A siRNA screen identifies RAD21, EIF3H, CHRAC1 and TANC2 as driver genes within the 8q23, 8q24.3 and 17q23 amplicons in breast cancer with effects on cell growth, survival and transformation

Sardar Faisal Mahmoon, Nadège Gruel, Elodie Chapeaublanc, Aurianne Lescure, Thouis Jones, Fabien Reyal, Anne Vincent-Salomon, Virginie Raynal, Gaëlle Pierron, Franck Perez, Jacques Camonis, Élaine Del Nery, Olivier Delattre, Virginie Raynal, Elaine Del, Fabien Reyal, Elodie Chapeaublanc, and Isabelle Bernard-Pierrot

1 CNRS, UMR 144, 2 Centre de Recherche, 3 Translational Research Department, 4 Biophenics Platform, Saint-Louis Hospital, 5 Surgery Department, 6 Anatomopathology Department and 7 INSERM, U830, Institut Curie, 26 rue d’Ulm, 75248 Cedex 05, Paris, France

Introduction

Identification of oncogene dependence, in which tumor cells become overly dependent on an activated oncogene for their proliferation and survival, could provide new targets for treatment (1). Gene amplification, leading to overexpression, is an important mechanism of oncogene activation in human solid cancers, including breast cancer (2–4). However, many passenger genes are amplified together with the tumor-promoting genes and functional studies are required to identify driver genes. In breast cancer, the major recurrent amplicons include 8p11-12, 8q24 (MYC), 11q13 (CCND1), 17q12 (ERBB2/HER-2) and 20q13 (5). The 17q12 amplicon has been extensively characterized and contains the ERBB2 gene, which provides one of the best examples of oncogene dependence and therapeutic efficacy (6). Two other genes within the 17q12 amplicon, GRB7 and STARD3, have also been shown to play a causal role in tumor development (7). Several recent studies have also identified key driver genes of the 8p11-12 amplicon, including PPAPDC1B and WHSC1L1, which present enzymatic activity and could, therefore, serve as potential therapeutic targets (8–10).

In this study, we aimed to identify new driver genes/potential therapeutic targets within the most frequent recurrent amplicons in human breast cancer, by following the same strategy we used to identify PPAPDC1B and WHSC1L1 as driver genes of the 8p11-12 amplicon, but applied at a larger scale (8). This approach involved the identification of amplified and overexpressed genes within recurrent regions of amplification, together with a high-throughput (HT) small interfering RNA (siRNA) screen for identifying genes for which silencing results in a decrease in cell viability. We focused on two chromosomes presenting regions frequently amplified in breast cancer: chromosomes 8 and 17. We defined eight recurrent minimal regions of amplification encompassing 101 genes overexpressed with respect to non-amplified tumors or normal samples. We then screened these regions, with a cell-based functional assay, in three breast cancer cell lines. We identified eight driver genes within five amplicons as critical for breast tumor cell proliferation or survival when amplified and overexpressed. We studied six of these genes in more detail (the other two being the well-known drivers of the 17q12 amplicon, ERBB2 and GRB7). We showed that one gene was a false-positive driver gene coming out from the screen and one gene was an essential gene required for cell survival independently of its expression level. We demonstrated that four genes induced cell apoptosis and/or cell cycle arrest following the knockdown of their expression in a cell line in which they were amplified and overexpressed. Finally, we demonstrated that these genes were also critical for cell transformation, by assessing cell-attachment-independent growth.

Materials and methods

Breast tumor samples and cell lines

We used frozen samples of 185 primary T1T2 infiltrating ductal breast carcinomas and 11 normal breast tissue samples. Normal breast samples were obtained during reduction mammoplasty. Samples of tumor or normal breast tissues were flash frozen after surgery and stored at −80°C. All tumor samples contained >50% cancer cells and the normal samples contained >50% breast epithelial cells, as evaluated by anatomopathological examination. This study was approved by the institutional review board of Institut Curie. In total, 23 cell lines derived from sporadic ductal breast carcinomas were obtained from the American Type Culture Collection (18485, BT474, BT549, CAMA1, HCC1937, H5787, MCF 10A, MCF7, MDA-MB-134VI, MDA-MB-175VI, MDA-MB-231, MDA-MB-361, MDA-MB-415, MDA-MB-436, MDA-MB-453, MDA-MB-468, SKBR3, T47D, ZR-75-1, ZR-75-30, SUM149 and HCC1500).

DNA and RNA extraction from breast tumor samples and cell lines

DNA and RNA were extracted from frozen samples of breast carcinomas and normal breast by the cesium chloride method (11,12). RNA was isolated from cell lines with RNeasy minikit (Qiagen, Courtaboeuf, France), whereas DNA was extracted with a conventional phenol–chloroform procedure.

Comparative genomic hybridization array

The 3.4K CIT (Carte d’Identification des Tumeurs program of the Ligue Nationale Contre le Cancer) BAC array and the experimental procedures for spotting, hybridization and washing were as described previously (13).

Data analysis was based on the normalized ratios of Cy5/Cy3 signals observed for each BAC clone that previously passed the flag assessment procedure. For autosoal chromosomes, the loss of a given locus was defined by a ratio ≤0.8, a gain was defined by a ratio ≥1.2 and <2.0 and an amplification was defined by a ratio ≥2.0. Whenever a statistical hypothesis concerning the levels of the ratios was tested, the hypothesis was considered to be rejected for P values ≤0.05. Whenever P values were adjusted for multiple hypothesis
testing, the false discovery rate using the Benjamini–Hochberg procedure was applied (R-Multitest package) (14). For analysis of amplicons, the identification criteria were BAC loci consisting of at least one clone with a ratio ≥2.0 in one tumor. The minimal region of amplification was then determined. The comparative genomic hybridization (CGH) data for chromosomes 8 and 17 and the corresponding expression datasets used in this study are available upon demand.

