1116–1123 *Nucleic Acids Research, 1998, Vol. 26, No. 4*

Differential screening and suppression subtractive hybridization identified genes differentially expressed in an estrogen receptor-positive breast carcinoma cell line

Wayne W. Kuang, Devon A. Thompson, Renee V. Hoch and Ronald J. Weigel*

Department of Surgery, Stanford University, Stanford, CA 94305, USA

Received June 10, 1997; Revised and Accepted December 18, 1997

DDBJ/EMBL/GenBank accession no. AF007170

ABSTRACT

Differences in gene expression are likely to explain the phenotypic differences between hormone-responsive and hormone-unresponsive breast cancer. We have identified differentially expressed cDNAs in the estrogen receptor (ER)-positive MCF7 breast carcinoma cell line compared with the ER-negative MDA-MB-231 breast carcinoma cell line. Differential screening isolated four differentially expressed genes: cytokeratin 8, cytokeratin 18, *Hsp27* and *GPCR*-Br. To identify differentially expressed genes of lower abundance, suppression subtractive hybridization was utilized and 29 differentially expressed clones were isolated. Sequence analysis revealed that 11 clones were from previously described genes: *HEK8*, neuropeptide Y receptor Y1, *p21*$^{WAF-1}$, *p55*^{PIK}, cytokeratin 18 (cloned twice), fructose-1,6-biphosphatase, cytokeratin 8, TGFβ1 binding protein, elongation factor 1α2 and *pS2*. The remaining 18 clones did not match sequences in the GenBank/EMBL database, indicating that they may be novel genes. Expression of *pS2*, neuropeptide Y receptor Y1 and three novel clones was induced by estradiol, indicating estrogen-responsiveness. The expression pattern of one novel gene, *DEME*-6, correlated with expression of *ER* and *ERF*-1/*AP*-2γ in a panel of breast carcinoma cell lines. A 2.6 kb cDNA of *DEME*-6 was sequenced and contains an open reading frame of 574 amino acids that demonstrates 62.4% similarity with a gene from *Caenorhabditis elegans* chromosome III. Expression of *DEME*-6 was also detected in primary breast carcinomas but not in normal breast tissue, as determined by RT-PCR. These findings support the hypothesis that a set of genes coordinately regulated with *ER*, but not necessarily estradiol-responsive, are characteristic of the hormone-responsive breast cancer phenotype.

INTRODUCTION

The estrogen receptor (ER) is routinely used as a prognostic and predictive marker in the clinical management of breast cancer patients. Women with ER-positive breast cancer have a better prognosis (1,2) and generally have hormone-responsive tumors that are more likely to respond to endocrine therapy (3,4). In addition, ER-positive tumors are more highly differentiated and are found more frequently in post-menopausal women (5,6). The molecular basis for these observed differences between ER-positive and ER-negative tumors remains unclear and the role of ER in hormone-responsive breast cancer remains to be delineated.

These findings have led to the hypothesis that alterations in ER may account for biological differences between hormone-responsive and hormone-unresponsive tumors. ER is known to be an estrogen-induced transcription factor that transactivates expression of ER-regulated genes such as progesterone receptor (*PgR*), *pS2* and heat shock protein 27000 (*Hsp27*) that have a role in proliferation and differentiation (7–9). Studies of *ER* mRNA structure have identified alternately spliced 5′ exons (10–13) and have also characterized exon skipping variants that have been found not only in breast tumors, but also in normal breast tissue (14–17). The genomic structure of *ER* has also been analyzed; however, these studies have been unable to demonstrate a consistent correlation between amplifications, deletions, rearrangements or restriction fragment length polymorphisms of the *ER* gene and the inheritance of breast cancer, ER-negative status or overexpression of *ER* (18–20).

These observations have led us to propose that *ER* is only one of a set of expressed genes that are responsible for the phenotype of hormone-responsive breast cancer. In support of this hypothesis is the finding that ER-negative cells transfected with *ER* demonstrate paradoxical responses to estradiol (21). Further support for this hypothesis is that expression of *ER* in ER-positive breast carcinomas is controlled at the level of transcription (22,23) and transcriptional regulation of the *ER* gene may be controlled by the ERF-1/AP-2γ transcription factor, which binds the 5′-untranslated region of the *ER* gene (24–27). In addition, expression of *ERF*-1/*AP*-2γ correlates with *ER* expression in multiple breast and endometrial carcinoma cell lines (24). Consequently, the hypothesis has been extended to include the concept that factors such as ERF-1/AP-2γ may control transcription of a set of genes, in addition to *ER*, whose expression gives rise to the well-differentiated phenotype of hormone-responsive breast tumors.

In this present study we set out to compare the patterns of gene expression in ER-positive and ER-negative breast cancer cell lines by isolating differentially expressed genes in a human breast carcinoma cell line system. We used differential screening (DS) (28) and suppression subtractive hybridization (SSH) (29) to identify genes that are expressed in hormone-responsive ER-positive MCF7 cells that are absent or minimally expressed in hormone-unresponsive ER-negative MDA-MB-231 cells. Further analysis of these genes identified a subset that are estrogen-responsive as

*To whom correspondence should be addressed. Tel: +1 650 723 9799; Fax: +1 650 725 3918; Email: ron.w@forsythe.stanford.edu

well as a subset that are not estrogen-responsive but that are coordinately regulated with *ER*. These genes may encode proteins that play a critical role in establishing the clinical phenotype of hormone-responsive breast cancer.

MATERIALS AND METHODS

Cell lines

Cell lines MCF7, T-47D, BT-20, MDA-MB-231 and HBL-100 (American Type Culture Collection, Rockville, MD) were maintained in minimal essential medium (Gibco BRL, Gaithersburg, MD); ZR-75-1 was maintained in RPMI 1640 (Gibco BRL); MDA-MB-361 was maintained in Leibovitz's L-15 medium (Gibco BRL). Media were supplemented with 10% fetal calf serum (Hyclone, Logan, UT), 10 U/ml penicillin G (Gibco BRL), 10 μg/ml streptomycin (Gibco BRL) and 6 ng/ml bovine insulin (Sigma Chemical Co., St Louis, MO). All cells were incubated at 37°C in 5% CO_2 except MDA-MB-361, which were maintained in a CO_2-free environment. MCF7 cells that were induced with estradiol were grown under normal conditions as described above until ~25% confluent. The medium was subsequently changed to phenol red-free minimal essential medium supplemented with 10% charcoal-stripped fetal calf serum, 10 U/ml penicillin G, 10 μg/ml streptomycin, 6 ng/ml bovine insulin for 6 days. MCF7 cells were induced with 1×10^{-8} M water soluble β-estradiol (Sigma Chemical Co.) for 3 days.

mRNA isolation

Polyadenylated RNA was isolated using the Fast Track® Kit (Invitrogen® Corp., Carlsbad, CA) as per the recommendations of the manufacturer.

