Gene expression profilers and conventional clinical markers to predict distant recurrences for premenopausal breast cancer patients after adjuvant chemotherapy (CMF)

Emma Niméus-Malmström[1+], Cecilia Ritz[2+], Patrik Edén[2], Anders Johnsson[1], Mattias Ohlsson[2], Carina Strand[1], Görel Östberg[3], Mårten Fernö[1*], Carsten Peterson[2]

[1]Department of Oncology, Institute of Medical Sciences, University Hospital, Lund, Sweden

[2]Department of Theoretical Physics, Lund, Sweden

[3]Department of Pathology, Hospital, Halmstad, Sweden

+These authors have contributed equally

*Corresponding author: Mårten Fernö, Department of Oncology, Institute of Medical Sciences, Lund University Hospital, SE 221 85 Lund, Sweden. Email: marten.ferno@med.lu.se.

Abbreviations: CMF, cyclophosphamide, methotrexate and 5-fluorouracil, ER, estrogen receptor, PgR, progesterone receptor, SPF, S-phase fraction, TNM, tumor size, lymph node status, metastases, NPI, Nottingham prognostic index, OR, odds ratio, SNR, signal-to-noise ratio, PCA, principal component analysis, ANN, artificial neural network, ROC, receiver operating characteristics, SD, standard deviation, FDR, false-discovery rate, CI, confidence interval

Keywords: breast cancer, cDNA microarray, drug resistance, prognostic markers

Running title: Predicting distant recurrences in CMF treated breast cancer

ABSTRACT

A large proportion of breast cancer patients are treated with adjuvant chemotherapy after the primary operation, but some will recur in spite of this treatment. In order to achieve an improved and more individualized therapy, our knowledge in mechanisms for drug resistance needs to be increased. We have investigated to what extent cDNA microarray measurements could distinguish the likelihood of recurrences after adjuvant CMF (cyclophosphamide, methotrexate and 5-fluorouracil) treatment of premenopausal, lymph node positive breast cancer patients, and have also compared with the corresponding performance using conventional clinical variables.

We tried several gene selection strategies, and built classifiers using the resulting gene lists. The best performing classifier with odds ratio=6.5 (95% CI=1.4-62) did not outperform corresponding classifiers based on clinical variables. For the clinical variables, calibrated on the samples, either using all the clinical parameters or the Nottingham Prognostic Index (NPI) parameters, the areas under the ROC curve were 0.78 and 0.79, respectively. The odds ratios at 90% sensitivity were 15 (95% CI=3.1-140) and 10 (95% CI=2.1-97), respectively. Our data have provided evidence for a comparable prediction of clinical outcome in CMF-treated breast cancer patients using conventional clinical variables and gene expression based markers.

INTRODUCTION

Breast cancer is a heterogeneous disease with a large variability in clinical outcome. Adjuvant polychemotherapy [e.g. with cyclophosphamide, methotrexate and 5-fluorouracil (CMF)] or anthracycline-containing regimes, produce substantial reduction in recurrence and mortality. In the metaanalysis, performed by the Early Breast Cancer Trialist´s Collabration Group, the absolute improvement in 15-year breast cancer survival after adjuvant polychemotherapy was 10 % (from 58% to 68%) for patients under the age of 50 (1). Besides an improvement in clinical outcome, these figures indicate that a large proportion of the patients will never recur after the primary operation and do, consequently, not need any further treatment with unnecessary side effects. Also, a considerable proportion of the patients will recur despite treatment with adjuvant polychemotherapy. Substantial efforts have been made to identify the group that does not need adjuvant systemic therapy, and to explain mechanisms why some patients recur in spite of chemotherapy. Possible mechanisms for recurrence after treatment are low initial drug sensitivity or an acquired drug resistance. In order to achieve a more effective and individualized chemotherapeutic treatment of breast cancer patients in the future, it is essential to increase our knowledge in mechanisms responsible for drug resistance, and to define reliable indicators for response to therapy. Commonly accepted prognostic factors are lymph node status, tumor size, histological grade, and patient age. Predictors for the effect of endocrine treatment, currently used in clinical routine, are estrogen (ER) and progesterone receptor (PgR) status, and for the effect of monoclonal antibodies (trastuzumab) c-erbB-2 is used. Useful markers for resistance and/or sensitivity of chemotherapy (CMF and/or antracyclin based regimes) have not, so far, been identified. Some markers have shown promising results in a limited number of studies, e.g. thymidylate synthase and thymidine kinase (2-4), c-erbB-2 (5-7), p53 (8-11), topoisomerase IIα, and multidrug resistance-associated protein (8, 12-14).

The development of techniques for gene expression analyses enables an extensive characterization of malignant tumors. Studies using these techniques in breast cancer have shown distinct differences in gene expression profile between hereditary and sporadic breast cancer (15), and between ER positive and ER negative cancer (16, 17). Promising results have also been obtained for predicting clinical outcome (17-21), both in patients not treated with adjuvant therapy (19-21) and in patients treated with adjuvant therapy, endocrine, chemotherapy, or both (17-19). Furthermore, gene expression analysis have identified genes involved in mediating the response to cytotoxic drugs, e.g. 5-fluoruracil in breast and colorectal cancer cell-lines, and esophageal cancer (11, 22), cisplatin in esophageal cancer (22), anthracyclines in breast cancer cell-lines (23), and neoadjuvant taxane treatment in breast cancer (24). However, before the above mentioned results can be applied in clinical routine, the data needs to be confirmed, using other array platforms and other patient materials. Furthermore, one important issue concerns whether the gene expression analysis provides information about clinical outcome and treatment sensitivity, in addition to the information obtained by conventional clinical factors, already in routine use. A recent publication from our group (25) has stressed this issue, by showing that clinical markers have similar power in predicting breast cancer prognosis as cDNA microarray gene expression profilers, using publicly available data (20).

In this study, we have used cDNA microarray analysis to predict recurrences after adjuvant treatment of CMF in a well-defined cohort of patients (premenopausal and lymph node positive). The ability to predict recurrences after CMF was also evaluated using clinical markers, publicly available cDNA expression data used for predicting clinical outcome (20),

and a gene expression profile associated with response to chemotherapy, based on prior knowledge, obtained after literature search.

METHOD AND PATIENTS

PATIENT SELECTION

According to treatment guidelines in the regional care program for breast cancer in Southern Sweden issued 1991, premenopausal lymph node positive (N+) breast cancer patients were recommended postoperative radiation and adjuvant chemotherapy. Radiotherapy was delivered to ipsilateral axillary and supraclavicular lymph nodes and the remaining breast parenchyma after breast conservation surgery or thoracic wall after mastectomy. The absorbed target dose was 50 Gy in 25 fractions in one series during five weeks. The standard chemotherapy at that time period was nine cycles of CMF. Patients for the present study were stringently selected in a stepwise manner to fulfill the following criteria: premenopausal women with primary breast carcinoma, stage T1-3N1-2M0, diagnosed 1992-97, frozen primary tumor samples were still available, referred to the department of Oncology in Lund or Malmö for adjuvant radiotherapy, treatment with nine cycles of CMF, either distant recurrence within 40 months after completion of CMF or remained free from distant recurrence for 40 months or longer, good quality of extracted RNA, and successful hybridization. After this selection process (Figure 1) we ended up with 29 recurrences and 56 recurrence-free patients that were included in the analysis (Table 1). The study was approved by the ethics committee at Lund University.

CHEMOTHERAPY

Patients were treated with an intravenous CMF schedule; cyclophosphamide 600 mg/m^2, methotrexate 40 mg/m^2 and 5-fluorouracil 600 mg/m^2, on day 1, every 3 weeks, for 9 cycles.

According to the regional guidelines, chemotherapy should be started within one month after surgery. Radiotherapy was started within one month after initiation of CMF. During the five-weeks of radiotherapy, cyclophosphamide was given at a dose of 850 mg/m^2 every three weeks, while methotrexate and 5-fluorouracil were omitted. The delivered chemotherapy doses were calculated and could be retrieved in 83 of the 85 patient´s records. The actual dose intensities mg/m^2/week were calculated and showed to be almost identical in the two groups; 93% of the planned doses for recurrence-free patients compared to 92% of the patients with recurrences. The main toxicity of CMF treatment was leucopenia. Dose reduction due to leucopenia (white blood cells <3.0 x 10^9/L) was performed in 65% of the recurrence-free patients and in 60% of the patients that later developed distant recurrence (p=0.63, chi-square-test).

METHODS

Conventional prognostic and treatment predictive factors

Histological grade was re-evaluated for all the samples by the same observer according to Elston and Ellis (26). The grading procedure consisted of judgment of tubule formation, nuclear plemorphism, and mitotic count. Each of these morphological features was given a score of 1 to 3 points. The overall histological grade was obtained by adding these points, and was categorized as follows: grade 1, 3-5 points, grade 2, 6-7 points, and grade 3, 8-9 points. The Nottingham Prognostic Index (NPI) is a linear combination of lymph node status, tumor size, and histological grade, according to the formula (27):

NPI = 0.2×tumor size (in cm) + lymph node status + histological grade,

where lymph node status is 1 for node negative, 2 for 1-3 tumor-involved nodes and 3 when 4 or more nodes are tumor-involved.

ER and PgR were analyzed routinely, at the time of the primary operation, with enzyme immunoassay according to kit instructions (Abbott Laboratories, Diagnostic Division, Chicago, IL, USA), and expressed as fmol per mg cytosol protein. Receptor values above or equal to 25 fmol/mg protein were considered positive.

The analysis of S-phase fraction (SPF) was also performed as part of clinical routine in an Ortho Cytoron Absolute flow cytometer (Ortho Diagnostic Systems, Raritan, NJ, USA). Samples with an SPF $\geq$12% were classified as high SPF, and those samples with values below these levels as low SPF (28).

RNA isolation and cDNA Microarray

Total RNA was isolated from fresh frozen tumors using Trizol (Invitrogen, Carlsbad, CA) and purified with the RNeasy® Midikit (Qiagen Inc, Valencia, CA). RNA quality was assessed with an Agilent 2100 Bioanalyzer RNA 6000 Lab.Chip kit (Agilent Technologies, Palo Alto, CA) and six samples were excluded due to poor RNA quality. The protocol for cDNA microarray has been reported previously (29). Briefly, the arrays were spotted with 27,648 sequence-verified cDNA clones (Unigene). Labeled cDNA was produced using 25 μg of tumor RNA and 10 μg Stratagene Reference RNA (Stratagene, La Jolla, CA) by anchored primed reverse transcriptase using CyscriptRT from the Cyscribe post labelling kit and Cy5-dUTP or Cy3-dUTP (Amersham Biosciences, Piscataway, NJ). Agilent software (Agilent technologies, Palo Alto, CA) was used for fluorescence scanning at 5 μm resolution and Gene pix Pro software (Axon Instruments, Inc., Union City, CA) for image analysis.

Data Mining Methods

Gene expression analysis

Gene expression analysis proceeded in 3 steps: (i) preprocessing, (ii) selection of significant

genes, and (iii) construction of classifier.

(i) *Preprocessing.* The data was stored in BASE (30) (BioArray Software Environment) after the initial image processing step. Pearson correlations of log reference intensities were calculated for all pairs of assays. The mean Pearson correlation for an assay ranged from 0.88 to 0.93, except for two assays, which had average Pearson correlation 0.73 and 0.13, respectively. These two assays were excluded from the following analysis. In BASE, a LOWESS normalization was applied to the log ratios (31). Replicate measurements x_i of the same reporter on an assay were merged as in (32) and represented by a weighted mean $m=\Sigma_i w_i x_i/\Sigma_i w_i$, where the weight w_i is $\exp(-3u_i^{1/2}/|x_i-m|)$, the estimated uncertainty of a spot u is $\mathrm{SNR}_1^{-2}+SNR_2^{-2}$, and SNR_i is the signal to background noise ratio for channel i. The set of equations for m was solved numerically by simple iteration. The error of the merged value was defined as $U=1/\Sigma_i(1/u_i)+\Sigma_i w_i^2(x_i-m)^2/(\Sigma_i w_i)^2$. We then modified expression values according to an error model (29) where expression values x_i, now representing the value merged on reporter, with large uncertainties u_i were moved towards the weighted mean m across assays for that reporter. The modified expression value was given by $x_i'=w_i(x_i-m)$. After reducing the importance of low-quality measurements in this way, the quality weights were not used in the following analysis. Reporters were excluded if missing in more than 10% of the samples or if the standard deviation of the modified log ratios was less or equal to 0.3. After these steps, 4,484 reporters remained for further processing.

(ii) *Selection of significant genes*. Reporters were ranked according to the Pearson correlation between (modified) gene expression log ratios and the clinical outcome M (M=1 for recurrence and M=0 for no recurrence). The false-discovery rate (FDR), defined as the fraction of reporters having a Pearson correlation higher than a chosen cut-off value by chance (33), was estimated from the Pearson correlation density of 1000 sample label

permutations.

