mRNA level. CDK2 forms complexes with cyclins E and A, indispensable for progression through mid- and late G_1 phase and entry into S phase (Safe 2001) (Fig. 3). Thus, E_2-induced increase in CDK2 expression should enhance the level of CDK2 protein available to form these complexes. The cyclin A/CDK2 complex was reported to activate b-myb involved in S-phase progression (Saville & Watson 1998). Overexpression of b-myb, observed here, could reinforce cell-cycle progression. Interestingly, we also observed the overexpression of ASK, another gene implicated in S-phase progression (Masai *et al.* 2000). Here, CKS1 induces the degradation of p27, a CDK2 inhibitor (Spruck *et al.* 2001), thus increasing CDK2 activity. The overexpression of CKS1 can also result from TGFβ underexpression, since TGFβ inhibits CKS1 transcription (Simon *et al.* 1995). Finally, decrease in TGFβ level induces additional inhibition of p27 expression (Polyak *et al.* 1994, Reynisdottir *et al.* 1995). Altogether, these different effects converge to enhance CDK2 activity in addition to the increases in CDK2 mRNA levels described above. In this study, we also observed the E_2-induced overexpression of nm23-H2, a nucleotide diphosphate kinase activating c-myc expression (Ji *et al.* 1995). This suggests that E_2 directly and indirectly stimulates c-myc expression.

One of the key features of an aggressive phenotype is the ability of the cancer cells to escape from the tumor stroma, thus promoting metastasis. The invasive process through the basement membrane requires changes in intercellular adhesion, cell motility and remodeling of the ECM (DePasquale *et al.* 1994, Lochter & Bissell 1995). In epithelium, desmosomes form strong adhesive junctions between cells by linking intermediate keratin filament networks to sites of intercellular adhesion. An essential structural component of desmosome is junction plakoglobin, the underexpression of which is associated with anchorage loss in cancers (Tada *et al.* 2000). Our results show that E_2 inhibits the expression of junction plakoglobin as well as that of keratins 8 and 18, which are interconnected to desmosomes via binding to desmoplakin, arguing for E_2-induced desmosome disorganization. In agreement with this, DePasquale *et al.* (1994) have observed a rearranging of vesicular plakoglobin staining after E_2 treatment of MCF7 cells, which could be a consequence of plakoglobin underexpression. We also observed in E_2-treated MCF7 cells, for the first time, to our knowledge, the underexpression of subunit 1B of the Arp2/3 complex, which is thought to promote actin filament assembly at the barbed end (Cooper & Schafer 2000), suggesting E_2-induced retardation of actin microfilament polymerization. Furthermore, the overexpression of Hsp27 (Porter *et al.* 2001) (Table 2), which inhibits the formation of basolateral microfilaments, suggests that E_2 treatment could alter microfilament homeostasis and microfilament dynamics (Piotrowicz & Levin 1997). Catenins form an important complex with cadherins, linking them to the actin filament network. Underexpression of β catenin (Gunin *et al.* 2003) in E_2-treated MCF7 cells could modify this process. Indeed, decrease in cadherin-associated β catenin has been associated with invasiveness in breast cancer (Gonzalez *et al.* 1999). Importantly, E_2 decreased γ-actin expression (Table 2), a phenomenon known to be associated with the progression from a nonmetastatic to a metastatic phenotype of cells in salivary gland adenocarcinoma (Suzuki *et al.* 1998). Finally, we observed, for the first time, to our knowledge, E_2-regulated expression of claudin 4, involved in tight junction constitution, the underexpression of which is associated with increased invasive potential (Michl *et al.* 2003). Altogether, our results and those of other workers indicate that, in MCF7 cells, E_2 treatment modulates the expression of major genes promoting the remodeling of cell adhesion structures, and leads, in consequence, to decreases in cell adhesiveness and increases in cell mobility, thereby setting off two major causal mechanisms associated with tumor progression and the acquisition of an invasive potential (Suzuki *et al.* 1998) through the loss of stable cell–cell adherent junctions (DePasquale *et al.* 1994).