Differential screening

An amplified oligo(dT)-primed MCF7 cDNA λ library was generated using the ZAP Express™ [*Eco*RI/*Xho*I] cDNA Synthesis Kit (Stratagene, La Jolla, CA). Aliquots of 2 μg MCF7 and MDA-MB-231 mRNA were reverse transcribed with random primers (Perkin Elmer, Foster City, CA). One twentieth (~100 ng) of MCF7 and MDA-MB-231 cDNA was labeled by random priming (Boehringer Mannheim, Indianapolis, IN) with 50 μCi [α-^{32}P]dCTP (3000 Ci/mmol; Amersham Life Science Inc., Arlington Heights, IL). Approximately 50 000 plaques of the MCF7 cDNA library were screened by hybridizing duplicate Protran™ (Schleicher & Schuell, Keene, NH) plaque lifts (30) with either MCF7 or MDA-MB-231 cDNA probes. All pre-hybridizations and hybridizations were performed under the following conditions: 50% formamide, 5× Denhardt's, 5× SSPE, 0.1% SDS and 50 μg/ml denatured salmon sperm DNA at 42°C. Membranes were washed with 2× SSC and 0.1% SDS at 42°C for 15 min, followed by 0.1× SSC and 0.5% SDS at 65°C for 30 min. Autoradiographs were compared and plaques that had a signal in the MCF7 but not the MDA-MB-231 cell line were isolated and purified by a secondary round of screening. Isolated plaques were processed to excise the cDNA inserts from the λ vector into pBK-CMV phagemids. Subsequent to repeated isolations of cytokeratin 8 and cytokeratin 18 cDNAs the strategy of using a third plaque lift from plates during the secondary round of screening was adopted. This third lift was probed sequentially with 25 ng labeled cytokeratin 8 and cytokeratin 18 cDNA respectively, thus allowing identification of clones that corresponded to these genes.

Suppression subtractive hybridization

SSH was performed with the Clontech PCR-Select™ cDNA Subtraction Kit (Clontech Laboratories Inc., Palo Alto, CA) as described by the manufacturer but with the following modifications. Starting material consisted of 2 μg MCF7 mRNA as tester and 2 μg MDA-MB-231 mRNA as driver. Primary and secondary PCR conditions were altered to increase specificity of amplification according to either plan A or B. Both A and B reduced the extension time and the number of cycles of the primary PCR to 2 min and 26 cycles and the primary PCR products were diluted 1/50 prior to use in the secondary PCR. All other aspects of plan A were as per the instructions of the manufacturer. Plan B diverged from plan A in two ways. First, the initial cycle of primary PCR was performed using annealing and extension times that had been reduced to 15 s and 1.5 min respectively. Second, for subsequent cycles the denaturing time was increased to 10 s while the annealing and extension times were reduced to 15 s and 1.5 min respectively. All PCR products generated using both plans A and B were subcloned into the pCR™II vector using the Original TA Cloning® Kit (Invitrogen®). Clones 1–11 and 17–48 were generated using plan A, while clones 12–16 were generated using plan B.

Isolation of a 2.6 kb *DEME*-6 cDNA

Approximately 1×10^6 plaques from the MCF7 cDNA library were screened using oligonucleotide probes derived from the 111 bp SSH-generated *DEME*-6 clone. Fifteen picomoles each of oligonucleotides o6-1 (5′-TGGGCCTTTCTCCAGCATCTCT-TCAGCC-3′) and o6-2 (5′-GTGATAATCTCAAGTATCCCAT-CCG-3′) were end-labeled individually with 75 μCi [γ-^{32}P]ATP (6000 Ci/mmol; Amersham) using T4 polynucleotide kinase (New England Biolabs, Beverly, MA). Labeled o6-1 and o6-2 were then combined and hybridized to plaque lifts on Optitran™ (Schleicher and Schuell) under the following conditions: 20% formamide, 5× Denhardt's, 5× SSPE, 0.1% SDS and 100 μg/ml denatured salmon sperm DNA at 42°C. Plaque lifts were washed in 2× SSC and 0.1% SDS at 42°C for 20 min, followed by one wash with 2× SSC and 0.1% SDS at 45°C for 20 min. Plaque lifts were placed on film with an intensifying screen at –80°C. Plaques that had a signal were isolated and purified by a secondary round of screening. Isolated plaques were processed to excise the cDNA inserts from the λ vector into pBK-CMV phagemids.

Northern analysis

Aliquots of 1 μg mRNA were electrophoresed on a 1% agarose–formaldehyde denaturing gel in 1× MOPS and then transferred to a Nytran® membrane (Schleicher & Schuell). Samples of 25–50 ng of each clone in pCR™II or pBK-CMV were ^{32}P-labeled by random priming (Boehringer Mannheim). In addition, 80 ng of a 530 bp PCR-generated fragment of ERF-1 cDNA (spanning a region from bp 2134 to 2664 in the 3′-untranslated region) (26) was labeled similarly. Northern blots were pre-hybridized and hybridized in 50% formamide, 5× Denhardt's, 5× SSPE, 1% SDS and 100 μg/ml denatured salmon sperm DNA at 42°C. Blots were washed in 2× SSC and 0.1% SDS at 42°C for 20 min, followed by two washes in 0.2× SSC and

0.1% SDS at 65°C for 20 min each. Northern blots were placed on film with an intensifying screen at –80°C. A Computing Densitometer 300A (Molecular Dynamics, Sunnyvale, CA) was utilized to determine relative expression. Values were normalized against β-actin.

Sequencing analysis

All sequencing was performed on double-stranded templates using the dideoxynucleotide chain termination method (31) with [α-^{35}S]dATP (1000 Ci/mmol; Amersham). Sequencing reactions were carried out with the Sequenase™ v.2.0 DNA Sequencing Kit (US Biochemical, Cleveland, OH). Sequence for the 2.6 kb *DEME*-6 cDNA was determined on both strands using an automated ABI 373 DNA sequencing system and a standard dye terminator AmpliTaq Kit. T3 and T7 promoter primers and custom sequence-specific primers were used for clones in pBK-CMV, whereas SP6 and T7 promoter primers were used for clones in pCR™II. The nucleotide sequence of the 2.6 kb *DEME*-6 cDNA was compared against the GenBank/EMBL and Expressed Sequence Tag (EST) (32) databases. In addition, the amino acid sequence of DEME-6 was compared against the PROSITE protein motif database (33) and analyzed by SOSUI (Secondary Structure Prediction of Membrane Proteins, http://www.tuat.ac.jp/~mitaku/sosui/).