(iii) *Construction of classifier.* This step was done following closely what was done earlier (16, 34, 35). The top 10 or top 100 genes with the highest Pearson correlation to clinical outcome was subject to principal component analyses (PCA), and the principal components with largest eigenvalues were used for construction of a committee of artificial neural network (ANN) classifiers. The performance was tested by applying the committee of networks to blind tests. In Khan et al (35) and Gruvberger et al (16) single test sets were used. Our goal was to compare different classifier performances, and multiple test set divisions then provide more reliable estimates (36). The need to retain sufficiently large training sets motivated small test sets. However, this leads to large variation between test set results (36), and many random test sets must be considered. Already facing substantial computational costs when ranking genes and selecting ANN designs, we therefore adopted a slightly different approach, where the ANN output values for all test samples were compiled and finally used to produce a single test result, as an estimate of the average result. With this approach, the test set size no longer poses a major problem, and we adopted a leave-one-out procedure.

In the cross-testing scheme, every member of a pre-defined pool of different ANN designs (and a new set of genes) was considered for each new blind test selection. The pool contained all combinations of the following parameters: number of inputs = 2, 4, 6, 8, 10; number of hidden nodes = 0, 2; and weight decay parameter = 0, 0.01, 0.03, 0.1. Back propagation (with learning rate = 0.75 and momentum parameter = 0.1) was used to minimize the error function during 50 training epochs, and for each iteration the learning rate was decreased by a factor of 0.98.

The performance of the classifiers from the different gene sets was measured by the area under the receiver operating characteristics curve (ROC area) (37). We also calculated the odds ratios (OR:s) after setting the thresholds corresponding to 10% misclassified in the distant recurrence group. The interpretation of ORs is known to be delicate (38) but they are included here for easier comparison to other studies (20, 25). All ORs in this paper are calculated at 90% sensitivity, making the comparison between them more straightforward than in the most general case. Compared to the ROC area, the OR is closer to a clinical reality, where a decision threshold must be implemented, but sensitive to noise in the studied data set. The ROC area represents a performance average over a wide range of thresholds, and is therefore less sensitive to noise and may better indicate which classifying approach that has the highest potential.

Clinical variables analysis

Five samples missing histological grade were excluded from the clinical variables analysis. Missing values for SPF were replaced with the mean over all samples. All tumors were annotated as T1 (≤ 20 mm), T2 (> 20 mm-50 mm), or T3 (> 50 mm). In two cases, T stage was the only available information of tumor size. In order to get a numerical value of size for all samples, these two missing values were replaced by the mean size over the samples with the same T stage annotation (T1 and T3, respectively).

ER, PgR, SPF, and tumor size were in the statistical analysis used as continuous variables.

Two approaches were taken using conventional variables only. In the first one, the NPI (27) was computed for all patients, without any learning steps. In the other approach, ANN models were constructed according to the cross-testing and cross validation scheme above, with the exception that the PCA step was not performed, since there were only 7 variables (number of

tumor-involved lymph nodes, tumor size, histological grade, age, ER, PgR, and SPF). The second approach was employed using all seven clinical variables, and also using only the three parameters included in the NPI (number of tumor-involved lymph nodes, tumor size, and histological grade).

Hybrid classifiers were constructed in the same way as for the gene expression data, but with the clinical variables added as input nodes. The pool of ANN designs was identical except for the number of inputs. The hybrid classifier with 7 clinical variables had 9, 11, 13, 15, or 17 number of inputs, and the hybrid classifier with the NPI variables had 5, 7, 9, 11 or 13.

Search strategy and selection criteria for drug associated genes

Data for the list of known drug associated genes (drug-genes) was identified during December 2003 - February 2004 (see Supplement) in two ways. First, already available articles within this subject were selected. Secondly, published articles, since 1997, were obtained by two separate searches of PubMed (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi). The first one included the search terms ”drug resistance”, ”cancer”, ”cyclophosphamide”, ”methotrexate”, and ”5-FU”. The second search included ”drug resistance” and ”cancer”, in order to find genes involved in sensitivity or resistance to other regimes than CMF (see Supplement) .

The selection of genes was performed prior to the data analyses. To find the reporters on our array corresponding to the genes on our pre-defined gene lists, we used the official gene symbol. The gene symbols for the van 't Veer gene list (16) were obtained through ACID (39) using UniGene build 176, and the official gene symbols for the drug-genes were found manually using Gene and Locus Link. All reporters on our array that according to UniGene build 180 had a gene symbol represented on the resulting list were selected. Among the

14,717 reporters that had less than 10% missing values, 245 matched the 253 initially pre-selected drug genes and 184 matched the 231 van ´t Veer genes. We confirmed that the 184 van 't Veer-genes available in our study had similar predictive power in the data set of van 't Veer and co-workers, as had the full set of 231 genes.

RESULTS

False-discovery rates in the different gene sets. The FDR was 6% for the top-100 reporters in the full reporter list containing 4,484 reporters (Table 2), but noticeably higher in the two pre-selected reporter sets. Restricting the analysis (both gene ranking and permutations test) to the drug-genes gave a 42% FDR, and correspondingly a 31% FDR for the van 't Veer genes. The top-100 reporters from the different reporter lists can be found in the supplementary information.

When using the top-10 reporters, the FDR:s were lower (Table 2). The unrestricted top 10-list is listed in Table 3a, and this list includes genes involved in functions such as signaling, gamma-aminobutyric acid metabolism, RNA processing, N-linked glycosylation via asparagines, electron transport, nucleotide binding, activation of T and natural killer cells, ATP binding, and metalloendopeptidase inhibitor activity. The top drug-genes were according to earlier studies important for resistance mechanisms to doxo- and epirubicin treatment (10, 23), methotrexate (10, 40) and docetaxel (24), cisplatin (41), 5-FU (11), vincristine (10), vindesin (10), mitomycin C (10) and thiotepa (10), and they are involved in cell proliferation, RNA processing, DNA-damage response, nucleotide biosynthesis, N-linked glycosylation via asparagines, estrogen receptor signaling pathway, and anti-apoptosis (Table 3b).

Predictive power of the different gene sets. Using PCA and ANN, the different lists of top-10 and top-100 reporters were used to classify the two groups of patients, with and without distant recurrences after adjuvant CMF, after proper division of data into training and validation tests (see Methods). For each blind test sample, a new ranking of reporters was performed, based on the remaining samples. This was done to avoid information leaks in the

analysis. Thus, the resulting predictions were not a test of a specific top reporter list, but rather a test of the full reporter set from which top lists were generated. The result in terms of ROC area (see Methods) was higher for the drug-genes top-10 reporters than for the other two top-10 gene selections, which had both similar results. The OR:s were significantly above 1 (>95% confidence, Fisher's exact test) for the drug-genes top-10 reporters and for the unrestricted top-10 reporters. Selecting the top-100 reporters gave worse prediction performance for all 3 reporter sets, both in terms of ROC area and OR.

Predictive power of the clinical variables. When using the same ANN procedure to build a classifier, including leave-one-out, NPI parameters and clinical markers yielded ROC areas comparable to the drug-genes top-10 reporter result, and higher OR:s than all tested classifiers based on gene expression (Table 4). The classifier using all 7 clinical markers performs better than the one using only the 3 NPI parameters. Using NPI directly, without calibrating any classifier on the data set, improved the results in terms of ROC area further. Using both clinical and gene expression data in hybrid classifiers did, however, not improve the results.

Gene ontology

The three top 100 gene lists (unrestricted, drug, and van ´t Veer) were functionally classified by annotating the genes with gene ontology followed by clustering into biological processes. Out of the most frequent biological processes, three processes were found on all three gene lists, mitosis, cytokinesis, and regulation of cell cycle, which are all processes related to cell proliferation. Data also indicates that the drug and van ’t Veer lists are more similar since several processes such as cell cycle and cell growth maintenance were uniquely common in these two gene lists. Some processes were represented in only one of the gene lists. In the unrestricted gene list several biological processes involving signaling were more common,

whereas in the drug gene list, biological processes involving protein modifications and regulation of cell proliferation were found. In the van ´t Veer gene list no clear trend could be found due to too few processes present only for this list.

DISCUSSION

The present study was focused on trying to explain why certain patients recur in spite of adjuvant chemotherapy (CMF). Currently available conventional factors are not considered sensitive enough for this selection. We constructed classifiers based on conventional markers and gene expression as measured by cDNA microarrays. We found that gene expression data could not improve the predictions. The strength of the conventional markers in relation to the gene expression profile is thus a confirmation of the results from a previous paper from our group (25), using publicly available data of van ´t Veer and coworkers (20).

The incapacity of gene expression analysis to improve predictive power may simply be due to a too small cohort in the study; one could hypothesize that a marker based on multidimensional gene expression data would benefit more from a larger study than would already established clinical markers. Our studied cohort was large enough to identify genes relevant for development of distant recurrences after adjuvant CMF (6% FDR among top 100 ranked genes), but may still be too small to fully avoid overtraining when building the classifiers. The fact that a combination of clinical parameters and gene expression data failed to improve the results, and sometimes reduced them, points in this direction. The relatively poor performances of the hybrid classifiers should therefore not be seen as any evidence of complete overlap between information from clinical and gene expression based markers. Among the classifiers investigated here, the NPI is the only one that has been calibrated using a large cohort of several thousand samples (27), while all other classifiers were calibrated on the data set of this paper, consisting of 85 samples. The rather big difference in performance of the NPI (ROC area=0.79) and the classifier based on the three NPI parameters, but calibrated using ANNs on the current data set (ROC area=0.74), illustrates the importance of large sample cohorts.

The apparent need for large sample cohorts when using gene expression analysis may be explained by the heterogeneity of breast cancer, with many subpopulations. Among clinical variables, some markers (e.g. ER status) mainly distinguish disease subtypes which correlate to outcome, while other markers (e.g. tumor size) may correlate more directly to the progression of the disease. The huge amount of information embedded in genome-wide studies should, in principle, allow for extraction of both kind of markers in gene expression data, but it is not inconceivable that genome-wide profiling is more related to disease subtypes (16, 42) than to progression. If so, gene expression analysis may be better suited for studies aiming at an improved biological insight to the mechanisms behind the studied disease and its subtypes, potentially leading to the discovery of new drug targets and development of new therapeutic protocols. A possible way to improve gene expression analysis (both for direct marker design and for gain of biological insight) is to interpret microarray data not in terms of individual genes, but in a way closer related to the underlying biology, e.g. pathways (43).

As an initial step in exploiting prior knowledge, we used literature genes and a gene list from a differently selected cohort of breast cancers (van 't Veer). Also, we interpreted the results in terms of gene ontology categories and find some categories in common for the different gene lists. When studying the gene ontology of the three different top 100 lists (unrestricted, drug, and van ´t Veer), mitosis, cytokinesis, and regulation of cell cycle existed on all lists. Since all lists are created for use of predicting recurrences/drug resistance this indicates that these well-known tumor genesis processes are also important for recurring tumors. Worth mentioning is that the top 100 drug genes and van ´t Veer genes have more processes in common, in comparison to the unrestricted genes.

The design of our study, involving only homogeneously treated premenopausal lymph node positive patients, helps focus on a well-defined medical question, but also implies that the recurrence-free group consists of two subgroups, one with an inherited good prognosis (already being cured by the primary operation and postoperative radiotherapy) and one subgroup with inherited bad prognosis, but also CMF-sensitivity (which without adjuvant CMF would have developed recurrence). The group having developed recurrences may be more homogeneous (inherited bad prognosis and CMF-resistant), but heterogeneity may still be a problem, since drug resistance in many cases is acquired, i.e. changes in gene expression are developed after the administration of the drug. Our study has thus only tried to identify those patients recurring in spite of adjuvant CMF, and for which alternative treatments should be recommended. The design of our study makes it impossible to answer which patients do not need adjuvant systemic therapy and which patients benefit from adjuvant CMF. CMF has nowadays, to a large extent, been replaced by other and more effective cytostatic treatments, e.g. anthracycline or taxane based regimes, but two out of the three drugs included in CMF, cyclophosphamide and 5-fluorouracil, are also included in many anthracycline based regimes. The reasons for included patients treated with CMF in the present study were to obtain a long follow up time and enough cases with frozen tumor tissue available. We furthermore hypothesize that the concept to test gene expression profile as a prognostic marker after adjuvant CMF could be generalized to other cytostatic regimes.

It should be emphasized that we have not pursued a survival analysis, since as discussed above the objective was to construct a classifier for somewhat extreme cases. In part this choice of procedure was dictated by the limited data set at our disposal for this question of CMF resistance. Our comparisons of different classifiers are not very sensitive to excluding the patients that lacked follow-up to the time threshold.

In conclusion, we have confirmed the strength of conventional markers compared to gene expression profilers for prognostic considerations, shown by similar performance in predicting clinical outcome after adjuvant cytostatic (CMF) therapy. We have also stressed important issues when interpreting gene expression data, including gene selection, overtraining, and study design.