For MCF7, SKBR3 and HCC1937, high-resolution CGH data from Affymetrix SNP6 arrays were obtained and treated by the Broad-Novartis cancer cell line encyclopedia as described by Barretina et al. (15). Data were downloaded from the website: http://www.broadinstitute.org/cell/home.

DNA microarrays and probe set filter criteria

Microarray experiments were conducted according to the manufacturer’s instructions and analyzed as described previously (16). For tumors, we used the Affymetrix Human Genome U133 set (HG-U133 A and B), consisting of two gene chip arrays containing 44 692 probe sets, corresponding to >16 000 unique genes, for DNA microarray analysis. Gene expression were computed by RMA (Robust Multi-array Average) algorithm (17) and quantile normalization (18) using gcma R package and the official Affymetrix probe set definition file (CDF). Data were generated by using R language environment (available at: http://www.r-project.org/) and bioconductor packages (available at: http://www.bioconductor.org/). For the SKBR3, MCF7 and HCC1937 cell lines, we used the Affymetrix Human Exon 1.0 ST DNA microarray according to the manufacturer’s instructions. Gene expression data were normalized and summarized using RMA algorithm with custom chip definition developed by Microarray Lab, BrainArray (19). BrainArray annotation ENTREZGENE (version 12, available at: http://brainarray.ncbi.nlm.nih.gov/brainarray/Database/CustomCDF/CDF_download.asp#v12) providing one remapped probe set per gene according to National Center for Biotechnology Information (NCBI) Homo sapiens ENTREZGENE build 36.1 was used. One log2-transformed signal value per gene was obtained.

Cell culture

SKBR3 and HCC1937 cells were cultured in RPMI supplemented with 10% fetal calf serum and 2 mM t-glutamine; whereas MCF7 cells were grown in Dulbecco’s modified Eagle’s medium with 10% fetal calf serum. All reagents were purchased from Gibco-BRL (Cergy Pontoise, France).

HT siRNA screening

A siRNA library of 404 siRNAs against 101 genes (four siRNAs per gene) was obtained from QIAGen in ‘Flexible siRNA format’. We also included a negative siRNA control (anti-luciferase, GL2) and a positive lethal siRNA (anti-KIF11) on each plate. Transfection was optimized for each cell line in order to obtain a transfection efficiency >80% and cell toxicity <5%. Transfection efficiency was estimated by measuring GM130 knockdown by immune staining and toxicity was estimated by comparing 4′,6-diamidino-2-phenylindole (DAPI) staining for negative siRNA-treated cells and for non-treated cells. The transfection reagent HiPerfect (0.5 μl per well; QIAGen) was used for the SKBR3 cell line, INTERFERin (0.2 μl per well; Polyplus Transfection) was used for MCF7 and Lipofectamine 2000 (0.25 μl per well; Invitrogen) was used for HCC1937 transfection. Forward transfection was performed for SKBR3 and MCF7, with 7500 and 5000 cells per well, respectively, whereas reverse transfection was performed for HCC1937, with 5000 cells per well. A robot was used for the automatic distribution of reagents on the 96-well plates and, 72 h after transfection, immunofluorescence staining was performed with an antibody specific for Ki67 (a proliferation marker) and DAPI (cell count). Images were acquired with an IN Cell ANALYZER 1000 (GE Healthcare).

Data analysis and hit calling

Data were first transformed using log or logit functions. B-score normalization was then applied to each replicate, separately, and including corrections for plates, rows and columns (20, 21). Median and median absolute deviation (MAD) were computed and used to compute Robust Z-scores (RZ-scores) for each expression dataset used in this study, according to the formula (R-score – median)/median absolute deviation (MAD) (22). A gene was identified as a ‘hit’, if the RZ-score for at least two of four siRNAs was >2 or ≤−2 in at least two of three replicates. Hits for altered driver oncogenes were further validated with two siRNAs per gene, purchased from Qiagen: C8orf83-1: CACCTGGATGACAGAAGCTCA, C8orf83-2: CTGCACTATCATTACCTAAT, CHRAC1-1: CAGCTATATC CGAGGCTTAA, CHRAC1-2: TAGACGAGCAGCCAGGTA, EFDH1-1: ATGGCCAGCGCTGCTGAGTA, EFDH1-2: CCAAGAGTCCTCCTCA ATAA, RAD21-1: CCTCCAAATCTGTCGACTAA, RAD21-2: ATCG ATGACACCTATTAGA, TAN2-1: TAGGTCTGGAGGAACACATAA, TAN2-2: AAGGCTTACTGCGAGCAAA.

Quantitative real-time reverse transcription–PCR

Reverse transcription was performed with 1 μg of total RNA and a high-capacity cDNA reverse transcription kit (Applied Biosystems). PCR was carried out in a Roche® Light Cycler 480 real-time thermal cycler, with SYBR Green Master Mix (Roche®). The sequences of the primers used are available upon request.

Proliferation assay

Cells were plated (1.5 × 10^4 per well) in a six-well plate and transfected with 20 nmol/l siRNA in the presence of the appropriate transfection reagent, as determined for HT screening. Cells were treated with trypsin 72 h after transfection and stained with trypan blue. Viable cells were counted in triplicate using a Malassez hematocytometer.

Soft agar assay

For MCF7 and HCC1937, 20 000 siRNA-transfected cells in Dulbecco’s modified Eagle’s medium or RPMI 1640 supplemented with 10% fetal calf serum and 0.3% agar were added to triplicate wells containing medium and 0.8% agar, in 12-well plates. The plates were incubated for 14–21 days and colonies with diameters >50 μm under a phase-contrast microscope equipped with a measuring grid were scored as positive.