In this study, we identified E_2-induced overexpression of cathepsin D and pS2, and underexpression of two new genes, USF2 and HGFA I2. Known to be associated with invasive human breast cancer (Elliott *et al.* 2002) and metastasis (Birchmeier *et al.* 1997), HGF can activate the transcription factor E1AF (Hanzawa *et al.* 2000), which, in turn, stimulates the production of various MMPs (matrix metalloproteinases) (Hanzawa *et al.* 2000), thereby promoting matrix degradation. HGF is synthesized as pro-HGF/SF, and it requires activation by factors such as HGF activator (HGFA).

Thus, through a decrease in HGFA I2 (inhibitor of HGFA) expression, E_2 appears to promote HGFA-mediated specific activation of HGF, resulting in ECM degradation. This process could be reinforced by the underexpression of USF2, observed here for the first time, which interacts, like other USF transcription factors, with the E1 box binding site on the HGF gene multiconsensus region. Jiang *et al.* (2000) have shown that only USF-1, and not USF-2, can stimulate HGF transcription. We can speculate that underexpression of USF-2 could diminish USF-2 competition for E1 box, facilitate USF-1-dependent HGF gene transcription and promote E1AF-dependent MMPs gene transcription, leading to increased matrix degradation potential. Furthermore, cathepsin D secretion by breast cancer cells has been shown to lead to matrix degradation via the formation of large acidic vesicles, which may facilitate matrix component digestion (Montcourrier *et al.* 1990). pS2 overexpression could be associated with an alteration in ECM deposition and increased cell mobility (Williams *et al.* 1996).

Genotoxic metabolites, such as catechol estrogens and quinone derivatives, arising from E_2 catabolism are suspected to promote carcinogenesis via production of genotoxic metabolites. Indeed, CYP1B1 (cytochrome P450) hydroxylates E_2 to the promutagenic metabolite 4-OH catechol estrogen (4-OH-E_2) (Hayes *et al.* 1996). E_2-induced underexpression of TGFβ, described here, could increase CYP1B1 expression (Dohr *et al.* 1997) and thus enhance E_2 catabolism. On the other hand, our data show that E_2 induces underexpression of UDP glucuronosyltransferase, which is involved in the catechol-quinone detoxification system (Raftogianis *et al.* 2000), thus contributing to increased production of semiquinone (E_2-3,4-semiquinone) and quinone (E_2-3,4-quinone) from 4-OH-E_2 (Fig. 4). However, quinones can be recycled to catechols in the presence of NAD(P)H menadione oxidoreductase (Long & Jaiswal 2000). Consequently, underexpression of this gene (Table 2) and decrease of its activity lead to increased levels of toxic E_2-3,4-derived quinones. Quinones can also be synthesized from ubiquinol conjugates in the presence of NAD(P)H dehydrogenase ubiquinone (Hatefi *et al.* 1985). Overexpression of this gene, in E_2-treated MCF7, reinforces the accumulation of quinone metabolites promoting tumor progression (Yager & Liehr 1996). Moreover, E_2 decreases both NAD(P)H:quinone oxidoreductase activity and thiol-specific antioxidant mRNA level (Table 3), an effect which could result in free radical damage (Yim *et al.* 1994, Ernster *et al.* 1995, Montano & Katzenellenbogen 1997). Thus, E_2 appears to play a further role in breast cancer development by modulating the expression of five enzymes involved in the control of E_2-dependent genotoxic metabolites and free radicals production, thereby increasing the probability that mutations occur during DNA synthesis. In conclusion, E_2 appears to promote breast cancer progression by stimulating cell growth, anchorage junctions remodeling and ECM disassembly, together with genotoxic metabolite accumulation. The slight changes in gene expression observed by microarray analyses (Tables 2 and 3) appear sufficient to promote tumor progression and invasiveness.