ACKNOWLEDGMENTS

This study was supported by Swedish Cancer Society, Gunnar, Arvid and Elisabeth Nilsson Foundation, Mrs Berta Kamprad Foundation, Swedish Research Council and the Swedish Foundation for Strategic Research through the Lund Center for Stem Cell Biology and Cell Therapy. We are also indebted to participating departments of surgery, oncology, and pathology of the South Sweden Breast Cancer Group, and Regional Tumor Registry for primary care and follow-up of the patients, and for handling of frozen breast cancer tissue.

FIGURE LEGENDS

Figure 1. The selection of patients included in the study.

Table 1. Clinical and biological characteristics of 85 premenopausal patients, with lymph node positive breast cancer, treated with adjuvant CMF, subdivided with respect to if they have developed distant recurrences or not.

Table 1

Clinical parameter	Distant rec	No rec
Age at diagnosis		
<40 years	6	9
40-50 years	16	44
>50 years	5	5
Tumor size		
T1, ≤ 20 mm	3	24
T2, >20-50 mm	21	33
T3, > 50 mm	0	1
missing value	3	0
Lymph nodes		
1-3 pos lymph nodes	16	45
≥4 pos lymph nodes	11	13
Histological grade		
1	1	12
2	2	15
3	23	27
missing value	1	4
ER		
<25 fmol/mg protein	19	20
≥25 fmol/mg protein	8	38
PgR		
<25 fmol/mg protein	18	22
≥25 fmol/mg protein	9	36
SPF		
<12%	7	29
≥12%	16	24
missing value	4	5

Tabel 2. Pearson correlation coefficient and number of false positive among 100-top genes for unrestricted, drug, and van ´t Veer genes.

Table 2

Reporter set	**Pearson correlation**	**False discovery rate (%)**
Unrestricted		
top 10	0.42	2.0
top 100	0.34	6.2
Drug-genes		
top 10	0.30	9.3
top 100	0.15	42
van ´t Veer		
top 10	0.34	1.8
top 100	0.15	31

Table 3. A list of the top-10 unrestricted genes (a) and drug genes (b) were ranked using Pearson correlation and classified with ANN. +/- indicates if the gene is up or down-regulated in the group with no distant recurrences.

a)

Gene name	Gene symbol	Acc number	Up/down
4-aminobutyrate aminotransferase	ABAT	BC008990	+
Serum/glucocorticoid regulated kinase-like	SGKL	H98714	+
Thyroid hormone receptor interactor 13	TRIP13	AA630784	-
Interleukin 12A (natural killer cell stimulatory factor 1)	IL12A	AI304577	-
Hypothetical protein FLJ40629	FLJ40629	AA417744	-
Dolichyl-diphosphooligosaccharide-protein glycosyltransferase	DDOST	H96437	-
Arginine-rich, mutated in early stage tumors	ARMET	R91550	-
RNA binding protein with multiple splicing	RBPMS	W67323	+
Chromosome 20 open reading frame 129	C20orf129	R96998	-
ERO1-like (S. cerevisiae)	ERO1L	AA186804	-

b)

Gene name	Gene symbol	Acc number	Up/down
Dolichyl-diphosphooligosaccharide-protein glycosyltransferase	DDOST	H96437	-
RNA binding protein with multiple splicing	RBPMS	W67323	+
Cell division cycle 27	CDC27	T81764	+
Baculoviral IAP repeat-containing 5 (survivin)	BIRC5	AA460859	-
Estrogen receptor 1	ESR1	AA291702	+
V-abl Abelson murine leukemia viral oncogene homolog 1	ABL1	H91096	-
Fusion (involved in t(12;16) in malignant liposarcoma)	FUS	W67581	+
X-ray repair complementing defective repair in Chinese hamster cells 1	XRCC1	AA425139	+
V-raf murine sarcoma viral oncogene homolog B1	BRAF	W88566	-
Dihydrofolate reductase	DHFR	N52980	-

Table 4. The effectiveness of variables for separating in recurrence *vs.* recurrence-free patient groups is measured using the ROC area and odds ratios (OR), using the top ranked reporters of the unrestricted (unrestr.), drug and van ´t Veer reporter sets, respectively. As a comparison, the corresponding values for NPI and the seven clinical variables, as well as the combinations of clinical variables and the unrestricted reporter set, are shown.

Table 4

Reporter set	**ROC**	**OR**	**95% CI (Fischer´s exact test)**
Unrestricted			
top 10	0.70	6.5	1.4-62
top 100	0.60	2.0	0.36-21
Drug-genes			
top 10	0.78	6.0	1.3-57
top 100	0.57	2.3	0.42-23
van ´t Veer			
top 10	0.69	3.9	0.80-38
top 100	0.65	1.9	0.36-21
Clinical variables and combinations			
All 7	0.78	15	3.1-140
incl.top 10 unrestr.	0.71	1.2	0.18-14
incl.top 100 unrestr.	0.66	1.5	0.24-16
3 NPI parameters	0.74	10	2.1-97
incl.top 10 unrestr.	0.72	5.0	1.0-48
incl.top 100 unrestr.	0.76	2.1	0.37-140
NPI	0.79	10	2.1-97

REFERENCES

1. Early Breast Cancer Trialists´Collaborative Group; Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005; **365**, 1687-717.

2. Clark JL, Berger SH, Mittelman A, et al. Thymidylate synthase gene amplification in a colon tumor resistant to fluoropyrimidine chemotherapy. *Cancer Treat Rep* 1987; **71**, 261-5.

3. Romain S, Martin PM, Klijn JG, et al. DNA-synthesis enzyme activity: a biological tool useful for predicting anti-metabolic drug sensitivity in breast cancer? *Int J Cancer* 1997; **74**, 156-61.

4. Washtien WL. Increased levels of thymidylate synthetase in cells exposed to 5-fluorouracil. *Mol Pharmacol* 1984; **25**, 171-7.

5. Di Leo A, Chan S, Paesmans M, et al. HER-2/neu as a predictive marker in a population of advanced breast cancer patients randomly treated either with single-agent doxorubicin or single-agent docetaxel. *Breast Cancer Res Treat* 2004; **86**, 197-206.

6. Konecny GE, Thomssen C, Luck HJ, et al. Her-2/neu gene amplification and response to paclitaxel in patients with metastatic breast cancer. *J Natl Cancer Inst* 2004; **96**, 1141-51.

7. Muss HB, Thor AD, Berry DA, et al. c-erbB-2 expression and response to adjuvant therapy in women with node-positive early breast cancer. *N Engl J Med* 1994; **330**, 1260-6.

8. el-Deiry WS. Role of oncogenes in resistance and killing by cancer therapeutic agents. *Curr Opin Oncol* 1997; **9**, 79-87.

9. Geisler S, Borresen-Dale AL, Johnsen H, et al. TP53 gene mutations predict the response to neoadjuvant treatment with 5-fluorouracil and mitomycin in locally advanced breast cancer. *Clin Cancer Res* 2003; **9**, 5582-8.

10. MacGrogan G, Mauriac L, Durand M, et al. Primary chemotherapy in breast invasive carcinoma: predictive value of the immunohistochemical detection of hormonal receptors, p53, c-erbB-2, MiB1, pS2 and GST pi. *Br J Cancer* 1996; **74**, 1458-65.

11. Maxwell PJ, Longley DB, Latif T, et al. Identification of 5-fluorouracil-inducible target genes using cDNA microarray profiling. *Cancer Res* 2003; **63**, 4602-6.

12. Burger H, Foekens JA, Look MP, et al. RNA expression of breast cancer resistance protein, lung resistance-related protein, multidrug resistance-associated proteins 1 and 2, and multidrug resistance gene 1 in breast cancer: correlation with chemotherapeutic response. *Clin Cancer Res* 2003; **9**, 827-36.

13. Fazeny-Dorner B, Piribauer M, Wenzel C, et al. Cytogenetic and comparative genomic hybridization findings in four cases of breast cancer after neoadjuvant chemotherapy. *Cancer Genet Cytogenet* 2003; **146**, 161-6.

14. Nooter K, Brutel de la Riviere G, Look MP, et al. The prognostic significance of expression of the multidrug resistance-associated protein (MRP) in primary breast cancer. *Br J Cancer* 1997; **76**, 486-93.

15. Hedenfalk I, Duggan D, Chen Y, et al. Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 2001; **344**, 539-48.

16. Gruvberger S, Ringner M, Chen Y, et al. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. *Cancer Res* 2001; **61**, 5979-84.

17. Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 2003; **100**, 10393-8.

18. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004; **351**, 2817-26.

19. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002; **347**, 1999-2009.

20. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; **415**, 530-6.

21. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005; **365**, 671-9.

22. Kihara C, Tsunoda T, Tanaka T, et al. Prediction of sensitivity of esophageal tumors to adjuvant chemotherapy by cDNA microarray analysis of gene-expression profiles. *Cancer Res* 2001; **61**, 6474-9.

23. Kudoh K, Ramanna M, Ravatn R, et al. Monitoring the expression profiles of doxorubicin-induced and doxorubicin-resistant cancer cells by cDNA microarray. *Cancer Res* 2000; **60**, 4161-6.

24. Chang JC, Wooten EC, Tsimelzon A, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* 2003; **362**, 362-9.

25. Eden P, Ritz C, Rose C, et al. "Good Old" clinical markers have similar power in breast cancer prognosis as microarray gene expression profilers. *Eur J Cancer* 2004; **40**, 1837-41.

26. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991; **19**, 403-10.

27. Blamey RW, Davies CJ, Elston CW, et al. Prognostic factors in breast cancer -- the formation of a prognostic index. *Clin Oncol* 1979; **5**, 227-36.

28. Sigurdsson H, Baldetorp B, Borg A, et al. Flow cytometry in primary breast cancer: improving the prognostic value of the fraction of cells in the S-phase by optimal categorisation of cut-off levels. *Br J Cancer* 1990; **62**, 786-90.

29. Andersson A, Eden P, Lindgren D, et al. Gene expression profiling of leukemic cell lines reveals conserved molecular signatures among subtypes with specific genetic aberrations. *Leukemia* 2005; **19**, 1042-50.

30. Saal LH, Troein C, Vallon-Christersson J, et al. BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data. *Genome Biol* 2002; **3**, SOFTWARE0003.

31. Yang YH, Dudoit S, Luu P, et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 2002; **30**, e15.

32. Fernebro J, Francis P, Eden P, et al. Gene expression profiles relate to SS18/SSX fusion type in synovial sarcoma. *Int J Cancer* 2006; **118**, 1165-72.

33. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 1995; **57**, 289-300.

34. Gruvberger-Saal SK, Eden P, Ringner M, et al. Predicting continuous values of prognostic markers in breast cancer from microarray gene expression profiles. *Mol Cancer Ther* 2004; **3**, 161-8.

35. Khan J, Wei JS, Ringner M, et al. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat Med* 2001; **7**, 673-9.

36. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005; **365**, 488-92.

37. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; **143**, 29-36.

38. Pepe MS, Janes H, Longton G, et al. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol* 2004; **159**, 882-90.

39. Ringner M, Veerla S, Andersson S, et al. ACID: a database for microarray clone information. *Bioinformatics* 2004; **20**, 2305-6.