Isolation of genomic clones of *DEME*-6

Approximately 1×10^6 λ phage from a human genomic library (generated from human placental tissue) in EMBL3 SP6/T7 cloning vector (Clontech) were screened with the oligonucleotides and *DEME*-6 cDNA under the conditions described for screening of the cDNA library. Clones that hybridized to the probes were isolated and plaque purified. These genomic clones were then compared with the cDNA to identify intron–exon borders.

RT-PCR from primary tumors

Primary human breast tumor tissue was collected fresh from mastectomy and biopsy specimens and snap frozen in liquid nitrogen. Normal human breast tissue was obtained from mastectomy specimens in a region of normal breast. Approximately 0.5 mg tissue were homogenized and total RNA was isolated using TRIzol® reagent (Gibco BRL) as per the manufacturer's recommendations. An aliquot of 1 µg RNA from each sample was reverse transcribed using random hexamers with the Advantage™ RT-for-PCR Kit (Clontech) as per the manufacturer's recommendations. The 20 µl RT sample was diluted to 100 µl with water and then 2 µl each sample analyzed by PCR for *DEME*-6, *ER* and *GAPDH* messages using the Advantage™ cDNA PCR Kit (Clontech) with Advantage™ KlenTaq Polymerase mix and gene-specific primers designed across intron–exon junctions. Primers for *DEME*-6 were o6B-5 (5′-GCTTCACCTACAAGGGCCAGTGGAAGATGTCC-3′) and o6B-6 (5′-TTCACCAAGCACTCGTCATCCACTGAGTACTCG-3′), which generated a 439 bp DNA fragment. Primers for ER were oER-1A-long (5′-GTGCCCTACTACCTGGAGAACGAGCCCAGC-3′) and oER-1B-long (5′-AGCATAGTCATTGCACACTGCACAGTAGCG-3′), which generated a 195 bp DNA fragment. Primers for *GAPDH* were GAPDH-5′ (5′-TGAAGGTCGGAGTCAACGGATTTGGT-3′) and GAPDH-3′ (5′-CATGTGGGCCATGAGGTCCACCAC-3′), which generated a 983 bp DNA fragment.

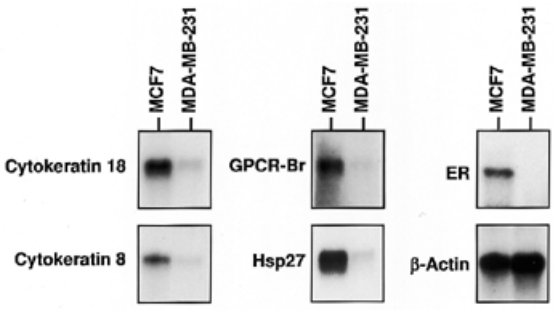

Figure 1. Northern blots of differentially expressed genes from differential screening. Northern blots demonstrate that clones isolated using differential screening detect mRNAs that are differentially expressed. The cDNAs were used as probes on Northern blots of mRNA isolated from ER-positive MCF7 and ER-negative MDA-MB-231 cells. The same blots were hybridized with β-actin cDNA to confirm similar loading and transfer of the mRNA, as well as with ER cDNA to serve as a positive control. Northern blots were placed on film at –80°C with a screen and the time needed to visualize hybridization signals ranged from 1 to 12 h.

Two-step PCR was utilized to amplify the genes and was performed on a Perkin Elmer 9600 DNA thermal cycler as follows: one cycle of 94°C for 1 min; 30 cycles of 94°C for 30 s, 68°C for 3 min; followed by 68°C for 3 min. PCR samples were then analyzed for *DEME*-6, *ER* and *GAPDH* by electrophoresis on 8% non-denaturing polyacrylamide gels.

RESULTS

Differential screening

The DS technique identified four genes that were over-expressed in MCF7 cells when compared with ER-negative MDA-MB-231 cells (Fig. 1). Sequence analysis identified these four genes as cytokeratin 8, cytokeratin 18, *Hsp27* and a novel member of the G protein-coupled receptor superfamily, *GPCR*-Br (34). Table 1 indicates that these genes each displayed 6-fold or greater expression in MCF7 than in MDA-MB-231 cells, as determined using Northern blots and densitometric analysis. Of the 157 clones isolated by DS, 123 (75%) were confirmed to be differentially expressed by Northern blot analysis. As shown in Table 1, the majority (110 of the 123) were found by sequence analysis to represent cytokeratin 18. In comparison, cytokeratin 8 was cloned 11 times, while *Hsp27* and *GPCR*-Br were each isolated once. Northern analysis hybridization signals for *GPCR*-Br were visualized in 12 h, while signals for cytokeratin 8, cytokeratin 18 and *Hsp27* were seen within 2 h. Isolation of cytokeratin 8, cytokeratin 18 and *Hsp27* was expected, since differential screening is known to isolate differentially expressed genes of high abundance (35) and these genes have been previously described in the literature as being among those genes that are differentially expressed at high levels (9,36). In an attempt to isolate genes such as *ER*, *PgR* and *pS2*, that are differentially expressed at lower levels, SSH was subsequently performed.

Suppression subtractive hybridization

The subtractive cloning strategy of SSH generated 332 clones with cDNA insert sizes ranging from 55 to ~1000 bp (Table 2). Forty eight of the 332 clones were further characterized. Six did not have cDNA inserts and 29 of the remaining 42 clones (69%) examined by Northern analysis of mRNA from MCF7 and

MDA-MB-231 cell lines were confirmed to be differentially expressed genes. These 29 clones have been designated *DEME* (differentially expressed in MCF7 with estradiol). Figure 2 shows that these genes were expressed in MCF7 cells while absent or minimally expressed in MDA-MB-231 cells. Among the 29 clones, differences in expression ranged from 6-fold to on/off. The time needed to visualize Northern analysis hybridization signals ranged from 1 h to 2 weeks.

Table 1. Summary of cDNA clones isolated from an MCF-7 cell line by differential screening

Gene[a]	mRNA (kb)[b]	No. of isolations[c]	Differential expression[d]
Cytokeratin 18	1.5	110	15-fold
Cytokeratin 8	1.8	11	19-fold
Hsp27	0.9	1	6-fold
GPCR-Br[e]	2.6	1	20-fold

[a]Sequence identity based on comparison with GenBank/EMBL database.
[b]Estimate of size (kb) of mRNA from Northern analysis.
[c]123 differentially expressed clones were isolated by differential screening of an MCF7 cDNA library with MCF7 and MDA-MB-231 cDNA.
[d]Relative expression is based on Northern analysis of MCF7 and MDA-MB-231 mRNA. Values attained by densitometry and normalized by comparison with β-actin (*x*-fold, degree of over-expression in MCF7 cells).
[e]*GPCR*-Br, G protein-coupled receptor–breast.