Flow cytometry analysis of the cell cycle and apoptosis

MCF7 and HCC1937 cells were transfected with 20 nmol/l siRNA against target genes in six-well plates. The medium from each well was transferred to a column tube 72 h after transfection. Cells were released from the wells by trypsin digestion and added to the corresponding collection tubes. Cells were fixed in 70% cold ethanol, treated with RNase A and stained with propidium iodide (PI), by incubation with a 200 μl Hanks' balanced salt solution/PI suspension [final concentration of PI: 20 μg/ml; RNase A (C): 10 mg/ml (Roche® already present)]. DNA content was analyzed in a FACS Caliber system. Data were collected and processed with ModFit (cell cycle) and FlowJo (apoptosis) software.

Statistical analysis

For each gene in the commonly amplified regions, the RNA levels of the amplified samples, as determined with Affymetrix arrays, were compared with those of tumor samples with no change in DNA copy number for the region considered and with those of normal samples, using the Wilcoxon rank sum tests.

All functional experiments were carried out in triplicate and repeated two or three times. Data are expressed as means ± SD. The data were subjected to statistical analysis by Student’s t-test. The control siRNA group was used as the reference. P values <0.05 were considered significant.

Results

Identification of candidate driver genes on chromosomes 8 and 17 in human breast cancer

Regions of recurrent amplification in human breast tumors were identified by assessing copy number alterations by CGH on DNA arrays (array-CGH hybridization) in 185 ductal breast carcinomas and 23 breast tumor-derived cell lines. Amplified regions were defined as regions with an array-CGH log2 signal >2.0 (more than four copies). We focused on two frequently altered chromosome arms, chromosomes 8q and 17q that were amplified in, respectively, 11% (21/185) and 14% (26/185) of breast tumors in our data set. For each chromosome arm, we defined the minimal recurrent regions of amplification and determined the genes included in these regions, according to the Build NCBI 136 Human (Hg18) Reference Sequence from NCBI (March 2006), as illustrated in Figure 1A for ampiclon A5 on chromosome 17 (17q23). In total, 10 recurrent amplicons were identified: five on chromosome 8q and five on chromosome 17q, corresponding to 143 amplified genes (Table 1). For two amplicons, none of the 23 cell lines analyzed by CGH displayed amplification, so no further cell-based functional studies of these amplified regions was possible. We, therefore, studied eight amplicons (four on chromosome 8q and four on chromosome 17q) containing, in total, 117 amplified genes (Table 1).

These eight recurrent regions of amplification presumably include both tumor-promoting and passenger genes. We hypothesized that tumor-promoting driver genes would be overexpressed due to DNA
Fig. 1. Strategy for candidate driver gene identification in breast carcinomas. Example of the recurrent 17q23 amplicon. (A) Genetic aberrations on chromosome 17q23 in breast cancers. CGH array of chromosome 17q23 in breast cancer samples (upper panel) and cell lines (lower panel). Cell lines are ordered as follow: MDA-MB-436, MDA-MB-134, MDA-MB-415, 184B5, BT549, CAMA1, MDA-MB-468, T47D, MDA-MB-175, BT474, Hs578T, MDA-MB-435, HCC1500, MDA-MB-231, MCF10A, ZR75.3, HCC1937, MDA-MB-453, SKBR3, SUM149, ZR-75-1, MDA-MB-361 and MCF7. Each column represents a clone on the array. Clones are ordered according to genome position, based on the NCBI human genome reference sequence 36, May 2006 from centromere (left) to telomere.
amplification. Based on this hypothesis, we derived a short list of candidate driver genes from the list of amplified genes, by comparing levels of gene expression in tumors with and without amplification and in normal breast tissue, as illustrated for ampiclon A5 on chromosome 17q23 (Figure 1B). RNA levels were analyzed with Affymetrix DNA microarrays in 167 of the 185 breast carcinomas and in 11 normal breast samples (Supplementary Table 1, available at Carcinogenesis Online). The 97 genes, significantly more strongly expressed in tumors with amplification than in tumors without amplification or normal samples (Wilcoxon rank sum test, \( P < 0.05 \)), were considered as candidate driver genes for systematic screening in a loss-of-function test (Table I and Figure 1B; Supplementary Table 1, available at Carcinogenesis Online). Genes within the studied amplicons for which no probe set was available on Affymetrix U133 arrays (Supplementary Figure 3B) or for which no probe set was available on chromosome 8q21, EIF3H and RAD21 on chromosome 8q23-q24, CHRAC1 on chromosome 8q24.3, STAC2 on chromosome 17q12 and TANC2 on chromosome 17q23.