TAM, an E_2 antagonist, reverses the effects of E_2 upon expression of genes involved in genotoxic metabolite accumulation and cytoskeletal ECM remodeling in addition to the overexpression of profilin, which is known to promote actin filament assembly at the barbed end (Kang *et al.* 1999) and of HSD17 β4, which is known to inactivate E_2 by conversion to estrone (Luu-The 2001) (Tables 2 and 3). This could explain, at least in part, how TAM can abrogate the E_2-induced phenotype. The fact that cell-cycle progression was not inhibited can be explained by recent findings that TAM can function as a molecular agonist inducing cell-cycle-associated gene in breast cancer (Hodges *et al.* 2003). Interestingly, TAM treatment also affects the expression of several genes not regulated by E_2, such as profilin, DDB-2, XR-CC1, Hsp90, REA, immunophilin, ERK1, HSD17 β4 and TGF-α. Some of these genes could be involved in TAM resistance. Overexpression of REA (Table 3), which competes with the coactivator SRC-1 in binding to ER and inhibits ER transcriptional activity (Delage-Mourroux *et al.* 2000), could explain the reversal of E_2-dependent transcriptional activities observed with TAM treatment. Although antiestrogenic treatment increases the life expectancy of breast cancer patients, resistance and relapse phenomena can frequently occur (Johnston 1997). Hypotheses of TAM resistance include the following: 1. ER structural and functional modifications (Chander *et al.* 1993); 2. post-receptor interaction modifications; 3. paracrine secretion modifications (Katzenellenbogen 1991); 4. pharmacologic changes (Osborne *et al.* 1992). Our results support the first hypothesis, since we observed, in MCF7 cells treated with TAM, decreases in expression of Hsp90 and immunophilin, both of which are required for the formation of functional ER heterocomplexes capable of interacting with hormones (Schiene-Fischer & Yu 2001). We showed here that TAM decreases the expression of TGFα (Table 3). Throughout the duration of TAM treatment, the inhibition of growth factor expression would promote selection of cells endowed with constitutively high expression of one or more growth factors or growth factor receptor (Wakeling 1990). Indeed, TGFα is constitutively expressed in many estrogen-independent cells (Bates *et al.* 1988). Evidence now shows that constitutive activity of growth factor can bypass the cell's dependence on estrogen and provide a mechanism for hormone-independent growth (Johnston *et al.* 1992), thus making TAM treatment inefficient. This could explain, at least in part, the arising of TAM

resistance. Moreover, TAM increases ERK1 gene expression (Table 3), which could in turn enhance cell proliferation and inhibit TGFβ-induced antiproliferative responses via SMAD 2/3 phosphorylation (Kretzschmar 2000). This may indicate a paradoxical effect of TAM, explaining at least in part, the phenomenon of TAM resistance (Rabenoelina *et al.* 2002). Finally, the downregulating effects of TAM treatment upon two genes involved in the base excision DNA repair system, DDB-2 (Itoh *et al.* 1999, Nichols *et al.* 2000) and XRCC1 (Miller *et al.* 2001), could lead to increased mutation rate in tumor cells.

In conclusion, our experimental approach allowed us to identify 32 early E_2-induced genes, including 19 new genes. It also allowed us to detect the TAM-induced expression reversal of 11 E_2-induced genes. In addition, we were able to report TAM-specific regulation of nine genes, including eight new ones. Integrative data analyses provided a global understanding of the cellular mechanisms alterations associated with E_2 and TAM treatments. This knowledge could be important to predict cell behavior associated with antiestrogenic treatment and subsequently determine new therapeutic targets. In particular, an understanding of mechanisms leading to TAM resistance/relapse may facilitate the development of novel therapeutic strategies addressed to estrogen-regulated tumorigenesis.

References

Bates SE, Davidson NE, Valverius EM, Freter CE, Dickson RB, Tam JP, Kudlow JE, Lippman ME & Salomon DS 1988 Expression of transforming growth factor alpha and its messenger ribonucleic acid in human breast cancer: its regulation by estrogen and its possible functional significance. *Molecular Endocrinology* **2** 543–555.

Birchmeier W, Brinkmann V, Niemann C, Meiners S, DiCesare S, Naundorf H & Sachs M 1997 Role of HGF/SF and c-Met in morphogenesis and metastasis of epithelial cells. *Ciba Foundation Symposium* **212** 230–246.