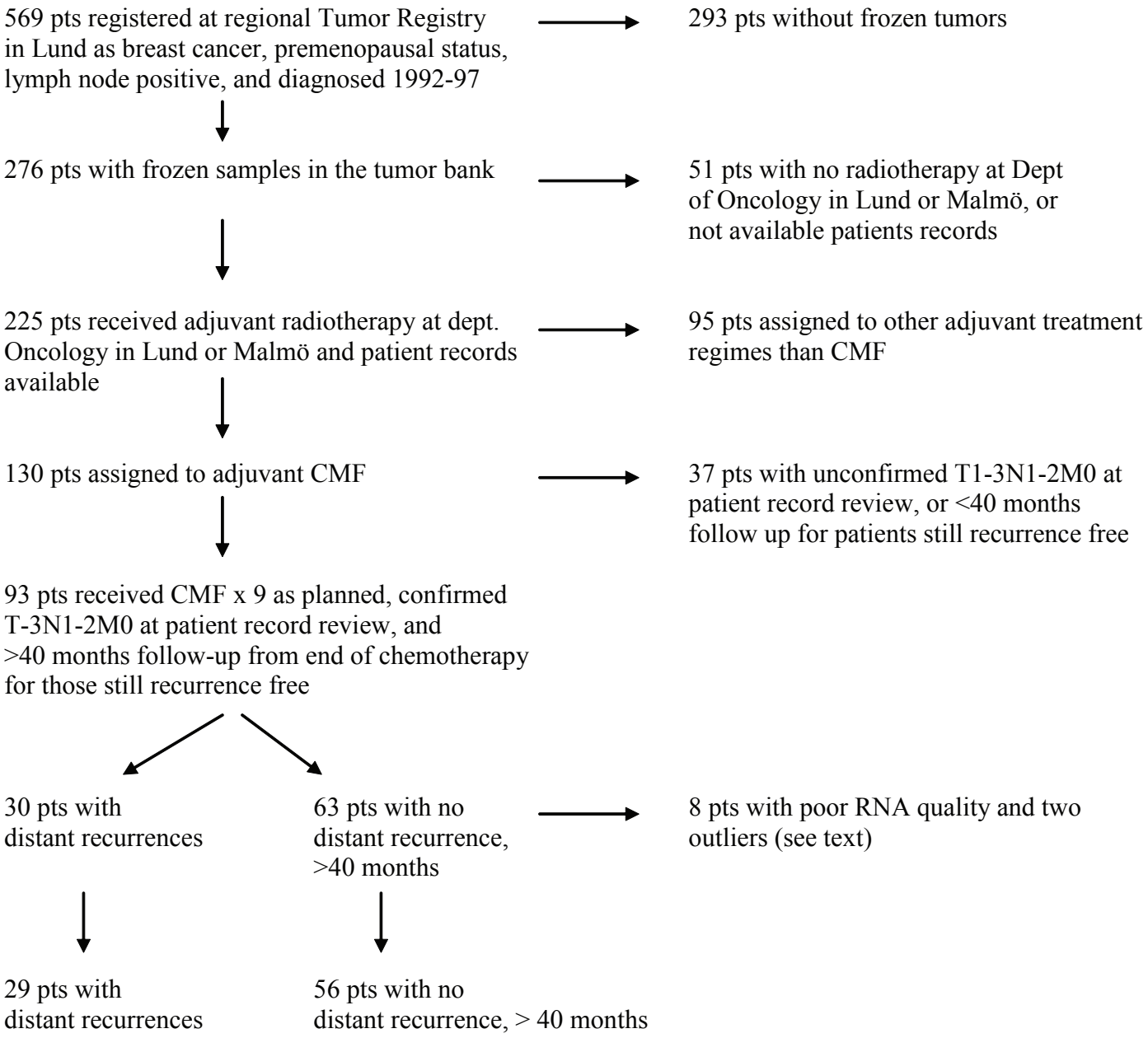
40. Zhao SC, Banerjee D, Mineishi S, et al. Post-transplant methotrexate administration leads to improved curability of mice bearing a mammary tumor transplanted with marrow transduced with a mutant human dihydrofolate reductase cDNA. *Hum Gene Ther* 1997; **8**, 903-9.

41. Nakamura M, Tsuji N, Asanuma K, et al. Survivin as a predictor of cis-diamminedichloroplatinum sensitivity in gastric cancer patients. *Cancer Sci* 2004; **95**, 44-51.

42. Zhao H, Langerod A, Ji Y, et al. Different gene expression patterns in invasive lobular and ductal carcinomas of the breast. *Mol Biol Cell* 2004; **15**, 2523-36.

43. Breslin T, Krogh M, Peterson C, et al. Signal transduction pathway profiling of individual tumor samples. *BMC Bioinformatics* 2005; **6**, 163.

Figure 1: Nimeus et al



Supplement 1. Genes known from the literature to be of importance in drug resistance. If no official gene symbol was found the accession number was used.

Gene symbol	Gene aliases	Gene name	Protein	Ref.
RBM17	SPF45			(1)
ABCB1	MDR1		P-glycoprotein, P-gp1	(2-6)
ABCG2	BCRP			(4)
ABCC1	MRP1			(2, 4, 5, 7)
ABCC2	MRP2			(2, 8, 9)
ABCC3	MRP3			(7)
ABCC5	MRP5			(10)
SLC19A1	RCF1			(4)
DHFR	DHFR	dihydrofolate reductase		(3, 11)
ERBB2	ERBB2			(3, 12-15)
MVP	LRP		lung resistant protein	(2)
BAD	bcl-2			(3, 16, 17)
BCL2L1	bcl-xL			(3, 16, 17)
GSTP1	GSTPi			(3, 18)
GSTM1	GSTM1			(18)
GSTT1	GST T1			(18)
SAT	SSAT, SAT			(19)
ANXA2	annexin II			(19)
TMSB10	TMSB10	thymosin beta-10		(19)
HSPE1	HSPE1	chaperonin-10		(19)
FXYD3	MAT-8	MAT-8 protein		(19)
BRAF, RAF1		Raf		(19)
KRAS2		K-ras		(19)

SLA	SLAP, SLA			(19)
PIK3CG		phosphoinositide 3-kinase		(19)
COPS8		HCOP9		(19)
APG5L		apoptosis specific protein		(19)
TNFRSF6		APO-1 cell surface antigen	fas	(17, 19)
CFLAR		FLIP protein		(19)
CCNG2		cyclin G		(19)
CDC2	CDC2			(19)
CDK2	CDK-2	cyclin-dependent protein kinase-2		(19)
MYLK2		myosin light chain		(19)
GSN	GSN	Gelsolin		(19)
TMSB4X, TMSB4Y		thymosin beta-4		(19)
SRM	SRM	spermidine synthase		(19, 20)
LOC442230, LOC440697, LOC391799		spermidine aminopropyltransferase		(19)
FGFR2	FGFR2	FGF receptor 2		(19)
TM4SF		transmembrane 4 superfamily protein		(19)
AUH	AUH	enoyl-CoA hydratase		(19)
NNT	NNT	nicotinamide nucleotide transhydrogenase		(19)
MRPS28	MRPS28	ribosomal protein S28		(19)
MRPL37	MRPL37	ribosomal protein L37		(19)
RPL7		ribosomal protein L7		(19)
MYC	myc			(3, 4, 21)
JUN	c-jun			(3, 12)
FOS	c-fos			(3)
HRAS, KRAS, NRAS	ras			(3)
FTP53	TP53		p53	(3, 13,

			19, 22)
CDKN1A	CDKN1A	p21cip1	(3, 15)
E2F1	E2F1		(23)
MDM2	mdm-2		(3, 24)
NOTCH4	NOTCH4		(20)
ITGA5	ITGA5		(20)
HMGA1	HMGA1		(20)
TCF8	TCF8, AREB6		(20)
SULT1E1	STE	Estrogen sulfotransferase	(20)
B2M	B2M	beta-2 micorglobulin precursor	(20)
XRCC1	XRCC1		(20)
CDC27	CDC27		(20)
GRN	GRN	granulin	(20, 25)
RCV1	RCV1		(20)
CRMP1	CRPM1	collapsin response mediator protein 1	(20)
PSMC1	PSMC1		(20)
TAF4		TBP-associated factor (hTAFII130)	(20)
CALM1, CALM2, CALM3		Calmodulin	(20)
TRIP		TRAF-interacting protein 1	(20)
CLTA		Clathrin light chain A	(20)
PIK3R1		PI-3-kinase associated p85	(20)
ARPC4		Arp2/3 protein complex subunit p20-Arc	(20)
CRK, CRKL		V-crk	(20)
PPIA, PPIB, PPIC, PPID, PPIE, PPIF, PPIG, PPIH		Peptidylprolyl cis-trans isomerase	(20)
MT1E		Metallothionein Ie gene (hMT-Ie)	(20)
NUDT1	hMTH1	human mutT homologue	(3)

MLH1	hMLH1	human mutLhomologue		(3)
BLMH	BLMH		bleomycin hydrolase	(3)
MGMT	MGMT	O6-metyl-guanine-DNA methyl-transferase		(3, 26)
TOP2A,TOP2B		DNA topoisomerase II	topoisomeras II	(3, 6)
ANXA1	ANXA1	Annexin1		(6)
BIRC5		Surviving gene		(27)
MKI67	MKI67		Ki-67	(28)
BAX	BAX	BCL2-associated X protein	bax	(17, 25, 29)
BAG1			bag-1	(17)
TNFSF6			fasL	(17)
ESR1	ER			(13)
MIB1			Mib1	(13)
TYMS	TYMS	thymidilate synthase	thymidilate synthase	(30)
UBE2M	UBE2M	ubiquitin-conjugating enzyme E2M (UBC12 homologue, yeast)		(25)
CUL1	CUL1	cullin 1		(25)
CSNK2B	CSNK2B	casein kinase 2, beta polypeptide		(25)
DDB1	DDB1			(25)
ABL1	ABL1			(25)
PRKDC	PRKDC	protein kinase, DNA-activated, catalytic polypeptide		(25)
ALDH1A1,ALDH3A1		aldehyde dehydrogenase		(31)
SH3KBP1	SH3KBP1	SH3-domain kinase binding protein 1	C-Cbl-interacting	(32)
HIST1H4E	HIST1H4E	Histone 1, H4e		(32)
EPB41L2	EPB41L2	erythrocyte membrane protein band 4.1-like 2		(32)
DPYSL3			Dihydropyrimidinase related protein-3	(32)
CDKN1C	CDKN1C	cyclin-dependent kinase inhibitor 1C (p57, Kip2)		(32)

EZH1	EZH1	enhancer of zeste homolog 1 (Drosophila)		(32)
GPS2, LOC392281	GPS2	G protein pathway suppressor 2		(32)
OS-9		OS-9 precursor		(32)
MXD4	MXD4	MAX dimerization protein 4		(32)
HLA-F		major histocompatibility complex, class I, F		(32)
A0366, A29103		TNF-binding polypeptide		(32)
CTTN	EMS1	ems1 sequence (mammary tumor and squamous cell carcinoma-associated (p80/85 src substrate)		(32)
RRAGA	RRAGA	Ras-related GTP binding A	FIP-1	(32)
RNPS1			human (clone E5.1) RNA binding protein	(32)
SHOX2			homeodomain protein (OG12)	(32)
GPC1			heparan sulfate proteoglycan (glypican)	(32)
MSH6	MSH6		hMSH6 protein	(32)
ODC1	ODC1	ornithine decarboxylase 1		(32)
AIP1	AIP1	atrophin-1 interacting protein 1		(32)
PCCA			propionyl coenzyme A carboxylase, beta polypeptide	(32)
GSTA3	GSTA3	glutathione S-transferase A3		(32)
SDHD		cytochrome b small subunit of complex		(32)
ATP5G3		mitochondrial ATP synthase subunit 9, P3 gene		(32)

		copy		
USP13	USP13	ubiquitin specific protease 13 (isopeptidase T-3)		(32)
MTX2	MTX2	metaxin 2		(32)
NTS		human proneurotensin/proneuromedin N		(32)
DFFA		DNA fragmentation factor-45		(32)
MAT2A		S-adenosylmethionine synthetase		(32)
OCRL	OCRL	oculocerebrorenal syndrome of Lowe		(32)
USP2	USP2	ubiquitin specific protease 2		(32)
A0043, AF028840		Kruppel-associated box protein		(32)
ERCC3	ERCC3	DNA repair helicase		(32)
MAGEA5	MAGEA5	melanoma antigen, family A, 5		(32)
AKR1C3	AKR1C3	aldo-keto reductase family 1, member C3		(32)
HNRPA2B1			hnRNP A2 protein	(32)
Hs.109059		mrp17		(32)
B0647, Hs.116412	EST			(32)
B4413, Hs.15871	EST			(32)
B8795, Hs. 72444	EST			(32)
B0557, Hs.115880	EST			(32)
B7553, Hs.12866	EST			(32)
A8884, Hs.12151	EST			(32)
B4895, Hs.3452	EST			(32)
B4272, Hs.8839	EST			(32)
A4146, AA586974	EST			(32)
B2430, Hs.118966	EST			(32)
C0670, Hs.31655	EST			(32)
A8952, Hs.109253	EST			(32)
B0829, Hs.107884	EST			(32)

A6528, Hs.8215	EST			(32)
A6811, Hs.100734	EST			(32)
B2316, Hs.117381	EST			(32)
RRM2	RRM2	ribonucleotide reductase	ribonucleotide reductase subunit 2	(33-35)
DCK			deoxycytidine kinase, dCK	(36-38)
NT5C2			cytoplasmic 5´ nucleotidase, 5NT	(39, 40)
DCTD	DCTD	dCMP deaminase		(41)
CDA			cytidine deaminase, CD	(42, 43)
SLC29A1	hENT1	human equilinbrate nucleoside transporter		(44)
SLC28A1	hCNT1	human concentrative nucleoside transporter		(45)
ERCC1	ERCC1	excision repair cross-complementing1		(46)
DPYD	DPO, DPYD	dihydropyrimidine dehydrogenase		(47)
PGR	PgR	Progesterone receptor		(48)
UGCG		UDP-glucose ceramide glucosyltransferase	glucosylceramide synthase GCS	(49)
PRKR		protein kinase, interferon-inducible double stranded RNA dependent		(25)
EIF4A1		eukaryotic translation initiation factor 4A, isoform 1		(25)
RAP1GDS1		RAP1, GTP-GDP dissociation stimulator 1		(25)
CALR		Calreticulin		(25)
CTNNA1		catenin (cadherin-associated protein), alpha 1 (102kD)		(25)
FCGRT		Fc fragment of IgG, receptor, transporter, alpha		(25)

Supplement 1. Genes known from the literature to be of importance in drug resistance. If no official gene symbol was found the accession number was used.