*HT siRNA screening of the identified candidate driver genes*

For the identification of driver genes inducing oncogene dependence when amplified in breast tumors, from the 101 candidates, we carried out loss-of-function screening with a HT siRNA platform. We selected three cell lines—SKBR3, HCC1937 and MCF7—for studies of the role of genes in at least one amplified and one non-amplified cell line per amplicon (Table 1). Robust HT siRNA assays were developed in which cells of the three cell lines were very efficiently transfected with siRNAs using lipid transfection reagents, in 96-well plates (see Materials and methods). Each gene was targeted by four independent siRNAs, resulting in a library of 404 (101 × 4) siRNAs. Positive (KIF11 siRNA) and negative (GL2 siRNA) controls were also added. The effect of silencing on cell growth and cell viability was addressed by high-content immunofluorescence imaging, after staining with a specific antibody against Ki67 (for cell growth) and DAPI (for cell viability) (Figure 2A). We performed three independent HT RNAi screens for each cell line, to obtain biological replicates. Data were normalized by Z-score analysis (see Materials and methods), and the functional effects of each siRNA were considered significant if the median Z-score was \(<-2\) or \(>2\) (corresponding to Z-score > 2 in at least two of the three experiments) (Figure 2B). Genes were identified as ‘oncogenic hits’ if at least two of the four siRNAs targeting them decreased significantly cell viability and/or cell proliferation. Five hits were identified in HCC1937, 29 in MCF7 and 11 in SKBR3. In total, 33 different ‘oncogenic hits’ were identified. Most were identified in only one cell line (25 of 33), four were common to two cell lines and four were common to all three cell lines (Table II). As expected, hits included several already well-known oncogenes/driver genes in breast carcinoma: ERBB2, GRB7 and MYC, confirming the efficacy of the screening procedure (Table II). Surprisingly, some genes (four in SKBR3 cells and 10 in MCF7) increased significantly cell viability or cell growth when silenced (Supplementary Table 2, available at Carcinogenesis Online). As an intermediate step in the validation of oncogenic hits, we determined the copy number of these genes and compared their levels of expression in cell lines in which they were identified as hits with those in normal breast cell lines (Figure 2C; Supplementary Table 3, available at Carcinogenesis Online). Indeed, if a hit was not overexpressed, the observed effect of its knockdown on cell growth or proliferation was considered to indicate a possible role of a gene essential for the survival of all cells, rather than that of an oncogenic driver gene. Alternatively, this effect might be due to off-target effects, particularly if the hit gene was expressed at a very low level (Affymetrix signals of \( <4 \) could be considered to correspond to noise).

Hits were, therefore, validated as oncogenic driver genes if they were more strongly expressed in the breast tumor cell lines in which they are hits than in normal breast cell lines and if this overexpression was linked to DNA amplification. In total, five hits in SKBR3, one hit in HCC1937 and three hits in MCF7 cells were validated on the basis of expression data analysis (Figure 2C and Table II). Finally, as RAD21 was a prevalidated hit in two cell lines, all these criteria resulted in the validation of eight different genes within five different recurrent amplified regions, the overexpression of which led to oncogene dependence in breast carcinomas (Table II). ERBB2 and GRB7 have already been identified as driver genes of the chromosome 17q12 ampiclon (7), so no further validation of these hits was necessary. We then carried out a functional validation of the role in breast cancer of the six other previously validated oncogene dependencies: CHRAC1 on chromosome 8q21, EIF3H and RAD21 on chromosome 8q23-q24, CHRAC1 on chromosome 8q24.3, STAC2 on chromosome 17q12 and TANC2 on chromosome 17q23.

Silencing of RAD21, EIF3H, CHRAC1 and TANC2 induces cell cycle arrest and apoptosis in breast tumor cells

We carried out functional validation studies for C8orf83, RAD21, EIF3H, CHRAC1 and TANC2 in the MCF7 and HCC1937 cell lines, making it possible to study CHRAC1 and TANC2 in one amplified and one non-amplified cell line and C8orf83, EIF3H and RAD21 in one DNA-copy-gain and one amplified cell line (Table II). For each gene, we carried out reverse transcription–quantitative PCR to check that its expression levels were consistent with the DNA copy number in these cell lines. This was found to be the case for all genes except C8orf83, which was expressed very weakly in HCC1937 cells (4-fold less) than in MCF7 cells, despite the similarity of the log2(DNA copy number) values for these two cell lines (0.85 versus 1) (Figure 3A). We validated these five driver gene hits further, by resynthesizing two siRNAs/gene and confirming, by reverse transcription–quantitative PCR, that transfection with each specific siRNA markedly decreased messenger RNA (mRNA) levels (75–85% inhibition; Supplementary Figure 1, available at Carcinogenesis Online). The silencing of a single gene had no significant effect on the expression levels of the other four genes (data not shown). Then, using different readouts from those used for the HT screening (cell counts and cell cycle analysis by fluorescence-activated cell sorting after PI incorporation here, versus DAPI and Ki67 labeling for the screen), we demonstrated that the specific knockdown of CHRAC1 expression with two different siRNAs yielded significantly fewer viable HCC1937 cells than were obtained with the control siRNA (50–60%) (Figure 3B), whereas this knockdown had a very limited effect on MCF7 cells, in which this gene was not amplified and no more strongly expressed than in normal cells (Supplementary Table 3, available at Carcinogenesis Online). These results suggest that the effects observed following transfection with a specific siRNA resulted from specific gene silencing rather than an off-target effect and that CHRAC1 is not an essential gene. The same conclusion could be drawn for TANC2 knockdown, which induced a significant decrease in cell viability (50–60%) in MCF7 cells, in which it was amplified and overexpressed, whereas it had only a limited effect on HCC1937 cells (Figure 3B), in which it was not amplified or overexpressed (Supplementary Table 3, available at Carcinogenesis Online). Knockdown of the expression of EIF3H and RAD21 decreased viability in both cell lines (50–70%) (Figure 3B), consistent with the level of expression and DNA amplification/gain for these genes in these cell lines (Figure 3A; Supplementary Table 3, available at Carcinogenesis Online). EIF3H and RAD21 belong to the same ampiclon, 8q23-q24, but no additional or synergistic effect

(right). Log2 DNA copy numbers are shown on a color scale. White indicates an absence of data. Amplified regions were defined by an array-CGH log2 signal \( \geq 2.0 \) (more than four copies). A commonly amplified region, which extends from 57.82 to 59.11 Mb, was present in seven of the 185 infiltrating ductal breast carcinomas (3.8%) and two of the 23 breast cancer cell lines (8.7%) studied. Known genes present in the common region of amplification, based on build 36 from NCBI, are listed. (B) Identification of candidate driver genes within the minimal region of amplification on 17q23, on the basis of expression data. Comparison of gene expression levels between tumors presenting DNA amplification for 17p23, tumors without amplification of 17q23 and normal samples, based on Wilcoxon rank sum tests. \( *P < 0.05 \); **\( P < 0.001 \); ***\( P < 0.0001 \). In cases in which several probe sets were available for a gene, we selected the set displaying the most significant difference to ensure that no candidate gene was excluded for spurious reasons, n.s., non-significant.
Table I. Candidate driver genes within the minimal recurrent regions of amplification on chromosomes 8q and 17q in breast ductal carcinomas for screening in a HT cell-based functional assay