Chander SK, Sahota SS, Evans TR & Luqmani YA 1993 The biological evaluation of novel antioestrogens for the treatment of breast cancer. *Critical Reviews in Oncology/Hematology* **15** 243–269.

Colditz GA 1998 Relationship between estrogen levels, use of hormone replacement therapy, and breast cancer. *Journal of the National Cancer Institute* **90** 814–823.

Cooper JA & Schafer DA 2000 Control of actin assembly and disassembly at filament ends. *Current Opinion in Cell Biology* **12** 97–103.

Couse JF & Korach KS 1999 Estrogen receptor null mice: what have we learned and where will they lead us? *Endocrine Reviews* **20** 358–417.

Delage-Mourroux R, Martini PG, Choi I, Kraichely DM, Hoeksema J & Katzenellenbogen BS 2000 Analysis of estrogen receptor interaction with a repressor of estrogen receptor activity (REA) and the regulation of estrogen receptor transcriptional activity by REA. *Journal of Biological Chemistry* **275** 35848–35856.

DePasquale JA, Samsonoff WA & Gierthy JF 1994 17-beta-Estradiol induced alterations of cell–matrix and intercellular adhesions in a human mammary carcinoma cell line. *Journal of Cell Science* **107** 1241–1254.

Dohr O, Sinning R, Vogel C, Munzel P & Abel J 1997 Effect of transforming growth factor-beta1 on expression of aryl hydrocarbon receptor and genes of Ah gene battery: clues for independent down-regulation in A549 cells. *Molecular Pharmacology* **51** 703–710.

Doisneau-Sixou SF, Sergio CM, Caroll JS, Hui R, Musgrove EA & Sutherland RL 2003 Estrogen and antiestrogen regulation of cell cycle progression in breast cancer cells. *Endocrine-Related Cancer* **10** 179–186.

El Etreby MF & Liang Y 1998 Effect of antiprogestins and tamoxifen on growth inhibition of MCF-7 human breast cancer cells in nude mice. *Breast Cancer Research and Treatment* **49** 109–117.

Elliott BE, Hung WL, Boag AH & Tuck AB 2002 The role of hepatocyte growth factor (scatter factor) in epithelial-mesenchymal transition and breast cancer. *Canadian Journal of Physiology and Pharmacology* **80** 91–102.

Ernster L & Dallner G 1995 Biochemical, physiological and medical aspects of ubiquinone function. *Biochimica et Biophysica Acta* **1271** 195–204.

Gadal F, Bozic C, Pillot-Brochet C, Malinge S, Wagner S, Le Cam A, Buffat L, Crépin M & Iris F 2003 Integrated transcriptome analysis of the cellular mechanisms associated with Ha-ras-dependent malignant transformation of the human breast epithelial MCF7 cell line. *Nucleic Acids Research* **31** 5789–5804.

Gibson DF, Gottardis MM & Jordan VC 1990 Sensitivity and insensitivity of breast cancer to tamoxifen. *Journal of Steroid Biochemistry and Molecular Biology* **37** 765–770.

Gonzalez MA, Pinder SE, Wencyk PM, Bell JA, Elston CW, Nicholson RI, Robertson JF, Blamey RW & Ellis IO 1999 An immunohistochemical examination of the expression of E-cadherin, alpha- and beta/gamma-catenins, and alpha2- and beta1-integrins in invasive breast cancer. *Journal of Pathology* **187** 523–529.

Gunin AG, Emelianov VU & Tolmachev AS 2003 Expression of estrogen receptor-alpha, glucocorticoid receptor, beta-catenin and glycogen synthase kinase-3 beta in the uterus of mice following long-term treatment with estrogen and glucocorticoid hormones. *European Journal of Obstetrics, Gynecology, and Reproductive Biology* **107** 62–67.