Sequence analysis revealed that 11 of the 29 clones matched previously described genes in the GenBank/EMBL database, as indicated in Table 2. Cytokeratin 8 and cytokeratin 18, that have been previously cloned using the alternative technique of DS, were isolated again using SSH. A known estrogen-responsive gene, *pS2* (*DEME*-40), which had previously been characterized as a differentially expressed gene (8), was also isolated. The sequences of the other 18 clones did not match any entries in the GenBank/EMBL database and are likely to be novel. Sixteen of the novel clones showed on/off differences in expression between the MCF7 and MDA-MB-231 cell lines.

Estrogen-responsiveness and pattern of expression in a panel of breast carcinoma cell lines

Each gene that had been identified as being differentially expressed in MCF7 cells as compared with MDA-MB-231 cells was used as a probe for Northern blots of mRNA from MCF7 cells that had been grown in the presence or absence of β-estradiol. Figure 3 shows that expression of five of these genes was augmented by β-estradiol treatment. The expression of *pS2* (*DEME*-40), neuropeptide Y receptor Y1 (NPY Rc Y1; *DEME*-12) and three novel clones (*DEME*-2, *DEME*-31 and *DEME*-47) was shown to be estrogen-responsive. The induction of *pS2* by estradiol has been previously described (8). While these five clones exhibited 4-fold or on/off induction, the majority of differentially expressed genes showed minimal (0.5- to 2.5-fold) or no response to estradiol treatment.

The differentially expressed genes identified were based only on a two cell line comparison, thus it was important to examine each on a panel of breast cancer cell lines. As anticipated, the estrogen-responsive genes demonstrated a correlation with *ER* expression in the panel of cell lines examined (data not shown). However, we were particularly interested in identifying genes that were not estrogen-responsive but exhibited a pattern of expression that correlated with the ER-positive phenotype. Genes coordinately regulated with *ER* are likely to play an important role in hormone-responsive tumors and may be transcriptionally regulated by mechanisms similar to those controlling *ER* expression. Of the subset of differentially expressed genes that are not estrogen-responsive only one gene, *DEME*-6, demonstrated a pattern of expression that correlated with *ER* expression in the panel of breast carcinoma cell lines examined.

Table 2. Summary of *DEME* cDNA clones generated from an MCF7 cell line by SSH

DEME[a]	Size (bp)[b]	Gene[c]	mRNA (kb)[d]	Differential expression[e]
2	297	–	1.7	±
6	111	–	2.7	±
7	297	*HEK8*	6.9/4.6	9-fold
8	358	–	3.5	±
9	341	–	1.4/1.0	±
10	206	–	1.4/1.0	±
12	125	NPY Rc Y1	8.7/2.7	±
13	(350)	–	9.0/4.4/3.3/1.8	±
15	148	$p21^{WAF-1}$	2.1	6-fold
16	(800)	$p55^{PIK}$	6.0	21-fold
19	396	Cytokeratin 18	1.5	8-fold
23	384	–	5.1/3.6	±
24	114	Fructose-1,6-bisphosphatase	1.4	±
25	274	Cytokeratin 8	1.8	19-fold
27	201	–	6.6/4.3/2.4	7-fold
29	111	–	1.2	±
30	(800)	–	8.5	±
31	63	–	1.8	±
33	162	TGFβ1 binding protein	6.8	8-fold
35	95	Cytokeratin 18	1.5	7-fold
36	62	Elongation factor 1α2	1.8	25-fold
37	148	–	7.1	7-fold
40	122	*pS2*	0.7	±
41	135	–	3.6	±
42	98	–	0.5	±
43	82	–	3.7	±
46	(750)	–	5.3	±
47	55	–	1.7	±
48	(1000)	–	8.3	±

[a]Of 48 clones six had no PCR-generated cDNA insert. The remaining 42 were analyzed by Northern blots. Differential expression was confirmed for the 29 *DEME* clones above.
[b]Size of PCR-generated insert, if known, or estimated size in parentheses.
[c]Sequence indentity based on comparison with GenBank/EMBL database. –, novel sequence; NPY Rc Y1, neuropeptide Y receptor type Y1; $p55^{PIK}$, P55 phosphotidylinositol kinase.
[d]Estimate of size (kb) of mRNA by Northern analysis.
[e]Relative expression is based on Northern analysis of MCF7 and MDA-MB-231 mRNA. Values attained by densitometry and normalized by comparison with β-actin. (±, expression in MCF7 and not in MDA-MB-231; *x*-fold, degree of overexpression in MCF7).

Isolation of *DEME*-6 cDNA

An MCF7 cDNA library was screened with an oligonucleotide probe prepared from the partial sequence of *DEME*-6. Fifteen

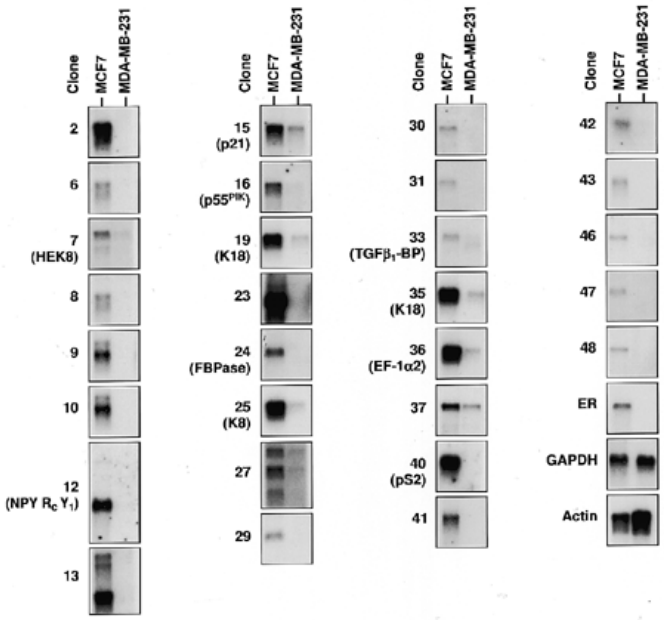

Figure 2. Northern blots of differentially expressed *DEME* genes from SSH. Northern blots demonstrate that the *DEME* clones generated from an MCF7 minus MDA-MB-231 subtraction using SSH detect mRNAs that are differentially expressed. Of the 48 SSH *DEME* clones examined 29 were used as cDNA probes for Northern blots of mRNA isolated from ER-positive MCF7 and ER-negative MDA-MB-231 cells. The other 19 clones either had no insert or were equally expressed in the two cell lines. The same blots were hybridized with β-actin and GAPDH cDNA to confirm similar loading and transfer of the mRNA as well as with ER cDNA to serve as a positive control. Northern blots were placed on film at –80°C with a screen and the time needed to visualize hybridization signals ranged from 1 h to 2 weeks. *DEME*-7, *HEK8*; *DEME*-12, neuropeptide Y receptor Y1; *DEME*-15, *p21*$^{WAF-1}$; *DEME*-16, *p55*$^{phosphotidylinositol\ kinase}$; *DEME*-19, cytokeratin 18; *DEME*-24, fructose-1,6-biphosphatase; *DEME*-25, cytokeratin 8; *DEME*-33, TGFβ1 binding protein; *DEME*-35, cytokeratin 18; *DEME*-36, elongation factor 1α2; *DEME*-40, *pS2*.