<table>
<thead>
<tr>
<th>Minimal region of amplification</th>
<th>Amplicon</th>
<th>start_pos (Mb)</th>
<th>end_pos (Mb)</th>
<th>Amplified tumors (%)</th>
<th>Amplified cell lines (%)</th>
<th>Genes in amplicon</th>
<th>No. of genes</th>
<th>Candidate driver genes</th>
<th>No. of genes screened</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome 8q</td>
<td>A1/8q21</td>
<td>91.06</td>
<td>95.49</td>
<td>6/185 (3.2%)</td>
<td>1/23 (4.3%)</td>
<td>C8orf83, CALB1, CDH17, DECR1, EFCBP1, FAM92A1, GEM, LRRC69, NBN, NECAB1, OSGIN2, PPM2C, RAD54B, RBM12B, RUNX1T1, SLC26A7, TMEM55A, TMEM64, TMEM67</td>
<td>19</td>
<td>C8orf83, CALB1, CDH17, DECR1, EFCBP1, GEM, NBN, NECAB1, OSGIN2, PPM2C, RAD54B, RBM12B, RUNX1T1, SLC26A7, TMEM55A, TMEM64, TMEM67</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>A2/8q22</td>
<td>97.06</td>
<td>99.73</td>
<td>7/185 (3.8%)</td>
<td>0</td>
<td>C8orf47, GDF6, HRSP12, KCNS2, LAPT4M4B, MATN2, MTDH, MTERFD1, NPA2, PGCP, POP1, PTDS51, RPL130, SDC2, STK3, TSPYL5, UQCRB</td>
<td>17</td>
<td>No amplified cell line for screening</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>A3/8q23-q24</td>
<td>116.49</td>
<td>120.23</td>
<td>7/185 (3.8%)</td>
<td>2/23 (8.7%)</td>
<td>SKBR3, HCC1937, C8orf85, COLEC10, EIF3H, EX1T, MED30, RAD21, SAMD12, SLC30A8, TNFRSF11B, TRPS1, UTP23</td>
<td>11</td>
<td>UTP23, COLEC10, EIF3H, EX1T, RAD21, SAMD12, THRAP6, TNFRSF11B, TRPS1</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>A4/8q24.21</td>
<td>127.79</td>
<td>128.80</td>
<td>8/185 (4.3%)</td>
<td>2/23 (8.7%)</td>
<td>SKBR3, HCC1937, MYC, POU5F1</td>
<td>2</td>
<td>MYC, POU5F1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>A5/8q24.23, 8q24.3</td>
<td>138.57</td>
<td>142.46</td>
<td>9/185 (4.9%)</td>
<td>1/23 (4.3%)</td>
<td>HCC1937, C8orf17, C8orf60, C8ORF32, CHRAC1, COL22A1, DENND3, EIF2C2, NIBP, GRF20, KNCN9, LOC286109, PTK2, SLC45A4, TRAPPC9</td>
<td>14</td>
<td>C8orf17, CHRAC1, COL22A1, DENND3, EIF2C2, KNCN9, NIBP, PTK2, SLC45A4, TRAPPC9, C8orf60</td>
<td>12</td>
</tr>
</tbody>
</table>
Table I. Continued

<table>
<thead>
<tr>
<th>Amplicon</th>
<th>start_pos (Mb)</th>
<th>end_pos (Mb)</th>
<th>Amplified tumors (%)</th>
<th>Amplified cell lines (%)</th>
<th>Genes in amplicon</th>
<th>No. of genes</th>
<th>Candidate driver genes</th>
<th>No. of genes screened</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome 17q A2/17q12</td>
<td>34.43</td>
<td>35.24</td>
<td>17/185 (9.2%)</td>
<td>4/23 (17.4%)</td>
<td>MDA-MB-361, MDA-MB-453, BT474, SKBR3</td>
<td>ARL5C, C17orf37, CACNB1, CDK12, ERBB2, FBXL20, GRB7, IKZF3, NEUROD2, PERLD1, PLCDC1, PNNMT, PPARGP2, PPP1R1B, PPP1R1B9, RPL19, STAC2, STARD3, TACAP</td>
<td>19</td>
<td>ARL5C, C17orf37, CACNB1, CRKRS, ERBB2, FBXL20, GRB7, IKZF3, NEUROD2, PERLD1, PNNMT, PPARGP2, PPP1R1B, PPP1R1B9, RPL19, STAC2, STARD3, TACAP, FBXL20, DLX3, DLX4, FM117A, FLJ45513, ITGA3, MYST2, PDK2, PPP1R1B9, SAMDJ4, TAC4</td>
</tr>
<tr>
<td>A3/17q21</td>
<td>45.15</td>
<td>45.56</td>
<td>6/185 (3.2%)</td>
<td>2/23 (8.7%)</td>
<td>MDA-MB-361, BT474</td>
<td>DLX3, DLX4, FM117A, FLJ45513, ITGA3, MYST2, PDK2, PPP1R1B9, SAMDJ4, TAC4</td>
<td>10</td>
<td>No amplified cell line for screening</td>
</tr>
<tr>
<td>A4/17q23-q23</td>
<td>51.36</td>
<td>53.67</td>
<td>4/185 (2.2%)</td>
<td>0</td>
<td>MDA-MB-361</td>
<td>CUEDC1, DYNLL2, EPX, MKS1, MRP23, OR4D1, OR4D2, SRSF1, VEZF1</td>
<td>9</td>
<td>ACE, CYB561, EFCA3, KCHN6, MAPK3, MARCH10, METTL2A, MRC2, TACO1, TANC2, TLK2, WDR68, CCDC44</td>
</tr>
<tr>
<td>A5/17q23</td>
<td>57.82</td>
<td>59.11</td>
<td>7/185 (3.8%)</td>
<td>2/23 (8.69%)</td>
<td>MCF7, MDA-MB-361</td>
<td>ACE, CYB561, EFCA3, KCHN6, MAPK3, MARCH10, METTL2A, MRC2, TACO1, TANC2, TLK2, WDR68, CCDC44</td>
<td>13</td>
<td>No amplified cell line for screening</td>
</tr>
<tr>
<td>A6/17q25</td>
<td>73.84</td>
<td>76.32</td>
<td>3/185 (1.6%)</td>
<td>1/23 (4.34%)</td>
<td>SKBR3</td>
<td>C17orf27, C1QTNF1, CANTI, CARD14, CBX2, CBX4, CBX8, CCDC40, CYTH1, DNAH17, EBF4A3, ENPP7, FLJ21865, FLJ355220, KIAA1303, KIAA1618, LOC146713, LOC284120, NPTX1, PSG1, SDHA, GAA, SGSH, SOCS3, TBC1D16, TIMP2, USP36, SLC26A11, LGALS3BP</td>
<td>29</td>
<td>No amplified cell line for screening</td>
</tr>
</tbody>
</table>