Hanzawa M, Shindoh M, Higashino F, Yasuda M, Inoue N, Hida K, Ono M, Kohgo T, Nakamura M, Notani K, Fukuda H, Totsuka Y, Yoshida K & Fujinaga K 2000 Hepatocyte growth factor upregulates E1AF that induces oral squamous cell carcinoma cell invasion by activating matrix metalloproteinase genes. *Carcinogenesis* **21** 1079–1085.

Hatefi Y, Ragan CI & Galante YM 1985 The enzymes and the enzyme complexes of the mitochondrial oxidative phosphorylation system. In *The Enzymes of Biological Membranes*, 5th edn, pp 1–70. Ed AN Martonosi. New York: Wiley.

Hayes CL, Spink DC, Spink BC, Cao JQ, Walker NJ & Sutter TR 1996 17β-Estradiol hydroxylation catalyzed by human cytochrome P450 1B1. *PNAS* **93** 9776–9781.

Hodges LC, Cook JD, Lobenhofer EK, Li L, Bennett L, Bushel PR, Aldaz CM, Afshari CA & Walker CL 2003 Tamoxifen functions as a molecular agonist inducing cell cycle-associated genes in breast cancer cells. *Molecular Cancer Research* **1** 300–311.

Inadera H, Hashimoto S, Dong HY, Suzuki T, Nagai S, Yamashita T, Toyoda N & Matsushima K 2000 WISP-2 as a novel estrogen-responsive gene in human breast cancer cells. *Biochemical and Biophysical Research Communications* **275** 108–114.

Itoh T, Mori T, Ohkubo H & Yamaizumi M 1999 A newly identified patient with clinical xeroderma pigmentosum phenotype has a non-sense mutation in the DDB2 gene and incomplete repair in (6–4) photoproducts. *Journal of Investigative Dermatology* **113** 251–257.

Ji L, Arcinas M & Boxer L 1995 The transcription factor, Nm23H2, binds to and activates the translocated c-myc allele in Burkitt's lymphoma. *Journal of Biological Chemistry* **270** 13392–13398.

Jiang JG, Gao B & Zarnegar R 2000 The concerted regulatory functions of the transcription factors nuclear factor-1 and upstream stimulatory factor on a composite element in the promoter of the hepatocyte growth factor gene. *Oncogene* **19** 2786–2790.

Johnston SR 1997 Acquired tamoxifen resistance in human breast cancer – potential mechanisms and clinical implications. *Anti-Cancer Drugs* **8** 911–930.

Johnston SR, Dowsett M & Smith IE 1992 Towards a molecular basis for tamoxifen resistance in breast cancer. *Annals of Oncology* **3** 503–511.

Kang F, Purich DL & Southwick FS 1999 Profilin promotes barbed-end actin filament assembly without lowering the critical concentration. *Journal of Biological Chemistry* **274** 36963–36972.

Katzenellenbogen BS 1991 Antiestrogen resistance mechanisms by which breast cancer cells undermine the effectiveness of endocrine therapy. *Journal of the National Cancer Institute* **83** 1434–1435.

Key TJ & Pike MC 1988 The role of oestrogens and progestagens in the epidemiology and prevention of breast cancer. *European Journal of Cancer and Clinical Oncology* **24** 29–43.

Kretzschmar M 2000 Transforming growth factor-β and breast cancer. Transforming growth factor-β/SMAD signaling defects and cancer. *Breast Cancer Research* **2** 107–115.

Kuang WW, Thompson DA, Hoch RV & Weigel RJ 1998 Differential screening and suppression subtractive hybridization identified genes differentially expressed in an estrogen receptor-positive breast carcinoma cell line. *Nucleic Acids Research* **26** 1116–1123.

Leung BS & Potter AH 1987 Mode of estrogen action on cell proliferation in CAMA-1 cells. II. Sensitivity of G1 phase population. *Journal of Cellular Biochemistry* **34** 213–225.

Lippman ME, Bolan G & Huff K 1975 Human breast cancer responsive to androgen in long term tissue culture. *Nature* **258** 339–341.

Lochter A & Bissell MJ 1995 Involvement of extracellular matrix constituents in breast cancer. *Seminars in Cancer Biology* **6** 165–173.