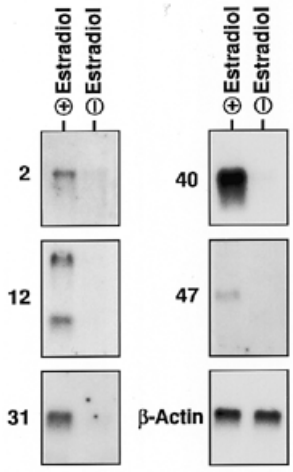

Figure 3. Northern blots demonstrating estrogen-responsiveness. The cDNAs of five *DEME* clones were used as probes for Northern blots of mRNA from the ER-positive MCF7 cells that had been grown in the presence (+) or absence (–) of β-estradiol. The remaining differentially expressed clones examined showed minimal (0.5- to 2.5-fold) or no response to β-estradiol treatment (data not shown). The same blots were hybridized with β-actin to confirm similar loading and transfer of the mRNA. *DEME*-40, *pS2*; *DEME*-12, neuropeptide Y receptor Y1; *DEME*-2, *DEME*-31 and *DEME*-47 are novel.

cDNAs were isolated, the largest of which was 2.6 kb. The sequence of the 2.6 kb *DEME*-6 cDNA is shown in Figure 4A. This cDNA contains an open reading frame of 574 amino acids, that may extend 5′ to the end of this cDNA clone, and a 3′-untranslated region of 917 bp. The sequence was compared with the GenBank/EMBL database and does not match any previously reported gene; however, the predicted amino acid sequence demonstrates an overall 62.4% similarity (36.6% identity and 25.8% conservative substitution) with the *Caenorhabditis elegans C32D5.6* gene (GenBank accession no. 746469) (Fig. 4B). The predicted amino acid sequence also shows homology to various sequences in the human EST database and further analysis has identified a putative RNA binding region RNP-1 signature and a highly hydrophobic region that may represent an anchoring transmembrane domain (Fig. 4A). The sequence for *DEME*-6 has been submitted to the GenBank/EMBL database and assigned accession no. AF007170.

The 2.6 kb cDNA for *DEME*-6 was hybridized to a panel of breast carcinoma cell lines. As shown in Figure 5, *DEME*-6 is expressed in the ER-positive cell lines MCF7, T47-D, MDA-MB-361 and ZR-75–1. *DEME*-6 is also expressed in the ER-negative BT-20 cell line, which expresses low levels of mutant *ER* mRNA (37). Expression of *DEME*-6 is absent or minimal in the ER-negative cell lines MDA-MB-231 and HBL-100. As can be seen in Figure 5, *DEME*-6 expression also parallels the expression pattern of the ERF-1/AP-2γ transcription factor.

RT-PCR from primary tumors

DEME-6 expression was examined in a panel of primary breast carcinomas. Mapping of the genomic clones indicated that an intron occurred in the region between nt 999 and 1313. The identification of this intron was an important feature used to design primers that amplified *DEME*-6 cDNA across this splice site. RT-PCR was used to examine *DEME*-6, *ER* and *GAPDH* expression in 12 primary breast tumors and two samples of normal breast tissue as shown in Figure 6. *DEME*-6 was expressed in half of the tumor samples (top panel) but was not detected in normal breast tissue. As reported previously (38), *ER* mRNA was detected in all tumor samples but was more abundant in tumors that were classified as ER-positive by immunoenzyme assay (middle panel). The cell lines MCF7 and MDA-MB-231 were included as positive and negative controls respectively. Although the primary tumors examined are comprised of a very heterogeneous population of cell types, these data demonstrate that *DEME*-6 was expressed in a significant fraction of the primary breast tumors but was not detected in normal mammary tissue.

DISCUSSION

The molecular basis for the hormone-responsive phenotype and improved prognosis associated with ER-positive status is poorly understood. To shed light on the underlying tumor biology we have compared the pattern of gene expression in two breast carcinoma cell lines. DS and SSH were used to isolate a panel of genes that are over-expressed in hormone-responsive ER-positive MCF7 cells as compared with hormone-unresponsive ER-negative MDA-MB-231 cells. Since SSH can isolate cDNAs from different regions of the same gene, we cannot be certain that each SSH clone represents a different gene. However, based on the transcript size, differential expression and estrogen-responsiveness we estimate that this panel