Gene symbol	Gene aliases	Gene name	Protein	Ref.
RBM17	SPF45			(1)
ABCB1	MDR1		P-glycoprotein, P-gp1	(2-6)
ABCG2	BCRP			(4)
ABCC1	MRP1			(2, 4, 5, 7)
ABCC2	MRP2			(2, 8, 9)
ABCC3	MRP3			(7)
ABCC5	MRP5			(10)
SLC19A1	RCF1			(4)
DHFR	DHFR	dihydrofolate reductase		(3, 11)
ERBB2	ERBB2			(3, 12-15)
MVP	LRP		lung resistant protein	(2)
BAD	bcl-2			(3, 16, 17)
BCL2L1	bcl-xL			(3, 16, 17)
GSTP1	GSTPi			(3, 18)
GSTM1	GSTM1			(18)
GSTT1	GST T1			(18)
SAT	SSAT, SAT			(19)
ANXA2	annexin II			(19)
TMSB10	TMSB10	thymosin beta-10		(19)
HSPE1	HSPE1	chaperonin-10		(19)
FXYD3	MAT-8	MAT-8 protein		(19)
BRAF, RAF1		Raf		(19)
KRAS2		K-ras		(19)

SLA	SLAP, SLA			(19)
PIK3CG		phosphoinositide 3-kinase		(19)
COPS8		HCOP9		(19)
APG5L		apoptosis specific protein		(19)
TNFRSF6		APO-1 cell surface antigen	fas	(17, 19)
CFLAR		FLIP protein		(19)
CCNG2		cyclin G		(19)
CDC2	CDC2			(19)
CDK2	CDK-2	cyclin-dependent protein kinase-2		(19)
MYLK2		myosin light chain		(19)
GSN	GSN	Gelsolin		(19)
TMSB4X, TMSB4Y		thymosin beta-4		(19)
SRM	SRM	spermidine synthase		(19, 20)
LOC442230, LOC440697, LOC391799		spermidine aminopropyltransferase		(19)
FGFR2	FGFR2	FGF receptor 2		(19)
TM4SF		transmembrane 4 superfamily protein		(19)
AUH	AUH	enoyl-CoA hydratase		(19)
NNT	NNT	nicotinamide nucleotide transhydrogenase		(19)
MRPS28	MRPS28	ribosomal protein S28		(19)
MRPL37	MRPL37	ribosomal protein L37		(19)
RPL7		ribosomal protein L7		(19)
MYC	myc			(3, 4, 21)
JUN	c-jun			(3, 12)
FOS	c-fos			(3)
HRAS, KRAS, NRAS	ras			(3)
FTP53	TP53		p53	(3, 13,

			19, 22)
CDKN1A	CDKN1A	p21cip1	(3, 15)
E2F1	E2F1		(23)
MDM2	mdm-2		(3, 24)
NOTCH4	NOTCH4		(20)
ITGA5	ITGA5		(20)
HMGA1	HMGA1		(20)
TCF8	TCF8, AREB6		(20)
SULT1E1	STE	Estrogen sulfotransferase	(20)
B2M	B2M	beta-2 micorglobulin precursor	(20)
XRCC1	XRCC1		(20)
CDC27	CDC27		(20)
GRN	GRN	granulin	(20, 25)
RCV1	RCV1		(20)
CRMP1	CRPM1	collapsin response mediator protein 1	(20)
PSMC1	PSMC1		(20)
TAF4		TBP-associated factor (hTAFII130)	(20)
CALM1, CALM2, CALM3		Calmodulin	(20)
TRIP		TRAF-interacting protein 1	(20)
CLTA		Clathrin light chain A	(20)
PIK3R1		PI-3-kinase associated p85	(20)
ARPC4		Arp2/3 protein complex subunit p20-Arc	(20)
CRK, CRKL		V-crk	(20)
PPIA, PPIB, PPIC, PPID, PPIE, PPIF, PPIG, PPIH		Peptidylprolyl cis-trans isomerase	(20)
MT1E		Metallothionein Ie gene (hMT-Ie)	(20)
NUDT1	hMTH1	human mutT homologue	(3)

MLH1	hMLH1	human mutLhomologue		(3)
BLMH	BLMH		bleomycin hydrolase	(3)
MGMT	MGMT	O6-metyl-guanine-DNA methyl-transferase		(3, 26)
TOP2A,TOP2B		DNA topoisomerase II	topoisomeras II	(3, 6)
ANXA1	ANXA1	Annexin1		(6)
BIRC5		Surviving gene		(27)
MKI67	MKI67		Ki-67	(28)
BAX	BAX	BCL2-associated X protein	bax	(17, 25, 29)
BAG1			bag-1	(17)
TNFSF6			fasL	(17)
ESR1	ER			(13)
MIB1			Mib1	(13)
TYMS	TYMS	thymidilate synthase	thymidilate synthase	(30)
UBE2M	UBE2M	ubiquitin-conjugating enzyme E2M (UBC12 homologue, yeast)		(25)
CUL1	CUL1	cullin 1		(25)
CSNK2B	CSNK2B	casein kinase 2, beta polypeptide		(25)
DDB1	DDB1			(25)
ABL1	ABL1			(25)
PRKDC	PRKDC	protein kinase, DNA-activated, catalytic polypeptide		(25)
ALDH1A1,ALDH3A1		aldehyde dehydrogenase		(31)
SH3KBP1	SH3KBP1	SH3-domain kinase binding protein 1	C-Cbl-interacting	(32)
HIST1H4E	HIST1H4E	Histone 1, H4e		(32)
EPB41L2	EPB41L2	erythrocyte membrane protein band 4.1-like 2		(32)
DPYSL3			Dihydropyrimidinase related protein-3	(32)
CDKN1C	CDKN1C	cyclin-dependent kinase inhibitor 1C (p57, Kip2)		(32)

EZH1	EZH1	enhancer of zeste homolog 1 (Drosophila)		(32)
GPS2, LOC392281	GPS2	G protein pathway suppressor 2		(32)
OS-9		OS-9 precursor		(32)
MXD4	MXD4	MAX dimerization protein 4		(32)
HLA-F		major histocompatibility complex, class I, F		(32)
A0366, A29103		TNF-binding polypeptide		(32)
CTTN	EMS1	ems1 sequence (mammary tumor and squamous cell carcinoma-associated (p80/85 src substrate)		(32)
RRAGA	RRAGA	Ras-related GTP binding A	FIP-1	(32)
RNPS1			human (clone E5.1) RNA binding protein	(32)
SHOX2			homeodomain protein (OG12)	(32)
GPC1			heparan sulfate proteoglycan (glypican)	(32)
MSH6	MSH6		hMSH6 protein	(32)
ODC1	ODC1	ornithine decarboxylase 1		(32)
AIP1	AIP1	atrophin-1 interacting protein 1		(32)
PCCA			propionyl coenzyme A carboxylase, beta polypeptide	(32)
GSTA3	GSTA3	glutathione S-transferase A3		(32)
SDHD		cytochrome b small subunit of complex		(32)
ATP5G3		mitochondrial ATP synthase subunit 9, P3 gene		(32)

		copy		
USP13	USP13	ubiquitin specific protease 13 (isopeptidase T-3)		(32)
MTX2	MTX2	metaxin 2		(32)
NTS		human proneurotensin/proneuromedin N		(32)
DFFA		DNA fragmentation factor-45		(32)
MAT2A		S-adenosylmethionine synthetase		(32)
OCRL	OCRL	oculocerebrorenal syndrome of Lowe		(32)
USP2	USP2	ubiquitin specific protease 2		(32)
A0043, AF028840		Kruppel-associated box protein		(32)
ERCC3	ERCC3	DNA repair helicase		(32)
MAGEA5	MAGEA5	melanoma antigen, family A, 5		(32)
AKR1C3	AKR1C3	aldo-keto reductase family 1, member C3		(32)
HNRPA2B1			hnRNP A2 protein	(32)
Hs.109059		mrp17		(32)
B0647, Hs.116412	EST			(32)
B4413, Hs.15871	EST			(32)
B8795, Hs. 72444	EST			(32)
B0557, Hs.115880	EST			(32)
B7553, Hs.12866	EST			(32)
A8884, Hs.12151	EST			(32)
B4895, Hs.3452	EST			(32)
B4272, Hs.8839	EST			(32)
A4146, AA586974	EST			(32)
B2430, Hs.118966	EST			(32)
C0670, Hs.31655	EST			(32)
A8952, Hs.109253	EST			(32)
B0829, Hs.107884	EST			(32)

A6528, Hs.8215	EST			(32)
A6811, Hs.100734	EST			(32)
B2316, Hs.117381	EST			(32)
RRM2	RRM2	ribonucleotide reductase	ribonucleotide reductase subunit 2	(33-35)
DCK			deoxycytidine kinase, dCK	(36-38)
NT5C2			cytoplasmic 5´ nucleotidase, 5NT	(39, 40)
DCTD	DCTD	dCMP deaminase		(41)
CDA			cytidine deaminase, CD	(42, 43)
SLC29A1	hENT1	human equilinbrate nucleoside transporter		(44)
SLC28A1	hCNT1	human concentrative nucleoside transporter		(45)
ERCC1	ERCC1	excision repair cross-complementing1		(46)
DPYD	DPO, DPYD	dihydropyrimidine dehydrogenase		(47)
PGR	PgR	Progesterone receptor		(48)
UGCG		UDP-glucose ceramide glucosyltransferase	glucosylceramide synthase GCS	(49)
PRKR		protein kinase, interferon-inducible double stranded RNA dependent		(25)
EIF4A1		eukaryotic translation initiation factor 4A, isoform 1		(25)
RAP1GDS1		RAP1, GTP-GDP dissociation stimulator 1		(25)
CALR		Calreticulin		(25)
CTNNA1		catenin (cadherin-associated protein), alpha 1 (102kD)		(25)
FCGRT		Fc fragment of IgG, receptor, transporter, alpha		(25)

GAMT	guanidinoacetate N-methyltransferase	(25)
KIAA0138	KIAA0138 gene product	(25)
AK3	adenylate kinase 3	(25)
CLTB	clathrin, light polypeptide (Lcb)	(25)
FLII	flightless I homologue (Drosophila)	(25)
MRPS12	mitochondrial ribosomal protein S12	(25)
FLJ21168	hypothetical protein FLJ21168	(25)
RAB31	RAB31, member RAS oncogene family	(25)
DDX19	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 19 (DBP5 homologue, yeast)	(25)
GPX4	glutathione peroxidase 4 (phospholipids hydroperoxidase)	(25)
RBPMS	RNA-binding protein gene with multiple splicing	(25)
PCBP1	poly(rC) binding protein 1	(25)
EP400	trinucleotide repeat containing 12	(25)
CGI-51	CGI-51 protein	(25)
SGSH	N-sulfoglucosamine sulfohydrolase (sulfamidase)	(25)

CHS1	Chediak-Higashi syndrome 1	(25)
ACTR2	ARP2 actin-related protein 2 homologue (yeast)	(25)
CYBA	CYBA cytochrome b-245, alpha polypeptide	(25)
SDC4	syndecan 4 (amphiglycan, ryudocan)	(25)
RALY	RNA-binding protein (autoantigenic)	(25)
BRD2	bromodomain-containing 2	(25)
HSD17B4	hydroxysteroid (17-beta) dehydrogenase 4	(25)
LOXL1	lysyl oxidase-like 1	(25)
LSM7	U6 snRNA-associated Sm-like protein LSm7	(25)
PRIM2A	primase, polypeptide 2A (58kD)	(25)
SFRS4	splicing factor, arginine/serine-rich 4	(25)
GTF2H2	general transcription factor IIH, polypeptide 2 (44kD subunit)	(25)
ALAS1	aminolevulinate, delta-, synthase 1	(25)
ZNF288	zinc finger protein 288	(25)
STXBP2	syntaxin binding protein 2	(25)

Supplement 1. Genes known from the literature to be of importance in drug resistance. If no official gene symbol was found the accession number was used.

