Articles

Assessment of an RNA interference screen-derived mitotic and ceramide pathway metagene as a predictor of response to neoadjuvant paclitaxel for primary triple-negative breast cancer: a retrospective analysis of five clinical trials

Nicolai Juul*, Zoltan Szallasi*, Aron C Eklund*, Qiyuan Li, Rebecca A Burrell, Marco Gerlinger, Vicente Valero, Eleni Andreopoulou, Francisco J Esteva, W Fraser Symmans, Christine Desmedt, Benjamin Haibe-Kains, Christos Sotiriou, Lajos Pusztai, Charles Swanton

Summary

Background Addition of taxanes to preoperative chemotherapy in breast cancer increases the proportion of patients who have a pathological complete response (pCR). However, a substantial proportion of patients do not respond, and the prognosis is particularly poor for patients with oestrogen-receptor (ER)/progesterone-receptor (PR)/human epidermal growth factor receptor 2 (HER2; ERBB2)-negative (triple-negative) disease who do not achieve a pCR. Reliable identification of such patients is the first step in determining who might benefit from alternative treatment regimens in clinical trials. We previously identified genes involved in mitosis or ceramide metabolism that influenced sensitivity to paclitaxel, with an RNA interference (RNAi) screen in three cancer cell lines, including a triple-negative breast-cancer cell line. Here, we assess these genes as a predictor of pCR to paclitaxel combination chemotherapy in triple-negative breast cancer.

Methods We derived a paclitaxel response metagene based on mitotic and ceramide genes identified by functional genomics studies. We used area under the curve (AUC) analysis and multivariate logistic regression to retrospectively assess the metagene in six cohorts of patients with triple-negative breast cancer treated with neoadjuvant chemotherapy; two cohorts treated with paclitaxel (n=27, 30) and four treated without paclitaxel (n=88, 28, 48, 39).

Findings The metagene was associated with pCR in paclitaxel-treated cohorts (AUC 0·79 [95% CI 0·53–0·93], 0·72 [0·48–0·90]) but not in non-paclitaxel treated cohorts (0·53 [0·31–0·77], 0·59 [0·22–0·82], 0·53 [0·36–0·71], 0·64 [0·43–0·81]). In multivariate logistic regression, the metagene was associated with pCR (OR 19·92, 2·62–151·57; p=0·0039) with paclitaxel-containing chemotherapy.

Interpretation The paclitaxel response metagene shows promise as a paclitaxel-specific predictor of pCR in patients with triple-negative breast cancer. The metagene is suitable for development into a reverse transcription-PCR assay, for which clinically relevant thresholds could be established in randomised clinical trials. These results highlight the potential for functional genomics to accelerate development of drug-specific predictive biomarkers without the need for training clinical trial cohorts.

Funding UK Medical Research Council; Cancer Research UK; the National Institute for Health Research (UK); the Danish Council for Independent Research-Medical Sciences (FSS); Breast Cancer Research Foundation (New York); Fondation Luxembourgeoise contre le Cancer; the Fonds National de la Recherche Scientifique; Brussels Region (IRSIB-IP, Life Sciences 2007) and Walloon Region (Biowin-Keymarker); Sally Pearson Breast Cancer Fund; and the European Commission.

Introduction

Despite the use of modern cytotoxic agents, the proportion of patients who achieve a complete pathological response (pCR) to preoperative chemotherapy remains low, at 15–25% for all breast-cancer histopathological subtypes.[1] Rates of pCR in sporadic oestrogen-receptor (ER)/progesterone-receptor (PR)/human epidermal growth factor receptor 2 (HER2; ERBB2)-negative (triple-negative) breast cancer range from 12% for taxane monotherapy to 45% with combination neoadjuvant chemotherapy regimens.[2–4] Patients who achieve a pCR after chemotherapy have excellent disease-free and overall survival.[1] The outcome for patients who do not achieve a pCR varies, and

the prognosis for patients with triple-negative cancers and residual disease after preoperative chemotherapy is particularly poor.[2] Better understanding of breast-cancer biology is likely to expand the list of potentially effective chemotherapeutic agents in the neoadjuvant setting, and will help identify tailored chemotherapy schedules for distinct patient cohorts based on tumour molecular characterisation. If it can be reliably established that patients resistant to one type of therapy are sensitive to a different agent, then robust predictors of chemotherapy response will have an essential role in selecting the optimum treatment in the neoadjuvant setting. The identification of patients with disease that is resistant to

Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark (N Juul MSc, Prof Z Szallasi MD, A C Eklund MSc, Q Li MSc); Children's Hospital Informatics Program at the Harvard-MIT Division of Health Sciences and Technology, Harvard Medical School, Boston, MA, USA (Z Szallasi); Translational Cancer Therapeutics Laboratory, Cancer Research UK London Research Institute, London, UK (R A Burrell BA, M Gerlinger MD, C Swanton MRCP); Institute of Cancer, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK (M Gerlinger); Department of Breast Medical Oncology (V Valero MD, E Andreopoulou MD, Prof F J Esteva MD, Prof L Pusztai MD) and Department of Pathology, The University of Texas M D Anderson Cancer Center, Houston, TX, USA (Prof W F Symmans MD); Medical Oncology Department, Jules Bordet Institute, Brussels, Belgium (C Desmedt PhD, B Haibe-Kains PhD, C Sotiriou MD); Machine Learning Group, Université Libre de Bruxelles, Brussels, Belgium (B Haibe-Kains); and Royal Marsden Hospital, Department of Medicine, Breast Unit, Sutton, UK (C Swanton)

Correspondence to:
Dr Charles Swanton, Translational Cancer Therapeutics Laboratory, Cancer Research UK London Research Institute, 44 Lincoln's Inn Fields, London WC2A 3PX, UK
charles.swanton@cancer.org.uk

conventional chemotherapy combinations is the first step in this process.

Several predictive biomarkers have been discovered using associative-learning strategies. Predictive gene-expression signatures derived from an associative analysis approach are susceptible to chance associations which lead to overestimation of true clinical accuracy; therefore, two separate trial cohorts are required to train and validate the predictive signature.[5,6] Associative strategies developed from genomics signatures that are predictive of drug response in cell lines might circumvent the need for training and validation of trial cohorts.[7] Quantifying distinct biological processes within gene-expression datasets, instead of a gene-by-gene based associative analysis, may further avoid these problems and accelerate biomarker development.[8] However, potentially relevant biological processes or functional modules must be identified in advance for such an analysis to be possible.

RNA interference (RNAi) functional screening may be an applicable method to identify biological processes relevant to drug response in cancer medicine. To test this hypothesis, we revisited results of our previous study[9] of an RNAi drug-resistance screen across three cancer cell lines, including a triple-negative breast-cancer cell line, MDA-MB-231. In this screen, we identified two distinct gene sets regulating sensitivity to paclitaxel.[9] The first set of genes is involved in mitosis and the mitotic spindle assembly checkpoint (SAC), and the second set is involved in metabolism of the proapoptotic lipid, ceramide. The involvement of both of these gene sets is consistent with the current biological understanding of the mechanism of action of paclitaxel. An activated SAC orchestrates a paclitaxel-induced mitotic arrest, and in our RNAi screen, silencing several genes implicated in SAC control impaired the accumulation of cells in mitosis and subsequent cell death in response to paclitaxel. Identification of ceramide pathway genes as regulators of paclitaxel sensitivity[9] is consistent with published evidence that overexpression of glucosylceramide synthase (UGCG) promotes resistance to paclitaxel and repression promotes paclitaxel sensitivity.[10–12]

We created a **metagene** using established methodology[8,13] to quantify the activity of these two biological pathways identified by our RNAi screen, and tested its paclitaxel-predictive value in patients with triple-negative breast cancer who were treated in clinical trials with either a paclitaxel-containing combination regimen, T-FAC (paclitaxel followed by fluorouracil, doxorubicin, and cyclophosphamide) or by regimens without paclitaxel.