Total | | | | | | 143 | 101 |

Genes within the minimal region of amplification and overexpressed in amplified tumors with respect to non-amplified tumors or normal samples were considered for screening (Supplementary Table 1, available at Carcinogenesis Online).
was observed when the expression of both genes was knocked down simultaneously (Figure 3C). For each gene, both siRNAs had the same effect on cell proliferation; we, therefore, assumed that this was a specific, rather than an off-target effect (Figure 3B). However, as we did not test any non-amplified cell lines, we were unable to exclude the possibility of these genes being essential genes. A similar conclusion was raised for C8orf83 for which expression knockdown resulted in a decrease in cell viability in both cell lines (60–90%) in which it had been gained or amplified (Figure 3B). Surprisingly, however, the strongest effect was observed in the HCC1937 cell line, in which

---

**Fig. 2.** HT siRNA screening of candidate driver genes in breast cancer cell lines. (A) Schematic representation of the screen. HT siRNA screen based on immunofluorescence imaging. Identification of driver genes, the silencing of which causes cell cycle arrest and/or a decrease in cell viability. (B) Dot plot of the screens. Each screen was repeated three times to obtain biological replicates. Results are represented as median Z-scores for each siRNA. siRNAs with median Z-scores >2 or <−2 were considered significant hits, and genes with at least two hits for four siRNAs were considered to be hits genes (Table II). Genes for which knockdown induced an increase in cell proliferation/cell survival were not considered as oncogenic hits. (C) Intermediate validation step for oncogenic hits based on mRNA levels and DNA copy number. Hit genes were considered to be overexpressed when the fold difference with respect to mean expression in three normal breast cell lines (MCF10, MCF12 and HMEC) was >1.5 (Log2(FC) > 0.6); FC, fold change. Hit genes were considered to be gained/amplified if the CGH log2 ratio was >0.6 [gained from 0.6 to 1 (three to four copies) and amplified when >1 (more than four copies)]. Thresholds are indicated by dashed lines. Hits were prevalidated if amplified and/or overexpressed. The names of the prevalidated hits are indicated on the graphs.

---

S.F. Mahmood et al.
<table>
<thead>
<tr>
<th>Hit gene</th>
<th>Amplicon</th>
<th>Screening readout</th>
<th>Cell count index</th>
<th>G$_0$ arrest</th>
<th>Overexpression as compared with normal cell lines</th>
<th>Cell count index</th>
<th>G$_0$ arrest</th>
<th>Overexpression as compared with normal cell lines</th>
<th>Cell count index</th>
<th>G$_0$ arrest</th>
<th>Overexpression as compared with normal cell lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>KIF11</td>
<td>10q24</td>
<td>$-8.1$ to $-8.0$</td>
<td>7.2</td>
<td>4.70</td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>C8orf83</td>
<td>8q21</td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>RBM12B</td>
<td>TME67</td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>CALB1</td>
<td>8q22</td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>EF53H</td>
<td>8q23</td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>RAD21</td>
<td>q24</td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>MYC</td>
<td>8q24.21</td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>CHRAC1</td>
<td>8q24.3</td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>COIL22A1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>PTK2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>TRAPPC9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>ARL5C</td>
<td>17q12</td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>CDK12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>ERBB2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>GRB7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>PPP1R1B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>RPL19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>STAC2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>FAM117A</td>
<td>17q21</td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>FLJ45513</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>MYST2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>TAC4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>ACE</td>
<td>17q23</td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>EF53B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>KCN6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>MARCH10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>MRC2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>TANC2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>C1QTNF1</td>
<td>17q25</td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>CARD14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>FLJ13522</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>SOC53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
<tr>
<td>USP36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Positive control</td>
<td>No No Yes No</td>
<td>Yes Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
<td>No Yes No Yes</td>
</tr>
</tbody>
</table>