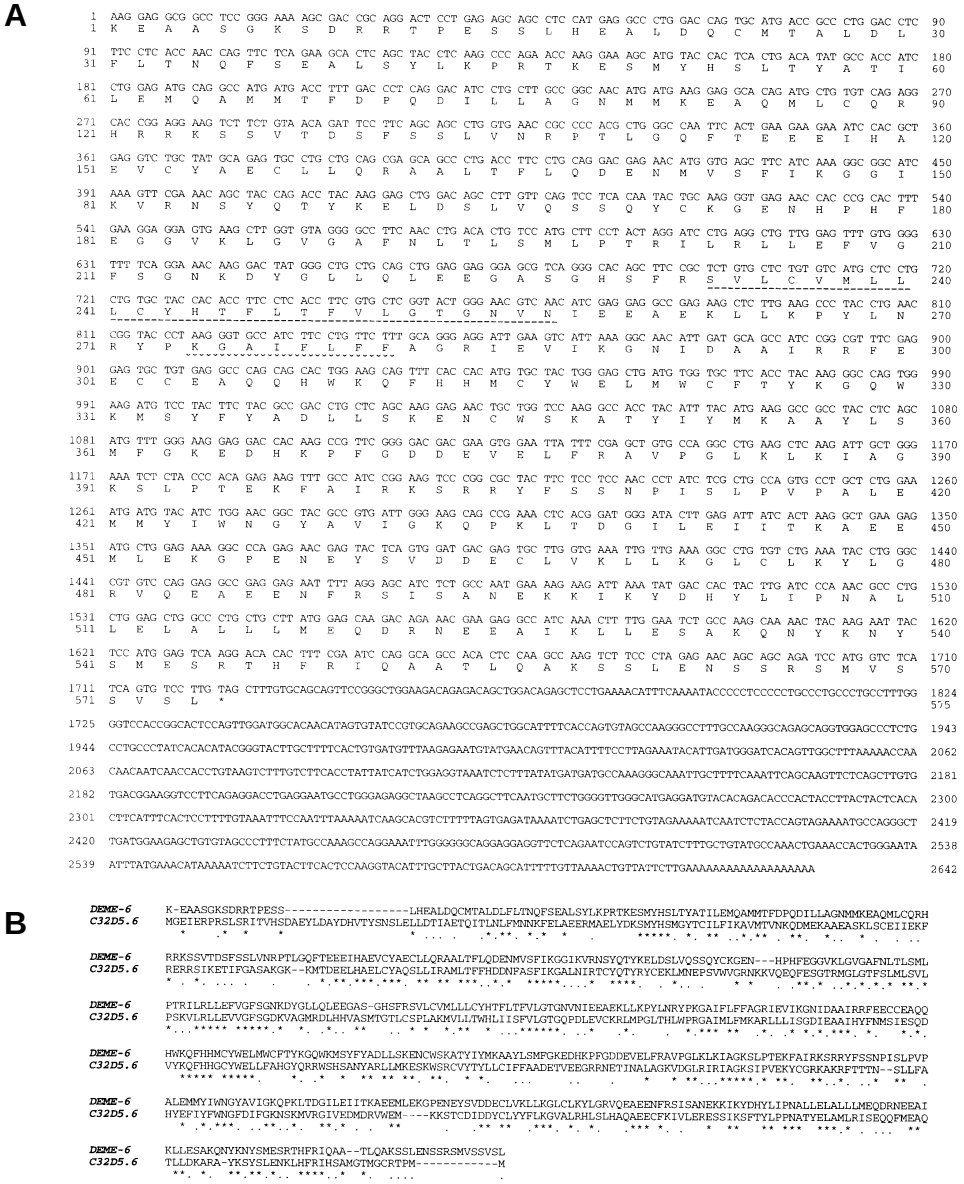

Figure 4. Nucleotide and deduced amino acid sequence of the human *DEME*-6 gene. (**A**) The 2642 bp nucleotide sequence of *DEME*-6 contains a long open reading frame which encodes a protein of 574 amino acids as indicated. Sequence analyses utilized the computer program DNA Strider 1.2 (47). Further examination of the DEME-6 protein using PROSITE (33) and SOSUI (http://www.tuat.ac.jp/~mitaku/sosui/) revealed a putative eukaryotic RNP-1 motif that is indicated with the demarcation (v v v v) under the amino acids and a candidate anchoring transmembrane domain that is indicated with the demarcation (- - - -) under the amino acids. (**B**) Using the computer program Clustal V (48) the longest open reading frame (574 amino acids) of *DEME*-6 was aligned with the 578 amino acid protein product from *C32D5.6* of *C.elegans* (GenBank accession no. 746469), demonstrating an overall similarity of 62.4%. Identical amino acids are indicated with an asterisk (*) and conserved amino acids are indicated with a period (.).

of genes includes at least 28 separate genes. Isolation of cytokeratin 8 and cytokeratin 18 by both DS and SSH attests to the validity of these two techniques, since these two genes have been previously described as being differentially expressed (36).

Comparison of DS and SSH

The simplicity of the DS technique is advantageous, however, DS is limited by its ability to isolate only those genes that are abundantly expressed, such as cytokeratin 8, cytokeratin 18 and *Hsp27* (9,35,36). An additional disadvantage is the inability to equalize the differing levels of mRNAs, which contributes to repeated isolation of abundant genes. As a result, cytokeratin 18 was isolated 110 times (Table 1). Furthermore, even though 75% of the genes isolated using differential screening were differentially expressed, none of them exhibited an on/off pattern of expression and these 123 differentially expressed genes represented only three known genes and one novel gene. This novel gene has been characterized as a member of the G protein-coupled receptor superfamily and is associated with *ER* expression in breast cancer cell lines and primary tumors (34). While differential screening is an effective strategy, known genes differentially expressed at low levels in an on/off fashion, such as *ER*, *PgR* and *pS2*, were not isolated. This prompted the use of the recently developed

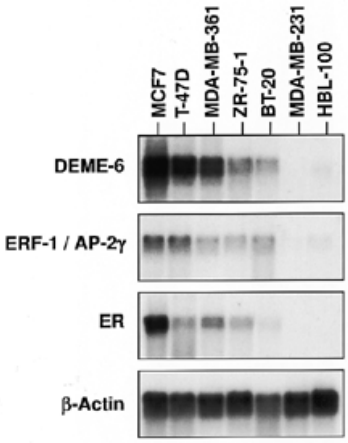

Figure 5. Northern blots demonstrating a correlation with estrogen receptor status. cDNA from the 2.6 kb clone of *DEME*-6 and *ERF*-1/*AP*-2γ were used to probe Northern blots containing mRNA from several ER-positive (MCF7, T-47D, MDA-MB-361 and ZR-75-1) and ER-negative (BT-20, MDA-MB-231 and HBL-100) breast carcinoma cell lines. Blots were re-hybridized with β-actin to confirm approximately equal loading and transfer as well as with *ER* cDNA to serve as a positive control.

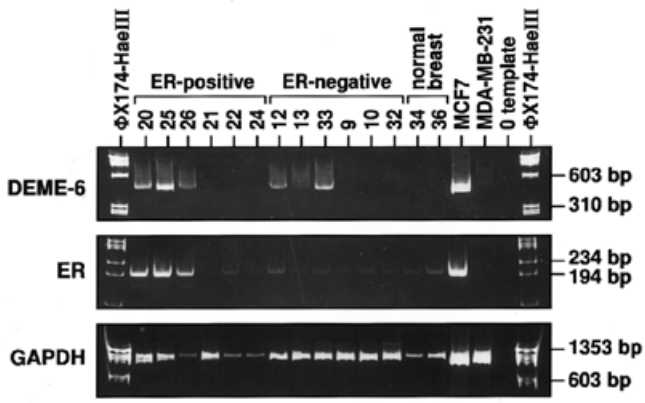

Figure 6. Examination of *DEME*-6 expression in primary breast tumors by RT-PCR. Electrophoresis of RT-PCR products for *DEME*-6, *ER* and *GAPDH* for six ER-positive (20, 25, 26, 21, 22 and 24) and six ER-negative (12, 13, 33, 9, 10 and 32) tumors determined by immunohistochemistry, two normal breast samples (34 and 35) and two breast carcinoma cell lines, MCF7 (ER-positive) and MDA-MB-231 (ER-negative). *GAPDH* was identified in all samples.

technique of SSH, which has been refined to successfully identify differentially expressed genes.