Methods

Patients and procedures

Gene-expression and treatment-response data were retrieved from six cohorts in five neoadjuvant clinical trials, referred to in this study as MDA1, MDA/MAQC-II, TOP, EORTC$_{FEC}$, EORTC$_{TET}$, and DFCI. In all trials, pCR was determined at the time of surgery (no evidence of residual invasive cancer in the breast or lymph nodes at time of

	MDA1	MDA/MAQC-II	TOP	DFCI	EORTC10944	
					FEC group	TET group
Microarray platform	U133A	U133A	U133 Plus 2.0	U133 Plus 2.0	X3P	X3P
Treatment	T-FAC	T-FAC	Epirubicin	Cisplatin	FEC	TET
Number of patients	133	100	120	28	102	58
Median age (years; range)	51 (28–79)	50 (26–73)	47 (27–68)	50 (29–69)	49 (26–70)	49 (34–70)
Histology; n (%)						
IDC	130 (98)	93 (93)	112 (93)	28 (100)	52 (51)	56 (97)
ILC	1 (1)	7 (7)	1 (1)	0 (0)	4 (4)	2 (3)
Other	2 (1)	0 (0)	7 (6)	0 (0)	7 (7)	0 (0)
Unknown	0 (0)	0 (0)	0 (0)	0 (0)	39 (38)	0 (0)
Stage; n (%)						
T1	13 (10)	10 (10)	17 (14)	1 (4)	2 (2)	1 (2)
T2	70 (52·5)	62 (62)	83 (69)	22 (79)	63 (62)	37 (64)
T3	22 (16·5)	13 (13)	5 (4)	4 (14)	34 (33)	20 (34)
T4	28 (21)	15 (15)	15 (13)	0 (0)	0 (0)	0 (0)
Unknown	0 (0)	0 (0)	0 (0)	1 (4)	3 (3)	0 (0)
Nodes; n (%)						
SLN positive	ND	ND	ND	13 (46)	ND	ND
N0	40 (30)	27 (27)	55 (46)	15 (54)	37 (36)	22 (38)
N1	62 (47)	47 (47)	60 (50)	ND	55 (54)	31 (53)
N2	14 (11)	13 (13)	3 (2)	ND	7 (7)	5 (9)
N3	17 (13)	13 (13)	2 (2)	ND	0 (0)	0 (0)
Unknown	0 (0)	0 (0)	0 (0)	0 (0)	3 (3)	0 (0)
Grade; n (%)						
1	2 (1)	11 (11)	2 (2)	0 (0)	2 (2)	0 (0)
2	54 (41)	42 (42)	20 (17)	0 (0)	21 (21)	16 (28)
3	77 (58)	47 (47)	92 (77)	28 (100)	31 (30)	37 (64)
Unknown	0 (0)	0 (0)	6 (5)	0 (0)	48 (47)	5 (9)
Oestrogen receptor status; n (%)						
Positive	82 (62)	61 (61)	1 (1)	0 (0)	37 (36)	0 (0)
Negative	51 (38)	39 (39)	119 (99)	28 (100)	65 (64)	58 (100)
Progesterone receptor status; n (%)						
Positive	75 (56)	51 (51)	1 (1)	0 (0)	25 (25)	0 (0)
Negative	55 (41)	49 (49)	75 (62)	28 (100)	75 (74)	58 (100)
Unknown	3 (2)	0 (0)	44 (37)	0 (0)	2 (2)	0 (0)
HER2 status; n (%)						
Amplified	33 (25)	7 (7)	32 (27)	0 (0)	26 (25)*	19 (33)*
Not amplified	99 (74)	93 (93)	61 (51)	28 (100)	76 (75)*	39 (67)*
Unknown	1 (1)	0 (0)	27 (23)	0 (0)	0 (0)*	0 (0)*
Total triple-negative; n (%)	27 (20)	30 (30)	88 (73)**	28 (100)	48 (47)	39 (67)
Response; n (%)						
pCR	34 (26)	15 (15)	17 (14)	5 (18)	39 (38)	26 (45)
No pCR	99 (74)	85 (85)	101 (84)	21 (75)	63 (62)	32 (55)
Non-evaluable	0 (0)	0 (0)	2 (2)	2 (7)	0 (0)	0 (0)

T stage was measured before treatment. Nodal status was confirmed at surgery for MDA1, MDA/MAQC-II, TOP, and EORTC. Nodal status for DFCI was assessed before treatment with SLN biopsy. T-FAC=paclitaxel, fluorouracil, doxorubicin, and cyllophosphamide. FEC=fluorouracil, epirubicin, and cyclophosphamide. TET=docetaxel followed by epirubicin and docetaxel. IDC=invasive ductal carcinoma. ILC=invasive lobular carcinoma. SLN=sentinel lymph node. ND=not determined. pCR=pathological complete response. *Receptor status inferred by gene expression. **Triple-negative breast cancer subtype inferred by expression when insufficient data by immunohistochemistry/FISH.