Gene symbol	Gene aliases	Gene name	Protein	Ref.
RBM17	SPF45			(1)
ABCB1	MDR1		P-glycoprotein, P-gp1	(2-6)
ABCG2	BCRP			(4)
ABCC1	MRP1			(2, 4, 5, 7)
ABCC2	MRP2			(2, 8, 9)
ABCC3	MRP3			(7)
ABCC5	MRP5			(10)
SLC19A1	RCF1			(4)
DHFR	DHFR	dihydrofolate reductase		(3, 11)
ERBB2	ERBB2			(3, 12-15)
MVP	LRP		lung resistant protein	(2)
BAD	bcl-2			(3, 16, 17)
BCL2L1	bcl-xL			(3, 16, 17)
GSTP1	GSTPi			(3, 18)
GSTM1	GSTM1			(18)
GSTT1	GST T1			(18)
SAT	SSAT, SAT			(19)
ANXA2	annexin II			(19)
TMSB10	TMSB10	thymosin beta-10		(19)
HSPE1	HSPE1	chaperonin-10		(19)
FXYD3	MAT-8	MAT-8 protein		(19)
BRAF, RAF1		Raf		(19)
KRAS2		K-ras		(19)

SLA	SLAP, SLA			(19)
PIK3CG		phosphoinositide 3-kinase		(19)
COPS8		HCOP9		(19)
APG5L		apoptosis specific protein		(19)
TNFRSF6		APO-1 cell surface antigen	fas	(17, 19)
CFLAR		FLIP protein		(19)
CCNG2		cyclin G		(19)
CDC2	CDC2			(19)
CDK2	CDK-2	cyclin-dependent protein kinase-2		(19)
MYLK2		myosin light chain		(19)
GSN	GSN	Gelsolin		(19)
TMSB4X, TMSB4Y		thymosin beta-4		(19)
SRM	SRM	spermidine synthase		(19, 20)
LOC442230, LOC440697, LOC391799		spermidine aminopropyltransferase		(19)
FGFR2	FGFR2	FGF receptor 2		(19)
TM4SF		transmembrane 4 superfamily protein		(19)
AUH	AUH	enoyl-CoA hydratase		(19)
NNT	NNT	nicotinamide nucleotide transhydrogenase		(19)
MRPS28	MRPS28	ribosomal protein S28		(19)
MRPL37	MRPL37	ribosomal protein L37		(19)
RPL7		ribosomal protein L7		(19)
MYC	myc			(3, 4, 21)
JUN	c-jun			(3, 12)
FOS	c-fos			(3)
HRAS, KRAS, NRAS	ras			(3)
FTP53	TP53		p53	(3, 13,

			19, 22)
CDKN1A	CDKN1A	p21cip1	(3, 15)
E2F1	E2F1		(23)
MDM2	mdm-2		(3, 24)
NOTCH4	NOTCH4		(20)
ITGA5	ITGA5		(20)
HMGA1	HMGA1		(20)
TCF8	TCF8, AREB6		(20)
SULT1E1	STE	Estrogen sulfotransferase	(20)
B2M	B2M	beta-2 micorglobulin precursor	(20)
XRCC1	XRCC1		(20)
CDC27	CDC27		(20)
GRN	GRN	granulin	(20, 25)
RCV1	RCV1		(20)
CRMP1	CRPM1	collapsin response mediator protein 1	(20)
PSMC1	PSMC1		(20)
TAF4		TBP-associated factor (hTAFII130)	(20)
CALM1, CALM2, CALM3		Calmodulin	(20)
TRIP		TRAF-interacting protein 1	(20)
CLTA		Clathrin light chain A	(20)
PIK3R1		PI-3-kinase associated p85	(20)
ARPC4		Arp2/3 protein complex subunit p20-Arc	(20)
CRK, CRKL		V-crk	(20)
PPIA, PPIB, PPIC, PPID, PPIE, PPIF, PPIG, PPIH		Peptidylprolyl cis-trans isomerase	(20)
MT1E		Metallothionein Ie gene (hMT-Ie)	(20)
NUDT1	hMTH1	human mutT homologue	(3)

MLH1	hMLH1	human mutLhomologue		(3)
BLMH	BLMH		bleomycin hydrolase	(3)
MGMT	MGMT	O6-metyl-guanine-DNA methyl-transferase		(3, 26)
TOP2A,TOP2B		DNA topoisomerase II	topoisomeras II	(3, 6)
ANXA1	ANXA1	Annexin1		(6)
BIRC5		Surviving gene		(27)
MKI67	MKI67		Ki-67	(28)
BAX	BAX	BCL2-associated X protein	bax	(17, 25, 29)
BAG1			bag-1	(17)
TNFSF6			fasL	(17)
ESR1	ER			(13)
MIB1			Mib1	(13)
TYMS	TYMS	thymidilate synthase	thymidilate synthase	(30)
UBE2M	UBE2M	ubiquitin-conjugating enzyme E2M (UBC12 homologue, yeast)		(25)
CUL1	CUL1	cullin 1		(25)
CSNK2B	CSNK2B	casein kinase 2, beta polypeptide		(25)
DDB1	DDB1			(25)
ABL1	ABL1			(25)
PRKDC	PRKDC	protein kinase, DNA-activated, catalytic polypeptide		(25)
ALDH1A1,ALDH3A1		aldehyde dehydrogenase		(31)
SH3KBP1	SH3KBP1	SH3-domain kinase binding protein 1	C-Cbl-interacting	(32)
HIST1H4E	HIST1H4E	Histone 1, H4e		(32)
EPB41L2	EPB41L2	erythrocyte membrane protein band 4.1-like 2		(32)
DPYSL3			Dihydropyrimidinase related protein-3	(32)
CDKN1C	CDKN1C	cyclin-dependent kinase inhibitor 1C (p57, Kip2)		(32)

EZH1	EZH1	enhancer of zeste homolog 1 (Drosophila)		(32)
GPS2, LOC392281	GPS2	G protein pathway suppressor 2		(32)
OS-9		OS-9 precursor		(32)
MXD4	MXD4	MAX dimerization protein 4		(32)
HLA-F		major histocompatibility complex, class I, F		(32)
A0366, A29103		TNF-binding polypeptide		(32)
CTTN	EMS1	ems1 sequence (mammary tumor and squamous cell carcinoma-associated (p80/85 src substrate)		(32)
RRAGA	RRAGA	Ras-related GTP binding A	FIP-1	(32)
RNPS1			human (clone E5.1) RNA binding protein	(32)
SHOX2			homeodomain protein (OG12)	(32)
GPC1			heparan sulfate proteoglycan (glypican)	(32)
MSH6	MSH6		hMSH6 protein	(32)
ODC1	ODC1	ornithine decarboxylase 1		(32)
AIP1	AIP1	atrophin-1 interacting protein 1		(32)
PCCA			propionyl coenzyme A carboxylase, beta polypeptide	(32)
GSTA3	GSTA3	glutathione S-transferase A3		(32)
SDHD		cytochrome b small subunit of complex		(32)
ATP5G3		mitochondrial ATP synthase subunit 9, P3 gene		(32)

		copy		
USP13	USP13	ubiquitin specific protease 13 (isopeptidase T-3)		(32)
MTX2	MTX2	metaxin 2		(32)
NTS		human proneurotensin/proneuromedin N		(32)
DFFA		DNA fragmentation factor-45		(32)
MAT2A		S-adenosylmethionine synthetase		(32)
OCRL	OCRL	oculocerebrorenal syndrome of Lowe		(32)
USP2	USP2	ubiquitin specific protease 2		(32)
A0043, AF028840		Kruppel-associated box protein		(32)
ERCC3	ERCC3	DNA repair helicase		(32)
MAGEA5	MAGEA5	melanoma antigen, family A, 5		(32)
AKR1C3	AKR1C3	aldo-keto reductase family 1, member C3		(32)
HNRPA2B1			hnRNP A2 protein	(32)
Hs.109059		mrp17		(32)
B0647, Hs.116412	EST			(32)
B4413, Hs.15871	EST			(32)
B8795, Hs. 72444	EST			(32)
B0557, Hs.115880	EST			(32)
B7553, Hs.12866	EST			(32)
A8884, Hs.12151	EST			(32)
B4895, Hs.3452	EST			(32)
B4272, Hs.8839	EST			(32)
A4146, AA586974	EST			(32)
B2430, Hs.118966	EST			(32)
C0670, Hs.31655	EST			(32)
A8952, Hs.109253	EST			(32)
B0829, Hs.107884	EST			(32)

A6528, Hs.8215	EST			(32)
A6811, Hs.100734	EST			(32)
B2316, Hs.117381	EST			(32)
RRM2	RRM2	ribonucleotide reductase	ribonucleotide reductase subunit 2	(33-35)
DCK			deoxycytidine kinase, dCK	(36-38)
NT5C2			cytoplasmic 5´ nucleotidase, 5NT	(39, 40)
DCTD	DCTD	dCMP deaminase		(41)
CDA			cytidine deaminase, CD	(42, 43)
SLC29A1	hENT1	human equilinbrate nucleoside transporter		(44)
SLC28A1	hCNT1	human concentrative nucleoside transporter		(45)
ERCC1	ERCC1	excision repair cross-complementing1		(46)
DPYD	DPO, DPYD	dihydropyrimidine dehydrogenase		(47)
PGR	PgR	Progesterone receptor		(48)
UGCG		UDP-glucose ceramide glucosyltransferase	glucosylceramide synthase GCS	(49)
PRKR		protein kinase, interferon-inducible double stranded RNA dependent		(25)
EIF4A1		eukaryotic translation initiation factor 4A, isoform 1		(25)
RAP1GDS1		RAP1, GTP-GDP dissociation stimulator 1		(25)
CALR		Calreticulin		(25)
CTNNA1		catenin (cadherin-associated protein), alpha 1 (102kD)		(25)
FCGRT		Fc fragment of IgG, receptor, transporter, alpha		(25)

Supplement 1. Genes known from the literature to be of importance in drug resistance. If no official gene symbol was found the accession number was used.

Gene symbol	Gene aliases	Gene name	Protein	Ref.
RBM17	SPF45			(1)
ABCB1	MDR1		P-glycoprotein, P-gp1	(2-6)
ABCG2	BCRP			(4)
ABCC1	MRP1			(2, 4, 5, 7)
ABCC2	MRP2			(2, 8, 9)
ABCC3	MRP3			(7)
ABCC5	MRP5			(10)
SLC19A1	RCF1			(4)
DHFR	DHFR	dihydrofolate reductase		(3, 11)
ERBB2	ERBB2			(3, 12-15)
MVP	LRP		lung resistant protein	(2)
BAD	bcl-2			(3, 16, 17)
BCL2L1	bcl-xL			(3, 16, 17)
GSTP1	GSTPi			(3, 18)
GSTM1	GSTM1			(18)
GSTT1	GST T1			(18)
SAT	SSAT, SAT			(19)
ANXA2	annexin II			(19)
TMSB10	TMSB10	thymosin beta-10		(19)
HSPE1	HSPE1	chaperonin-10		(19)
FXYD3	MAT-8	MAT-8 protein		(19)
BRAF, RAF1		Raf		(19)
KRAS2		K-ras		(19)

SLA	SLAP, SLA			(19)
PIK3CG		phosphoinositide 3-kinase		(19)
COPS8		HCOP9		(19)
APG5L		apoptosis specific protein		(19)
TNFRSF6		APO-1 cell surface antigen	fas	(17, 19)
CFLAR		FLIP protein		(19)
CCNG2		cyclin G		(19)
CDC2	CDC2			(19)
CDK2	CDK-2	cyclin-dependent protein kinase-2		(19)
MYLK2		myosin light chain		(19)
GSN	GSN	Gelsolin		(19)
TMSB4X, TMSB4Y		thymosin beta-4		(19)
SRM	SRM	spermidine synthase		(19, 20)
LOC442230, LOC440697, LOC391799		spermidine aminopropyltransferase		(19)
FGFR2	FGFR2	FGF receptor 2		(19)
TM4SF		transmembrane 4 superfamily protein		(19)
AUH	AUH	enoyl-CoA hydratase		(19)
NNT	NNT	nicotinamide nucleotide transhydrogenase		(19)
MRPS28	MRPS28	ribosomal protein S28		(19)
MRPL37	MRPL37	ribosomal protein L37		(19)
RPL7		ribosomal protein L7		(19)
MYC	myc			(3, 4, 21)
JUN	c-jun			(3, 12)
FOS	c-fos			(3)
HRAS, KRAS, NRAS	ras			(3)
FTP53	TP53		p53	(3, 13,

			19, 22)
CDKN1A	CDKN1A	p21cip1	(3, 15)
E2F1	E2F1		(23)
MDM2	mdm-2		(3, 24)
NOTCH4	NOTCH4		(20)
ITGA5	ITGA5		(20)
HMGA1	HMGA1		(20)
TCF8	TCF8, AREB6		(20)
SULT1E1	STE	Estrogen sulfotransferase	(20)
B2M	B2M	beta-2 micorglobulin precursor	(20)
XRCC1	XRCC1		(20)
CDC27	CDC27		(20)
GRN	GRN	granulin	(20, 25)
RCV1	RCV1		(20)
CRMP1	CRPM1	collapsin response mediator protein 1	(20)
PSMC1	PSMC1		(20)
TAF4		TBP-associated factor (hTAFII130)	(20)
CALM1, CALM2, CALM3		Calmodulin	(20)
TRIP		TRAF-interacting protein 1	(20)
CLTA		Clathrin light chain A	(20)
PIK3R1		PI-3-kinase associated p85	(20)
ARPC4		Arp2/3 protein complex subunit p20-Arc	(20)
CRK, CRKL		V-crk	(20)
PPIA, PPIB, PPIC, PPID, PPIE, PPIF, PPIG, PPIH		Peptidylprolyl cis-trans isomerase	(20)
MT1E		Metallothionein Ie gene (hMT-Ie)	(20)
NUDT1	hMTH1	human mutT homologue	(3)