SSH combines subtractive hybridization with PCR to generate a population of PCR fragments enriched for sequences from genes that are differentially expressed. While the underlying concept has been used before in techniques such as representational difference analysis (39), the strength and novelty of SSH stems from a process called normalization. This process equalizes the wide differences in abundance of different mRNA species (29). Consequently, differentially expressed genes of low abundance that cannot be detected by DS can be cloned. In addition, differentially expressed genes of high abundance are not isolated in excess. For example, only two of the 29 differentially expressed SSH clones were cytokeratin 18. Compared with other PCR-based cloning strategies, such as differential display (40), the problem of false positives in SSH seemed to be less of an issue; ~70% of cloned inserts represented differentially expressed genes. Our results indicate that SSH is an effective technique of high sensitivity that identifies differentially expressed genes of high and low abundance.

Estrogen-responsive genes

Using cells that are ER-positive we expected that a subset of these genes would be estrogen-responsive and perhaps regulated by ER. The estradiol-induced *pS2* gene was first described as an ER-regulated gene in MCF7 cells (41), therefore, isolation of this gene (*DEME*-40) validates our ability to identify estrogen-responsive genes. Expression of NPY Rc Y1 has not previously been reported as being induced by estradiol treatment, however, the secretion of gonadotropin releasing hormone associated with neuropeptide Y and *PgR* expression has been linked to estradiol treatment (42,43). Three novel genes have also been characterized as estrogen-responsive, indicating that they may be regulated by ER. Given the central role of *ER* expression in breast cancer, the identification of these novel genes that are candidate ER targets is an important contribution of this study. We are currently characterizing the structure of these genes in an attempt to understand their functional significance.

Expression of *DEME*-6

An important goal of this project was to identify a set of genes whose expression may be characteristic of the hormone-responsive breast cancer phenotype. ER and ER-regulated genes have long been associated with the hormone-responsive phenotype. Thus a particular interest was in identifying genes that are coordinately expressed with *ER* but not necessarily estradiol-inducible. One such gene identified is *DEME*-6. The expression pattern of *DEME*-6 parallels the expression of *ER* in cell lines, however, *DEME*-6 is not an estrogen-responsive gene. This suggests the possibility that *DEME*-6 and *ER* share common transcriptional mechanisms. A candidate transcription factor that may coordinately regulate these two genes is ERF-1/AP-2γ. It has been shown that ERF-1/AP-2γ is involved in regulating *ER* expression (24) and the expression pattern of *DEME*-6 correlates with the pattern of *ERF*-1/*AP*-2γ expression. Further experiments to determine if *DEME*-6 is regulated by ERF-1/AP-2γ are currently under investigation. It seems likely that *DEME*-6 may be one of several genes that are coordinately regulated with *ER*.

DEME-6 was shown to be expressed in primary tumors, although a clear correlation with *ER* expression was not exhibited in these carcinoma samples. One possible explanation for this discrepancy might be that tumor tissue is comprised of a highly heterogeneous population of cell types. Alternatively, this pattern of expression may ultimately be found to be related to hormone-responsiveness. Previously it has been shown that the levels of *ER* expression in primary breast tumors is highly variable (38). Although this was a limited number of primary tumors, those tumors determined to be ER-positive by enzyme immunoassay that also had high *ER* mRNA levels were found to express *DEME*-6. It is, therefore, possible that transcriptional mechanisms in ER-positive, hormone-responsive tumors which result in over-expression of *ER* mRNA may also be involved in *DEME*-6 expression. Further studies examining the promoter of *DEME*-6 may help to clarify mechanisms regulating *DEME*-6 expression in primary breast cancers.

Sequence analysis demonstrates extensive homology of *DEME*-6 with the *C32D5.6* gene of *C.elegans*. *C32D5.6* was

cloned as part of the *C.elegans* chromosome III genomic sequencing project and is of unknown function (44). In addition to its correlation with ER-positive status, the importance of *DEME*-6 is underscored by the fact that it is evolutionarily conserved. Further analysis revealed that *DEME*-6 contains a eukaryotic putative RNA binding RNP-1 signature as well as a candidate anchoring transmembrance domain. This RNP-1 motif is seen in various proteins, such as heterogeneous nuclear ribonucleoproteins, small nuclear ribonucleoproteins and pre-RNA- and mRNA-associated proteins (45,46) and the transmembrane domain suggests a putative role as a membrane protein. Further experiments are required before a functional role of DEME-6 can be confirmed.

In summary, we have identified a panel of genes that are differentially expressed in ER-positive MCF7 cells as compared with ER-negative MDA-MB-231 cells. Studies with estradiol indicate that NPY Rc Y1 and three novel clones are ER-regulated genes. In addition, the expression of one estrogen-unresponsive gene, *DEME*-6, correlates with *ER* and *ERF*-1/*AP*-2γ expression in a panel of breast carcinoma cell lines. These findings identify an additional candidate gene that may be involved in the phenotypic differences observed between ER-positive and ER-negative breast carcinomas.

The nucleotide sequence of *DEME*-6 published in this paper has been submitted to the GenBank/EMBL database at NCBI and assigned accession no. AF007170. The nucleotide sequence of all other *DEME* clones have been submitted to the EST database (dbEST) at NCBI and assigned accession nos AA506763–AA506795 in chronological order.

ACKNOWLEDGEMENTS

We thank Clontech Laboratories Inc. for financing sequencing of *DEME*-6 and for technical advice from Dr Luda Diatchenko in regards to suppression subtractive hybridization. Breast cancer tissue was obtained from the Breast Cancer Resource of the Department of Pathology, NYU Medical Center. The Resource is funded by the Department of the Army, grant DAMD 17-94-J-4177. This work was supported in part by NIH grant R29 CA63251 and a grant from the Department of the Army, DAMD 17-94-J-4353. Financial support for W.W.K. was provided through an Award from the Stanford Medical Scholars Program and a Howard Hughes Research Training Fellowship for Medical Students. D.A.T. was funded by NRSA grant F32 CA69751-01. R.J.W. is supported in part by a Clowes Career Development Award from the American College of Surgeons.