siRNAs were considered as hits if median Z-scores were $<-2$ or $>2$. Genes were considered as oncogenic hits if at least two of the four siRNAs were siRNA hits. Number of siRNA hits/gene and Z-scores for the best two siRNA hits/gene are indicated. For each gene, in each screened cell line, we indicate hit status, gene DNA copy number status, as determined from publicly available SNP6 array data (Broad-Novartis CCLC website) [genes were considered amplified if the normalized SNP array ratio was $>1$ (more than four copies) and were considered gained if this ratio was between 0.6 and 1 (from three to four copies)] and overexpression status with respect to normal cell lines, determined with U133plus2.0 DNA array data [genes were considered overexpressed if the fold difference with respect to mean expression in three normal breast cell lines (MCF10, MCF12 and HMEC) was $>1.5$ (log2FC $>0.6$); FC = fold change]. Data for mRNA levels and DNA copy number data are listed in Supplementary Table 3, available at Carcinogenesis Online. NA, not available.
Fig. 3. Specific knockdown of the expression of RAD21, EIF3H, CHRAC1 and TANC2 decreases cell viability via the induction of cell apoptosis and/or cycle arrest and colony formation in anchorage-independent growth conditions in breast cancer cell lines in which these genes are amplified and overexpressed. (A) Prevalidated hit gene expression in MCF7 and HCC1937 cells. We assessed mRNA levels by reverse transcription–PCR. (B–E) Effect of prevalidated gene knockdown on cell viability, cell cycle progression and cell apoptosis. The effect of single gene silencing on cell viability was assessed by cell counting 72 h after
Driver genes of chromosome 8 and 17 amplicons in breast cancer

CHRAC1, RAD21, EIF3H and TANC2 are driver genes involved in clonogenic cell growth

We then explored the role of these four genes in cloning efficiency. As for cell viability, RAD21 and EIF3H knockdown significantly decreased the ability of cells to form colonies in anchorage-independent conditions, in soft agar (Figure 3F), for both HCC1937 and MCF7 cells, in which these genes have been gained or amplified. The silencing of TANC2 and CHRAC1 also significantly inhibited the growth on soft agar of cell lines in which these genes were amplified and overexpressed (MCF7 and HCC1937 cells, respectively) but had no effect on cells in which these genes were neither amplified nor overexpressed with respect to normal breast cell lines (HCC1937 and MCF7 cells, respectively; Figure 3F).

RAD21 and EIF3H are prognostic markers in breast cancer

As RAD21, EIF3H, CHRAC1 and TANC2 were identified as driver genes for which amplification/overexpression regulated cell proliferation, cell survival and cell transformation, we investigated the potential value of these genes as prognostic markers in breast ductal carcinomas. Using the Kmiplot website (Kmplot.com) (23), we found a significant association between the level of gene expression and relapse-free survival at 10 years in 1866 breast cancer patients, for RAD21 (log rank test, P = 2.7E-15) and EIF3H (log rank test, P = 0.002), whereas this association was not significant for TANC2 (P = 0.24; Supplementary Figure 2, available at Carcinogenesis Online). No data were available for CHRAC1.

Discussion

Now that transcriptomic and genomic alteration data are simultaneously available for many tumor types, it is relatively easy to identify a particular class of candidate oncogenes: those that are overexpressed when amplified. By applying this approach, we identified 101 candidate driver oncogenes for breast cancer, located in eight different amplicons, on chromosomes 8q and 17q. Using a HT screening platform and a siRNA approach (four siRNAs per gene) in three breast tumor-derived cell lines, HCC1937, MCF7, SKBR3, we identified 33 'oncogenic hits': genes for which downregulation was associated with a cell growth arrest (with at least two of the four siRNAs used in at least two replicates of the three screening replicates). These hits included eight genes (CHRAC1, EIF3H, ERBB2, GRB7, RAD21, STAC2 and TANC2) that were amplified and overexpressed in the cell lines in which they were identified as hits. We tested six of these eight genes individually (CHRAC1, EIF3H, RAD21, STAC2 and TANC2), using a set of two siRNAs per gene. The other two hits, ERBB2 and GRB7, are well-known drivers of the 17q12 amplicon and, therefore, required no further validation. Involvement in cell growth, when amplified and overexpressed, was confirmed for four of these six genes (CHRAC1, EIF3H, RAD21 and TANC2). We also found that these genes were involved in cell attachment-independent growth. Consistent with their role in breast tumor cell proliferation/survival, the levels of expression of two of these genes, RAD21 and EIF3H, were associated with relapse-free survival. STAC2 was demonstrated to be a false-positive hit, the knockdown of which did not alter cell viability in the cell line in which it was amplified and overexpressed. C8orf83 was demonstrated to be an essential gene, the knockdown of which induced cell apoptosis whatever its level of expression in the cells concerned. A flowchart of the study is provided in Figure 4.

EIF3H and RAD21 have both previously been described as overexpressed and contributing to breast tumorigenesis (24,25), but no link to the chromosome 8q23-q24 amplicon was established. We show here that in cell lines with either an amplicon (MCF7) or a gained region (HCC1937) containing both genes, these two genes contribute to tumorigenesis in a non-additive manner. We also show here, for the first time, that CHRAC1 and TANC2 may play a role in carcinogenesis. The functions of TANC2 are unknown and its name is based on the structure of the protein (tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 2). CHRAC1 is involved in chromatin assembly (26) and, like RAD21, in DNA repair (27,28). RAD21 overexpression may confer resistance to radiotherapy and chemotherapy (29,30). The overexpression/amplification of CHRAC1 and RAD21 were not exclusive events. It would be interesting to compare resistance to various treatments as a function of the expression of these two genes. The initial concept that an amplicon in a given tumor contains only one driver gene has been shown to be incorrect in several cases (ERBB2, GRB7 and STARD13 in the 17q12 amplicon being the first example, followed by the 8p11-12 amplicon and more recently the 9p24 amplicon) (7,9,10,31). We provide here an additional example of an amplicon containing more than one driver gene: the 8q23-q24 amplicon, with EIF3H and RAD21.

HT platforms are increasingly being used for the functional validation of candidate oncogenes. We used three different cell lines for this screening: HCC1937, SKBR3 and MCF7. The numbers of hits ('oncogenic' hits for which inactivation inhibited cell viability and 'non-oncogenic' hit for which inactivation enhanced cell viability) identified differed considerably between the three cell lines: five in HCC1937, 15 in SKBR3 and 37 in MCF7 (four hits being common to two cell lines and four to three cell lines). The number of hits may depend on transfection efficiency, growth and proliferation rates and the number of false-positive and false-negative results, which may also depend on the cell line.