MLH1	hMLH1	human mutLhomologue		(3)
BLMH	BLMH		bleomycin hydrolase	(3)
MGMT	MGMT	O6-metyl-guanine-DNA methyl-transferase		(3, 26)
TOP2A,TOP2B		DNA topoisomerase II	topoisomeras II	(3, 6)
ANXA1	ANXA1	Annexin1		(6)
BIRC5		Surviving gene		(27)
MKI67	MKI67		Ki-67	(28)
BAX	BAX	BCL2-associated X protein	bax	(17, 25, 29)
BAG1			bag-1	(17)
TNFSF6			fasL	(17)
ESR1	ER			(13)
MIB1			Mib1	(13)
TYMS	TYMS	thymidilate synthase	thymidilate synthase	(30)
UBE2M	UBE2M	ubiquitin-conjugating enzyme E2M (UBC12 homologue, yeast)		(25)
CUL1	CUL1	cullin 1		(25)
CSNK2B	CSNK2B	casein kinase 2, beta polypeptide		(25)
DDB1	DDB1			(25)
ABL1	ABL1			(25)
PRKDC	PRKDC	protein kinase, DNA-activated, catalytic polypeptide		(25)
ALDH1A1,ALDH3A1		aldehyde dehydrogenase		(31)
SH3KBP1	SH3KBP1	SH3-domain kinase binding protein 1	C-Cbl-interacting	(32)
HIST1H4E	HIST1H4E	Histone 1, H4e		(32)
EPB41L2	EPB41L2	erythrocyte membrane protein band 4.1-like 2		(32)
DPYSL3			Dihydropyrimidinase related protein-3	(32)
CDKN1C	CDKN1C	cyclin-dependent kinase inhibitor 1C (p57, Kip2)		(32)

EZH1	EZH1	enhancer of zeste homolog 1 (Drosophila)		(32)
GPS2, LOC392281	GPS2	G protein pathway suppressor 2		(32)
OS-9		OS-9 precursor		(32)
MXD4	MXD4	MAX dimerization protein 4		(32)
HLA-F		major histocompatibility complex, class I, F		(32)
A0366, A29103		TNF-binding polypeptide		(32)
CTTN	EMS1	ems1 sequence (mammary tumor and squamous cell carcinoma-associated (p80/85 src substrate)		(32)
RRAGA	RRAGA	Ras-related GTP binding A	FIP-1	(32)
RNPS1			human (clone E5.1) RNA binding protein	(32)
SHOX2			homeodomain protein (OG12)	(32)
GPC1			heparan sulfate proteoglycan (glypican)	(32)
MSH6	MSH6		hMSH6 protein	(32)
ODC1	ODC1	ornithine decarboxylase 1		(32)
AIP1	AIP1	atrophin-1 interacting protein 1		(32)
PCCA			propionyl coenzyme A carboxylase, beta polypeptide	(32)
GSTA3	GSTA3	glutathione S-transferase A3		(32)
SDHD		cytochrome b small subunit of complex		(32)
ATP5G3		mitochondrial ATP synthase subunit 9, P3 gene		(32)

		copy		
USP13	USP13	ubiquitin specific protease 13 (isopeptidase T-3)		(32)
MTX2	MTX2	metaxin 2		(32)
NTS		human proneurotensin/proneuromedin N		(32)
DFFA		DNA fragmentation factor-45		(32)
MAT2A		S-adenosylmethionine synthetase		(32)
OCRL	OCRL	oculocerebrorenal syndrome of Lowe		(32)
USP2	USP2	ubiquitin specific protease 2		(32)
A0043, AF028840		Kruppel-associated box protein		(32)
ERCC3	ERCC3	DNA repair helicase		(32)
MAGEA5	MAGEA5	melanoma antigen, family A, 5		(32)
AKR1C3	AKR1C3	aldo-keto reductase family 1, member C3		(32)
HNRPA2B1			hnRNP A2 protein	(32)
Hs.109059		mrp17		(32)
B0647, Hs.116412	EST			(32)
B4413, Hs.15871	EST			(32)
B8795, Hs. 72444	EST			(32)
B0557, Hs.115880	EST			(32)
B7553, Hs.12866	EST			(32)
A8884, Hs.12151	EST			(32)
B4895, Hs.3452	EST			(32)
B4272, Hs.8839	EST			(32)
A4146, AA586974	EST			(32)
B2430, Hs.118966	EST			(32)
C0670, Hs.31655	EST			(32)
A8952, Hs.109253	EST			(32)
B0829, Hs.107884	EST			(32)

A6528, Hs.8215	EST			(32)
A6811, Hs.100734	EST			(32)
B2316, Hs.117381	EST			(32)
RRM2	RRM2	ribonucleotide reductase	ribonucleotide reductase subunit 2	(33-35)
DCK			deoxycytidine kinase, dCK	(36-38)
NT5C2			cytoplasmic 5´ nucleotidase, 5NT	(39, 40)
DCTD	DCTD	dCMP deaminase		(41)
CDA			cytidine deaminase, CD	(42, 43)
SLC29A1	hENT1	human equilinbrate nucleoside transporter		(44)
SLC28A1	hCNT1	human concentrative nucleoside transporter		(45)
ERCC1	ERCC1	excision repair cross-complementing1		(46)
DPYD	DPO, DPYD	dihydropyrimidine dehydrogenase		(47)
PGR	PgR	Progesterone receptor		(48)
UGCG		UDP-glucose ceramide glucosyltransferase	glucosylceramide synthase GCS	(49)
PRKR		protein kinase, interferon-inducible double stranded RNA dependent		(25)
EIF4A1		eukaryotic translation initiation factor 4A, isoform 1		(25)
RAP1GDS1		RAP1, GTP-GDP dissociation stimulator 1		(25)
CALR		Calreticulin		(25)
CTNNA1		catenin (cadherin-associated protein), alpha 1 (102kD)		(25)
FCGRT		Fc fragment of IgG, receptor, transporter, alpha		(25)

GAMT	guanidinoacetate N-methyltransferase	(25)
KIAA0138	KIAA0138 gene product	(25)
AK3	adenylate kinase 3	(25)
CLTB	clathrin, light polypeptide (Lcb)	(25)
FLII	flightless I homologue (Drosophila)	(25)
MRPS12	mitochondrial ribosomal protein S12	(25)
FLJ21168	hypothetical protein FLJ21168	(25)
RAB31	RAB31, member RAS oncogene family	(25)
DDX19	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 19 (DBP5 homologue, yeast)	(25)
GPX4	glutathione peroxidase 4 (phospholipids hydroperoxidase)	(25)
RBPMS	RNA-binding protein gene with multiple splicing	(25)
PCBP1	poly(rC) binding protein 1	(25)
EP400	trinucleotide repeat containing 12	(25)
CGI-51	CGI-51 protein	(25)
SGSH	N-sulfoglucosamine sulfohydrolase (sulfamidase)	(25)

CHS1	Chediak-Higashi syndrome 1	(25)
ACTR2	ARP2 actin-related protein 2 homologue (yeast)	(25)
CYBA	CYBA cytochrome b-245, alpha polypeptide	(25)
SDC4	syndecan 4 (amphiglycan, ryudocan)	(25)
RALY	RNA-binding protein (autoantigenic)	(25)
BRD2	bromodomain-containing 2	(25)
HSD17B4	hydroxysteroid (17-beta) dehydrogenase 4	(25)
LOXL1	lysyl oxidase-like 1	(25)
LSM7	U6 snRNA-associated Sm-like protein LSm7	(25)
PRIM2A	primase, polypeptide 2A (58kD)	(25)
SFRS4	splicing factor, arginine/serine-rich 4	(25)
GTF2H2	general transcription factor IIH, polypeptide 2 (44kD subunit)	(25)
ALAS1	aminolevulinate, delta-, synthase 1	(25)
ZNF288	zinc finger protein 288	(25)
STXBP2	syntaxin binding protein 2	(25)

Unnamed	Cluster Incl. U66042:Human clone 191B7 placenta expressed mRNA from chromosome X /cds=UNKNOWN /gb=U66042 /gi=15	(25)
CG1I	putative cyclin G1 interacting protein	(25)
LIMK2	LIM domain kinase 2	(25)
ATP6V0D1	ATPase, H+ transporting, lysosomal (vacuolar proton pump), member D	(25)
DICER1	helicase-moi	(25)
MUC1	mucin 1, transmembrane	(25)
DDOST	dolichyl-diphosphooligosaccharide-protein glycosyltransferase	(25)
EPO	Erythropoietin	(25)
Unnamed	Cluster Incl. M11119:Human endogenous retrovirus envelope region mRNA (PL1) /cds=UNKNOWN /gb=M11119 /gi=182205	(25)
AD024	AD024 protein	(25)
SLC25A1	solute carrier family 25 (mitochondrial carrier; citrate transporter), member 1	(25)
MGST3	microsomal glutathione S-transferase 3	(25)
RABAC1	Rab acceptor 1 (prenylated)	(25)
DRAP1	DR1-associated protein 1 (negative cofactor 2 alpha)	(25)
FUS	fusion, derived from t(12;16) malignant liposarcoma	(25)

Unnamed	Cluster Incl. U66042:Human clone 191B7 placenta expressed mRNA from chromosome X /cds=UNKNOWN /gb=U66042 /gi=15	(25)
CG1I	putative cyclin G1 interacting protein	(25)
LIMK2	LIM domain kinase 2	(25)
ATP6V0D1	ATPase, H+ transporting, lysosomal (vacuolar proton pump), member D	(25)
DICER1	helicase-moi	(25)
MUC1	mucin 1, transmembrane	(25)
DDOST	dolichyl-diphosphooligosaccharide-protein glycosyltransferase	(25)
EPO	Erythropoietin	(25)
Unnamed	Cluster Incl. M11119:Human endogenous retrovirus envelope region mRNA (PL1) /cds=UNKNOWN /gb=M11119 /gi=182205	(25)
AD024	AD024 protein	(25)
SLC25A1	solute carrier family 25 (mitochondrial carrier; citrate transporter), member 1	(25)
MGST3	microsomal glutathione S-transferase 3	(25)
RABAC1	Rab acceptor 1 (prenylated)	(25)
DRAP1	DR1-associated protein 1 (negative cofactor 2 alpha)	(25)
FUS	fusion, derived from t(12;16) malignant liposarcoma	(25)

EMP3	epithelial membrane protein 3	(25)
LOC56270	hypothetical protein 628	(25)
AP2S1	adaptor-related protein complex 2, sigma 1 subunit	(25)
DNAL4	dynein, axonemal, light polypeptide 4	(25)
TFAP4	transcription factor AP-4 (activating enhancer binding protein 4)	(25)
MRPL12	mitochondrial ribosomal protein L12	(25)
LIM	LIM protein (similar to rat protein kinase Cbinding enigma)	(25)
ALS2CR3	amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 3	(25)
ATP5A1	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, cardiac muscle	(25)
ZNF38	zinc finger protein 38 (KOX 25)	(25)
KPNB2	karyopherin (importin) beta 2	(25)
U5-100K	prp28, U5 snRNP 100 kd protein	(25)
SDBCAG84	serologically defined breast cancer antigen 84	(25)
IF2	translation initiation factor IF2	(25)
E2-EPF	ubiquitin carrier protein	(25)

EEF1A1		eukaryotic translation elongation factor 1 alpha 1	(25)
CTBP1		C-terminal binding protein 1	(25)
Cluster Incl. AI951946: x39f10.x1 Homo sapiens cDNA, 3 end /clone=IMAGE-2546059 /clone_end=3 /gb=AI951946 /gi	EST		(25)
CLPTM1		cleft lip and palate associated transmembrane protein 1	(25)
Cluster Incl. W72239:zd62h08.s1 Homo sapiens cDNA, 3 end /clone=IMAGE-345279 /clone_end=3 /gb=W72239 /gi=1382	EST		(25)
RER1		similar to S. cerevisiae RER1	(25)
SDF2		stromal cell-derived factor 2	(25)
Unnamed		Cluster Incl. AF007128:Homo sapiens clone 23870 mRNA sequence /cds=UNKNOWN /gb=AF007128 /gi=2852601 /ug=Hs.1246	(25)
VAPB		VAMP (vesicle-associated membrane protein)-associated protein B and C	(25)
FRAG1		FGF receptor activating protein 1	(25)

CRABP1	cellular retinoic acid binding protein 1	(25)
VWF	von Willebrand factor	(25)
CLK2	CDC-like kinase 2	(25)
P4HB	procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), beta polypeptide (protein disulfide	(25)
PPP2R1A	protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), alpha isoform	(25)