REFERENCES

1 Fisher,B., Redmond,C., Fisher,E.R., Caplan,R. and other contributing National Surgical Adjuvant Breast and Bowel Project Investigators. (1988) *J. Clin. Oncol.*, **6**, 1076–1087.
2 Pichon,M.F., Broet,P., Magdelenat,H., Delarue,J.C., Spyratos,F., Basuyau,J.P., Saez,S., Rallet,A., Courriere,P., Millon,R. and Asselain,B. (1996) *Br. J. Cancer*, **73**, 1545–1551.
3 Edwards,D.P., Chamness,G.C. and McGuire,W.L. (1979) *Biochim. Biophys. Acta*, **560**, 457–486.
4 Osborne,C.K., Yochmowitz,M.G., Knight,W.A. and McGuire,W.L. (1980) *Cancer*, **46**, 2884–2888.
5 Henry,J.A., Nicholson,S., Farndon,J.R., Westley,B.R. and May,F.E.B. (1988) *Br. J. Cancer*, **58**, 600–605.
6 Knight,W.A., Livingston,R.B., Gregory,E.J. and McGuire,W.L. (1977) *Cancer Res.*, **37**, 4669–4671.
7 May,F.E.B., Johnson,M.D., Wiseman,L.R., Wakeling,A.E., Kastner,P. and Westley,B.R. (1989) *J. Steroid Biochem.*, **33**, 1035–1041.
8 Brown,A.M.C., Jeltsch,J.-M., Roberts,M. and Chambon,P. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6344–6348.
9 Moretti-Rojas,I., Fuqua,S.A.W., Montgomery,R.A. and McGuire,W.L. (1988) *Breast Cancer Res. Treatment*, **11**, 155–163.
10 Keaveney,M., Klug,J., Dawson,M.T., Nestor,P.V., Neilan,J.G., Forde,R.C. and Gannon,F. (1991) *J. Mol. Endocrinol.*, **6**, 111–115.
11 Piva,R., Bianchi,N., Aguiari,G.L., Gambari,R. and del Senno,L. (1993) *J. Steroid Biochem. Mol. Biol.*, **46**, 531–538.
12 Grandien,K. (1996) *Mol. Cell. Endocrinol.*, **116**, 207–212.
13 Thompson,D.A., McPherson,L.A., Carmeci,C., deConinck,E.C. and Weigel,R.J. (1997) *J. Steroid Biochem. Mol. Biol.*, **62**, 143–153.
14 Fuqua,S.A.W., Fitzgerald,S.D., Chamness,G.C., Tandon,A.K., McDonnell,D.P., Nawaz,Z., O'Malley,B.W. and McGuire,W.L. (1991) *Cancer Res.*, **51**, 105–109.
15 Fuqua,S.A.W., Fitzgerald,S.D., Allred,D.C., Elledge,R.M., Tandon,A.K., Nawaz,Z., McDonnell,D.P., O'Malley,B.W. and McGuire,W.L. (1992) *Cancer Res.*, **52**, 483–486.
16 Pfeffer,U., Fecarotta,E., Arena,G., Forlani,A. and Vidali,G. (1996) *J. Steroid Biochem. Mol. Biol.*, **56**, 99–105.
17 Pfeffer,U. (1996) *Ann. NY Acad. Sci.*, **784**, 304–313.
18 Hill,S.M., Fuqua,S.A.W., Chamness,G.C., Greene,G.L. and McGuire,W.L. (1989) *Cancer Res.*, **49**, 145–148.
19 Nembrot,M., Quintana,B. and Mordoh,J. (1990) *Biochem. Biophys. Res. Commun.*, **166**, 601–607.
20 Yaich,L., Dupont,W.D., Cavener,D.R. and Parl,F.F. (1992) *Cancer Res.*, **52**, 77–83.
21 Jiang,S.-Y. and Jordan,V.C. (1992) *J. Natl. Cancer Inst.*, **84**, 580–591.
22 Weigel,R.J. and deConinck,E.C. (1993) *Cancer Res.*, **53**, 3472–3474.
23 Ottaviano,Y.L., Issa,J., Parl,F.F., Smith,H.S., Baylin,S.B. and Davidson,N.E. (1994) *Cancer Res.*, **54**, 2552–2555.
24 deConinck,E.C., McPherson,L.A. and Weigel,R.J. (1995) *Mol. Cell. Biol.*, **15**, 2191–2196.
25 Williamson,J.A., Bosher,J.M., Skinner,A., Sheer,D., Williams,T. and Hurst,H.C. (1996) *Genomics*, **35**, 262–264.
26 Bosher,J.M., Totty,N.F., Hsuan,J.J., Williams,T. and Hurst,H. (1996) *Oncogene*, **13**, 1701–1707.
27 McPherson,L.A., Baichwal,V.R. and Weigel,R.J. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 4342–4347.
28 Maser,R.L. and Calvet,J.P. (1995) *Semin. Nephrol.*, **15**, 29–42.
29 Diatchenko,L., Lau,Y.C., Campbell,A.P., Chenchik,A., Moqadam,F., Huang,B., Lukyanov,S., Lukyanov,K., Gurskaya,N., Sverdlov,E.D. and Siebert,P.D. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 6025–6030.
30 Benton,W.D. and Davis,R.W. (1977) *Science*, **196**, 180–182.
31 Sanger,F. and Coulson,A.R. (1975) *J. Mol. Biol.*, **94**, 441–448.
32 Adams,M.D. *et al.* (1995) *Nature*, **377** (suppl. 6547), 3–174.
33 Bairoch,A., Bucher,P. and Hofmann,K. (1997) *Nucleic Acids Res.*, **25**, 217–221.
34 Carmeci,C., Thompson,D.A., Ring,H.Z., Francke,U. and Weigel,R.J. (1997) *Genomics*, **45**, 607–617.
35 Watson,J.B. and Margulies,J.E. (1993) *Dev. Neurosci.*, **15**, 77–86.
36 Trask,D.K., Band,V., Zajchowski,D.A., Yaswen,P., Suh,T. and Sager,R. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 2319–2323.
37 Castles,C.G., Fuqua,S.A.W., Klotz,D. and Hill,S.M. (1993) *Cancer Res.*, **53**, 5934–5939.
38 Carmeci,C., deConinck,E.C., Lawton,T., Bloch,D.A. and Weigel,R.J. (1997) *Am. J. Pathol.*, **150**, 1563–1570.
39 Hubank,M. and Schatz,D.G. (1994) *Nucleic Acids Res.*, **22**, 5640–5648.
40 Zhao,S., Ooi,S.L. and Pardee,A.B. (1995) *BioTechniques*, **18**, 842–850.
41 Masiakowski,P., Breathnach,R., Bloch,J., Gannon,F., Krust,A. and Chambon,P. (1982) *Nucleic Acids Res.*, **10**, 7895–7903.
42 Kalra,S.P., Fuentes,M., Fournier,A., Parker,S.L. and Crowley,W.R. (1992) *Endocrinology*, **130**, 3323–3330.
43 Horvath,T.L., Shanabrough,M., Naftolin,F. and Leranth,C. (1993) *Endocrinology*, **133**, 405–414.
44 Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M., Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J., Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A., Fraser,A., Fulton,L., Gardner,A., Green,P., Hawkins,T. and Hillier,L. *et al.* (1994) *Nature*, **368**, 32–38.
45 Bandziulis,R.J., Swanson,M.S. and Dreyfuss,G. (1989) *Genes Dev.*, **3**, 431–437.
46 Burd,C.G. and Dreyfuss,G. (1994) *Science*, **265**, 615–621.
47 Marck,C. (1988) *Nucleic Acids Res.*, **16**, 1829–1836.
48 Higgins,D.G., Bleasby,A.J. and Fuchs,R. (1992) *Comput. Applic. Biosci.*, **8**, 189–191.