We identified 33 different hits leading to cell growth arrest (oncogenic hits). However, we also, unexpectedly, identified 14 hits for which proliferation rates were higher than for cells transfected with random siRNA. There are two possible reasons for this paradoxical
S.F. Mahmood et al.

Fig. 4. Flowchart of the study.

- 185 breast tumors, 23 breast cell lines, 11 normal breast samples
- CGH data for chr.8q, 17q
- Affymetrix U133 data
- 8 amplicons
- 101 genes over-expressed as consequence of DNA amplification in tumors

Cell based RNAi screen

- MCF7
- SKBR3
- HCC1937
- 29
- 11
- 5
- Genes that inhibited cell viability/proliferation when silenced

Expression and DNA copy number analysis in cell lines

- MCF7
- SKBR3
- HCC1937
- 5
- 7
- 2
- Oncogenic Hit over-expressed as compared to normal cell lines
- 3
- 5
- 1
- Oncogenic driver Hit amplified and over-expressed as compared to normal cell lines

- 6 new driver gene hits
- 2 well known driver of 17q12 amplicon: ERBB2 and GRB7

Analysis of consequences of knockdown on: cell viability/ cell cycle/ apoptosis/ colony formation on agar

- STAC2: false positive hit on chr.17q12
- C8orf83: essential gene on chr.8q21
- RAD21 and EIF3H: driver genes on chr.8q23
- CHRAC1: driver gene on chr.8q24.3
- TANC2: driver gene on chr.17q23
result: tumor cells are not optimally fit for in vitro growth, and many of these unexpected hits are false positives or the observed effect actually highlights the both tumor suppressor genes and proto-oncogenes with the amplicon concerned. This suggests that the cellular effect of proto-oncogenes may be affected by amplicon size and, thus, by the presence or absence of tumor suppressor gene within the region of amplification. However, it should be noted that 10 of these 14 hits were found in MCF7 cells, the cell line with the highest total number of hits. This suggests that most of these 14 hits are probably false-positive hits.

Several oncogenes had already been identified in the amplicons studied here (EIF3H, ERBB2, GRB7, MYC, RAD21). Interestingly, all these genes were identified as hits in our screening, suggesting that the number of false negatives was low. Among the 33 oncogenic hits, we focused here on six genes, each of which was amplified and overexpressed in at least two of the three cell lines used (HCC1937, MCF7, and SKBR3). For future studies, it would also be interesting to investigate the two gained and overexpressed hits: TRAPP9 on chromosome 8q24.3 and FAM117A on chromosome 17q21. The case of MYC, which was identified as a hit but not found to be overexpressed as a result of DNA amplification, strongly suggests that there may also be other true driver oncogenes among the hits that were amplified but not overexpressed in these three cell lines. This intermediate step in hit validation based on DNA copy number and mRNA level was aimed to short list the number of hit genes to further study individually but it could have excluded real driver gene hits. Indeed, we determine, for a given gene, overexpression with respect to ‘normal’ cell lines. In fact, these cell lines are not completely ‘normal’ according to other criteria (genomic alteration, unlimited growth), and they may overexpress some important genes for growth or immortalization by mechanisms other than genomic amplification. Furthermore, these genes were overexpressed in tumors due to DNA amplification. It would, therefore, be worthwhile to carry out further validation for all these amplified oncogenic hits (TMEM67 and CALB1 on chromosome 8q21; KCNH6, MARCH10, EFCAB3 and MRC2 on chromosome 17q23; CARD14 on chromosome 17q25) and gained oncogenic hits (COL2A2A1 and PTK2 on chromosome 8q23.4) (Table II) to determine whether they are false positives, essential genes or real oncogenic driver genes inducing cell proliferation/transformation when amplified. Most of these genes have not yet been shown to display oncogenic activity. However, MRC2 (Endo180) has been shown to be overexpressed in basal breast tumors and amplified in a subset of these tumors and to favor cell migration and tumor growth in vitro (32). PTK2 (FAK) has also been shown to be overexpressed and amplified in a subset of breast tumors and to promote breast cancer initiation and progression in vitro and in vivo (for review, see ref. 33). If, therefore, seems likely that MRC2 and PTK2 have, like MYC, been falsely excluded as hits and that they may, therefore, be driver genes of the 17q23 and 8q23.4 amplicons, respectively, in breast cancer.

Our initial goal was to identify oncogenes that might serve as potential treatment targets, but none of the genes we identified as new driver genes (RAD21, EIF3H, CHRAC1, TANC2) are enzymes or have an extracellular domain, proteins with these characteristics constituting the two most obvious classes of treatment targets, as they can be targeted by small molecules and monoclonal antibodies, respectively. However, additional studies on these oncogenes and, in particular, the pathways in which they are involved, could lead to the identification of future treatment targets.

Supplementary material

Supplementary Tables 1–3 and Figures 1 and 2 can be found at http://carcin.oxfordjournals.org/

Funding

Higher Education Commission (HEC) of Pakistan; Association pour la Recherche sur le Cancer (ARC) France for PhD research (to S.F.M.).

Acknowledgements

We thank Caroline Spraggon, Andrew Cassidy, Alexander Graham and Paul Elvin for the tumor microarray experiments, which were performed by AstraZeneca (Cancer and Infection Research Area, Alderley Park, Macclesfield, Cheshunt SKO 4TG, UK). We thank the staff of Charles Decraene for cell line microarray experiments (Translational Research Department, Institut Curie, Paris, France).

Conflict of Interest Statement: None declared.

References


Received April 22, 2013; revised September 25, 2013; accepted October 14, 2013