Long DJ 2nd & Jaiswal AK 2000 NRH quinone oxidoreductase2 (NQO2). *Chemico-Biological Interactions* **129** 99–112.

Luu-The V 2001 Analysis and characteristics of multiple types of human 17 beta-hydroxysteroid dehydrogenase. *Journal of Steroid Biochemistry and Molecular Biology* **76** 143–151.

Masai H & Arai K 2000 Regulation of DNA replication during the cell cycle roles of Cdc7 kinase and coupling of replication, recombination, and repair in response to replication fork arrest. *IUBMB Life* **49** 353–364.

Michl P, Barth C, Buchholz M, Lerch MM, Rolke M, Holzmann KH, Menke A, Fensterer H, Giehl K, Lohr, M, Leder G, Iwamura T, Adler G & Gress TM 2003 Claudin-4 expression decreases invasiveness and metastatic potential of pancreatic cancer. *Cancer Research* **63** 6265–6271.

Miller MC 3rd, Mohrenweiser HW & Bell DA 2001 Genetic variability in susceptibility and response to toxicants. *Toxicology Letters* **120** 269–280.

Montano MM & Katzenellenbogen BS 1997 The quinone reductase gene: a unique estrogen receptor-regulated gene that is activated by antiestrogens. *PNAS* **94** 2581–2586.

Montcourrier P, Mangeat PH, Salazar G, Morisset M, Sahuquet A & Rochefort H 1990 Cathepsin D in breast cancer cells can digest extracellular matrix in large acidic vesicles. *Cancer Research* **50** 6045–6054.

Nichols AF, Itoh T, Graham JA, Liu W, Yamaizumi M & Linn S 2000 Human damage-specific DNA-binding protein p48. Characterization of XPE mutations and regulation following UV irradiation. *Journal of Biological Chemistry* **275** 21422–21428.

Osborne CK, Coronado-Heinsohn EB, Hilsenbeck SG, McCue BL, Wakeling AE, McClelland RA, Manning DL & Nicholson RI 1995 Comparison of the effects of a pure steroidal antiestrogen with those of tamoxifen in a model of human breast cancer. *Journal of the National Cancer Institute* **87** 746–750.

Osborne CK, Wiebe VJ, McGuire WL, Ciocca DR & DeGregorio MW 1992 Tamoxifen and the isomers of 4-hydroxytamoxifen in tamoxifen-resistant tumors from breast cancer patients. *Journal of Clinical Oncology* **10** 304–310.

Piotrowicz RS & Levin EG 1997 Basolateral membrane-associated 27-kDa heat shock protein and microfilament polymerization. *Journal of Biological Chemistry* **272** 25920–25927.

Polyak K, Kato JY, Solomon MJ, Sherr CJ, Massague J, Roberts JM & Koff A 1994 p27 Kip1, a cyclin-Cdk inhibitor, links transforming growth factor-beta and contact inhibition to cell cycle arrest. *Genes and Development* **8** 9–22.

Porter W, Wang F, Duan R, Qin C, Castro-Rivera E, Kim K & Safe S 2001 Transcriptional activation of heat shock protein 27 gene expression by 17 beta-estradiol and modulation by antiestrogens and aryl hydrocarbon receptor agonists. *Journal of Molecular Endocrinology* **26** 31–42.

Rabenoelina F, Semlali A, Duchesne MJ, Freiss G, Pons M & Badia E 2002 Effect of prolonged hydroxytamxifen treatment of MCF-7 cells on mitogen activated kinase cascade. *International Journal of Cancer* **98** 698–706.

Raftogianis R, Creveling C, Weinshilboum R & Weisz J 2000 Estrogen metabolism by conjugation. *Journal of the National Cancer Institute. Monographs* **27** 113–124.

Reynisdottir I, Polyak K, Iavarone A & Massague J 1995 Kip/Cip and Ink4 Cdk inhibitors cooperate to induce cell cycle arrest in response to TGF-beta. *Genes and Development* **9** 1831–1845.