1. Sampath J, Long PR, Shepard RL, Xia X, Devanarayan V, Sandusky GE, et al. Human SPF45, a splicing factor, has limited expression in normal tissues, is overexpressed in many tumors, and can confer a multidrug-resistant phenotype to cells. Am J Pathol 2003;163(5):1781-90.
2. Burger H, Foekens JA, Look MP, Meijer-van Gelder ME, Klijn JG, Wiemer EA, et al. RNA expression of breast cancer resistance protein, lung resistance-related protein, multidrug resistance-associated proteins 1 and 2, and multidrug resistance gene 1 in breast cancer: correlation with chemotherapeutic response. Clin Cancer Res 2003;9(2):827-36.
3. el-Deiry WS. Role of oncogenes in resistance and killing by cancer therapeutic agents. Curr Opin Oncol 1997;9(1):79-87.
4. Fazeny-Dorner B, Piribauer M, Wenzel C, Fakhrai N, Pirker C, Berger W, et al. Cytogenetic and comparative genomic hybridization findings in four cases of breast cancer after neoadjuvant chemotherapy. Cancer Genet Cytogenet 2003;146(2):161-6.
5. Bredel M, Bredel C, Sikic BI. Genomics-based hypothesis generation: a novel approach to unravelling drug resistance in brain tumours? Lancet Oncol 2004;5(2):89-100.
6. Wang Y, Serfass L, Roy MO, Wong J, Bonneau AM, Georges E. Annexin-I expression modulates drug resistance in tumor cells. Biochem Biophys Res Commun 2004;314(2):565-70.
7. Young LC, Campling BG, Cole SP, Deeley RG, Gerlach JH. Multidrug resistance proteins MRP3, MRP1, and MRP2 in lung cancer: correlation of protein levels with drug response and messenger RNA levels. Clin Cancer Res 2001;7(6):1798-804.
8. Cui Y, Konig J, Buchholz JK, Spring H, Leier I, Keppler D. Drug resistance and ATP-dependent conjugate transport mediated by the apical multidrug resistance protein, MRP2, permanently expressed in human and canine cells. Mol Pharmacol 1999;55(5):929-37.

9. Kool M, de Haas M, Scheffer GL, Scheper RJ, van Eijk MJ, Juijn JA, et al. Analysis of expression of cMOAT (MRP2), MRP3, MRP4, and MRP5, homologues of the multidrug resistance-associated protein gene (MRP1), in human cancer cell lines. Cancer Res 1997;57(16):3537-47.
10. Davidson A, Dick G, Pritchard-Jones K, Pinkerton R. EVE/cyclosporin (etoposide, vincristine, epirubicin with high-dose cyclosporin)-chemotherapy selected for multidrug resistance modulation. Eur J Cancer 2002;38(18):2422-7.
11. Zhao SC, Banerjee D, Mineishi S, Bertino JR. Post-transplant methotrexate administration leads to improved curability of mice bearing a mammary tumor transplanted with marrow transduced with a mutant human dihydrofolate reductase cDNA. Hum Gene Ther 1997;8(8):903-9.
12. Chatterjee D, Liu CJ, Northey D, Teicher BA. Molecular characterization of the in vivo alkylating agent resistant murine EMT-6 mammary carcinoma tumors. Cancer Chemother Pharmacol 1995;35(5):423-31.
13. MacGrogan G, Mauriac L, Durand M, Bonichon F, Trojani M, de Mascarel I, et al. Primary chemotherapy in breast invasive carcinoma: predictive value of the immunohistochemical detection of hormonal receptors, p53, c-erbB-2, MiB1, pS2 and GST pi. Br J Cancer 1996;74(9):1458-65.
14. Thomas E, Berner G. Prognostic and predictive implications of HER2 status for breast cancer patients. Eur J Oncol Nurs 2000;4(Sa):10-7.
15. Yang W, Klos KS, Zhou X, Yao J, Yang Y, Smith TL, et al. ErbB2 overexpression in human breast carcinoma is correlated with p21Cip1 up-regulation and tyrosine-15 hyperphosphorylation of p34Cdc2: poor responsiveness to chemotherapy with cyclophoshamide methotrexate, and 5-fluorouracil is associated with Erb2 overexpression and with p21Cip1 overexpression. Cancer 2003;98(6):1123-30.
16. Simoes-Wust AP, Schurpf T, Hall J, Stahel RA, Zangemeister-Wittke U. Bcl-2/bcl-xL bispecific antisense treatment sensitizes breast carcinoma cells to doxorubicin, paclitaxel and cyclophosphamide. Breast Cancer Res Treat 2002;76(2):157-66.
17. Sjostrom J, Blomqvist C, von Boguslawski K, Bengtsson NO, Mjaaland I, Malmstrom P, et al. The predictive value of bcl-2, bax, bcl-xL, bag-1, fas, and fasL for chemotherapy response in advanced breast cancer. Clin Cancer Res 2002;8(3):811-6.
18. Su F, Hu X, Jia W, Gong C, Song E, Hamar P. Glutathion S transferase pi indicates chemotherapy resistance in breast cancer. J Surg Res 2003;113(1):102-8.
19. Maxwell PJ, Longley DB, Latif T, Boyer J, Allen W, Lynch M, et al. Identification of 5-fluorouracil-inducible target genes using cDNA microarray profiling. Cancer Res 2003;63(15):4602-6.
20. Kudoh K, Ramanna M, Ravatn R, Elkahloun AG, Bittner ML, Meltzer PS, et al. Monitoring the expression profiles of doxorubicin-induced and doxorubicin-resistant cancer cells by cDNA microarray. Cancer Res 2000;60(15):4161-6.
21. Klijn JG, Berns EM, Foekens JA. Prognostic factors and response to therapy in breast cancer. Cancer Surv 1993;18:165-98.

22. Geisler S, Borresen-Dale AL, Johnsen H, Aas T, Geisler J, Akslen LA, et al. TP53 gene mutations predict the response to neoadjuvant treatment with 5-fluorouracil and mitomycin in locally advanced breast cancer. Clin Cancer Res 2003;9(15):5582-8.
23. Han S, Park K, Bae BN, Kim KH, Kim HJ, Kim YD, et al. E2F1 expression is related with the poor survival of lymph node-positive breast cancer patients treated with fluorouracil, doxorubicin and cyclophosphamide. Breast Cancer Res Treat 2003;82(1):11-6.
24. Sjostrom J, Blomqvist C, Heikkila P, Boguslawski KV, Raisanen-Sokolowski A, Bengtsson NO, et al. Predictive value of p53, mdm-2, p21, and mib-1 for chemotherapy response in advanced breast cancer. Clin Cancer Res 2000;6(8):3103-10.
25. Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. Lancet 2003;362(9381):362-9.
26. Cayre A, Penault-Llorca F, De Latour M, Rolhion C, Feillel V, Ferriere JP, et al. O(6)-methylguanine-DNA methyl transferase gene expression and prognosis in breast carcinoma. Int J Oncol 2002;21(5):1125-31.
27. Nakamura M, Tsuji N, Asanuma K, Kobayashi D, Yagihashi A, Hirata K, et al. Survivin as a predictor of cis-diamminedichloroplatinum sensitivity in gastric cancer patients. Cancer Sci 2004;95(1):44-51.
28. Linn SC, Pinedo HM, van Ark-Otte J, van der Valk P, Hoekman K, Honkoop AH, et al. Expression of drug resistance proteins in breast cancer, in relation to chemotherapy. Int J Cancer 1997;71(5):787-95.
29. Sjostrom J, Krajewski S, Franssila K, Niskanen E, Wasenius VM, Nordling S, et al. A multivariate analysis of tumour biological factors predicting response to cytotoxic treatment in advanced breast cancer. Br J Cancer 1998;78(6):812-5.
30. Pestalozzi BC, Peterson HF, Gelber RD, Goldhirsch A, Gusterson BA, Trihia H, et al. Prognostic importance of thymidylate synthase expression in early breast cancer. J Clin Oncol 1997;15(5):1923-31.
31. Sladek NE, Kollander R, Sreerama L, Kiang DT. Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: a retrospective study. Rational individualization of oxazaphosphorine-based cancer chemotherapeutic regimens. Cancer Chemother Pharmacol 2002;49(4):309-21.
32. Kihara C, Tsunoda T, Tanaka T, Yamana H, Furukawa Y, Ono K, et al. Prediction of sensitivity of esophageal tumors to adjuvant chemotherapy by cDNA microarray analysis of gene-expression profiles. Cancer Res 2001;61(17):6474-9.
33. Goan YG, Zhou B, Hu E, Mi S, Yen Y. Overexpression of ribonucleotide reductase as a mechanism of resistance to 2,2-difluorodeoxycytidine in the human KB cancer cell line. Cancer Res 1999;59(17):4204-7.
34. Jung CP, Motwani MV, Schwartz GK. Flavopiridol increases sensitization to gemcitabine in human gastrointestinal cancer cell lines and correlates with down-regulation of ribonucleotide reductase M2 subunit. Clin Cancer Res 2001;7(8):2527-36.
35. Zhou B, Mo X, Liu X, Qiu W, Yen Y. Human ribonucleotide reductase M2 subunit gene amplification and transcriptional regulation in a homogeneous staining chromosome region responsible for the mechanism of drug resistance. Cytogenet Cell Genet 2001;95(1-2):34-42.

36. Blackstock AW, Lightfoot H, Case LD, Tepper JE, Mukherji SK, Mitchell BS, et al. Tumor uptake and elimination of 2',2'-difluoro-2'-deoxycytidine (gemcitabine) after deoxycytidine kinase gene transfer: correlation with in vivo tumor response. Clin Cancer Res 2001;7(10):3263-8.
37. Hapke DM, Stegmann AP, Mitchell BS. Retroviral transfer of deoxycytidine kinase into tumor cell lines enhances nucleoside toxicity. Cancer Res 1996;56(10):2343-7.
38. Ruiz van Haperen VW, Veerman G, Boven E, Noordhuis P, Vermorken JB, Peters GJ. Schedule dependence of sensitivity to 2',2'-difluorodeoxycytidine (Gemcitabine) in relation to accumulation and retention of its triphosphate in solid tumour cell lines and solid tumours. Biochem Pharmacol 1994;48(7):1327-39.
39. Dumontet C, Fabianowska-Majewska K, Mantincic D, Callet Bauchu E, Tigaud I, Gandhi V, et al. Common resistance mechanisms to deoxynucleoside analogues in variants of the human erythroleukaemic line K562. Br J Haematol 1999;106(1):78-85.
40. Galmarini CM, Mackey JR, Dumontet C. Nucleoside analogues: mechanisms of drug resistance and reversal strategies. Leukemia 2001;15(6):875-90.
41. Plunkett W, Huang P, Xu YZ, Heinemann V, Grunewald R, Gandhi V. Gemcitabine: metabolism, mechanisms of action, and self-potentiation. Semin Oncol 1995;22(4 Suppl 11):3-10.
42. Eliopoulos N, Cournoyer D, Momparler RL. Drug resistance to 5-aza-2'-deoxycytidine, 2',2'-difluorodeoxycytidine, and cytosine arabinoside conferred by retroviral-mediated transfer of human cytidine deaminase cDNA into murine cells. Cancer Chemother Pharmacol 1998;42(5):373-8.
43. Neff T, Blau CA. Forced expression of cytidine deaminase confers resistance to cytosine arabinoside and gemcitabine. Exp Hematol 1996;24(11):1340-6.
44. Mackey JR, Mani RS, Selner M, Mowles D, Young JD, Belt JA, et al. Functional nucleoside transporters are required for gemcitabine influx and manifestation of toxicity in cancer cell lines. Cancer Res 1998;58(19):4349-57.
45. Mackey JR, Yao SY, Smith KM, Karpinski E, Baldwin SA, Cass CE, et al. Gemcitabine transport in xenopus oocytes expressing recombinant plasma membrane mammalian nucleoside transporters. J Natl Cancer Inst 1999;91(21):1876-81.
46. Rosell R, Crino L, Danenberg K, Scagliotti G, Bepler G, Taron M, et al. Targeted therapy in combination with gemcitabine in non-small cell lung cancer. Semin Oncol 2003;30(4 Suppl 10):19-25.
47. Ma T, Zhu ZG, Ji YB, Zhang Y, Yu YY, Liu BY, et al. Correlation of thymidylate synthase, thymidine phosphorylase and dihydropyrimidine dehydrogenase with sensitivity of gastrointestinal cancer cells to 5-fluorouracil and 5-fluoro-2'-deoxyuridine. World J Gastroenterol 2004;10(2):172-6.
48. Gibelli N, Zibera C, Asti A, Maestri L, Bacchella L, Pedrazzoli P, et al. CG5/Dx human breast cancer cell line: characterization of a new doxorubicin-resistant variant. Anticancer Res 1996;16(4A):1675-81.

49. Liu YY, Cabot MC. Development of a mammalian Tet-on expression cell line: glucosylceramide synthase regulates TNF-alpha-induced apoptosis. Methods Mol Biol 2004;249:177-92.