Safe S 2001 Transcriptional activation of genes by 17 beta-estradiol through estrogen receptor-Sp1 interactions. *Vitamins and Hormones* **62** 231–252.

Saville MK & Watson RJ 1998 The cell-cycle regulated transcription factor B-Myb is phosphorylated by cyclin A/Cdk2 at sites that enhance its transactivation properties. *Oncogene* **17** 2679–2689.

Schiene-Fischer C & Yu C 2001 Receptor accessory folding helper enzymes: the functional role of peptidyl prolyl *cis/trans*-isomerases. *FEBS Letters* **495** 1–6.

Simon KE, Cha HH & Firestone GL 1995 Transforming growth factor beta down-regulation of CKShs1 transcripts in growth-inhibited epithelial cells. *Cell Growth and Differentiation* **6** 1261–1269.

Sommers CL, Papageorge A, Wilding G & Gelmann EP 1990 Growth properties and tumorigenesis of MCF-7 cells transfected with isogenic mutants of rasH. *Cancer Research* **50** 67–71.

Soule HD & McGrath CM 1980 Estrogen responsive proliferation of clonal human breast carcinoma cells in athymic mice. *Cancer Letters* **10** 177–189.

Soule HD, Vazguez J, Long A, Albert S & Brennan M 1973 A human cell line from a pleural effusion derived from a breast carcinoma. *Journal of the National Cancer Institute* **51** 1409–1416.

Spruck C, Strohmaier H, Watson M, Smith AP, Ryan A, Krek TW & Reed SI 2001 A CDK-independent function of mammalian Cks1 targeting of SCF(Skp2) to the CDK inhibitor p27 Kip1. *Molecular Cell* **7** 639–650.

Sutherland RL, Hall RE & Taylor IW 1983 Cell proliferation kinetics of MCF-7 human mammary carcinoma cells in culture and effects of tamoxifen on exponentially growing and plateau-phase cells. *Cancer Research* **43** 3998–4006.

Sutherland RL, Watts CKW & Clarke CL 1988 In *Hormones and Their Actions*, Part 1, pp 197–215. Eds HJ Van Der Molen, RJB King & BA Cooke. Amsterdam: Elsevier.

Suzuki H, Nagata H, Shimada Y & Konno A 1998 Decrease in gamma-actin expression, disruption of actin microfilaments and alterations in cell adhesion systems associated with acquisition of metastatic capacity in human salivary gland adenocarcinoma cell clones. *International Journal of Oncology* **12** 1079–1084.

Tada H, Hatoko M, Tanaka A, Kuwahara M & Muramatsu T 2000 Expression of desmoglein I and plakoglobin in skin carcinomas. *Journal of Cutaneous Pathology* **27** 24–29.

Wakeling AE 1990 Mechanisms of growth regulation of human breast cancer. *Baillieres Clinical Endocrinology and Metabolism* **4** 51–66.

Williams R, Stamp GW, Gilbert C, Pignatelli M & Lalani EN 1996 pS2 transfection of murine adenocarcinoma cell line 410·4 enhances dispersed growth pattern in a 3-D collagen gel. *Journal of Cell Science* **109** 63–71.

Yager JD & Liehr JG 1996 Molecular mechanisms of estrogen carcinogenesis. *Annual Review of Pharmacology and Toxicology* **36** 203–232.

Yang GP, Ross DT, Kuang WW, Brown PO & Weigel RJ 1999 Combining SSH and cDNA microarrays for rapid identification of differentially expressed genes. *Nucleic Acids Research* **27** 1517–1523.

Yang YH, Dudoit S, Luu P & Speed T 2000 Normalization for cDNA microarray data. http://oz.Berkeley.EDU/users/terry/zarray/Html/normspie.html.

Yim MB, Chae HZ, Rhee SG, Chock PB & Stadtman ER 1994 On the protective mechanism of the thiol-specific antioxidant enzyme against the oxidative damage of biomacromolecules. *Journal of Biological Chemistry* **269** 1621–1626.

Received 16 September 2004
Accepted 29 September 2004
Made available online as an Accepted Preprint 18 October 2004