Table 1: **Clinical and histopathological characteristics of the cohorts analysed**

surgery). Two clinical trial cohorts were treated with paclitaxel-containing regimens (T-FAC): the MDA1 cohort[14] and the MDA/MAQC-II cohort (GEO accession number GSE16716). Patients in the TOP cohort (GSE16446) were treated with neoadjuvant epirubicin. The gene expression substudy of the EORTC 10994 trial[7,8] randomised patients to FEC (fluorouracil, epirubicin, and cyclophosphamide) or TET (docetaxel followed by epirubicin and docetaxel), corresponding to the EORTC$_{FEC}$ (GSE6861) and EORTC$_{TET}$ (GSE4779) subgroups, respectively. In the DFCI cohort (GSE18864), triple-negative patients were treated with neoadjuvant cisplatin.[15] Patient characteristics for all cohorts are summarised in table 1, and characteristics for the subset of patients with triple-negative disease are shown in the webappendix p 1. We referred to microarray-based expression subtyping if we were unable to fully classify tumours by histology. All clinical trials and translational parameters were approved by the ethics committee and institutional review boards of the participating institutions. Further cohort-specific information is detailed in the webappendix pp 10–12.

Validated datasets from our previously published RNAi screen were used for the derivation of the paclitaxel response metagene (webappendix pp 10–12).[9] The RNAi screen focused on kinases and ceramide-pathway genes. In the kinase screen, we identified ten mitosis-associated genes that influenced sensitivity to paclitaxel in three different cell lines. In addition, we identified three genes from the ceramide pathway (COL4A3BP, GBA1, GBA3) that together with the published gene UGCG,[10–12] encode proximal regulators of ceramide metabolism and influence paclitaxel sensitivity. Of these 14 genes (figure 1), six mitotic genes and three ceramide genes could be assayed on the Affymetrix HGU133A platform, and these were combined into the mitotic and the ceramide gene sets. The number of genes within each gene set was further reduced to the mitotic module (four genes) and the ceramide module (two genes) by only including genes that were significantly correlated with each other across four independent breast cancer datasets.[16–19] The expression of the genes within each module was compressed by taking the mean. Since the two modules were expected to predict sensitivity in the opposite direction (high mitotic expression=sensitivity, high ceramide expression=resistance), the paclitaxel response metagene was defined as the mitotic module minus the ceramide module. As a comparison, 10 000 random combinations of nine-gene sets were subjected to the same correlative approach, across four clinical datasets, and the mean expression of significantly correlated genes were tested for their ability to predict pCR in T-FAC-treated cohorts. We also did a literature search and found 24 genes reported as affecting taxane sensitivity when either overexpressed or repressed (webappendix p 6). These genes were subjected to the same correlation analysis as the paclitaxel response metagene, and the reported direction of expression relating to paclitaxel sensitivity was included as weights (minus 1 if the gene induced sensitivity when repressed, 1 if the gene induced sensitivity when overexpressed). The genomic grade index and the stroma signatures were calculated as described previously.[8,20]

Statistical analysis

A binomial test was used to test for enrichment in genes that predicted response to treatment with paclitaxel-containing neoadjuvant chemotherapy, among the genes that substantially affected paclitaxel sensitivity across three cell lines. The significance of association between module or metagene scores and pCR was estimated with the one-sided Wilcoxon signed-rank test and plotted with receiver operating characteristics curves. We used a one-sided Wilcoxon test because we had a prior expectation about the direction of association, based on the RNAi screening results.[9] The effect size was estimated with AUC and logistic regression. Since the MDA1 and MDA/MAQC-II trials were done by the same investigators, at the same site, with identical gene-expression platforms, we combined the two T-FAC
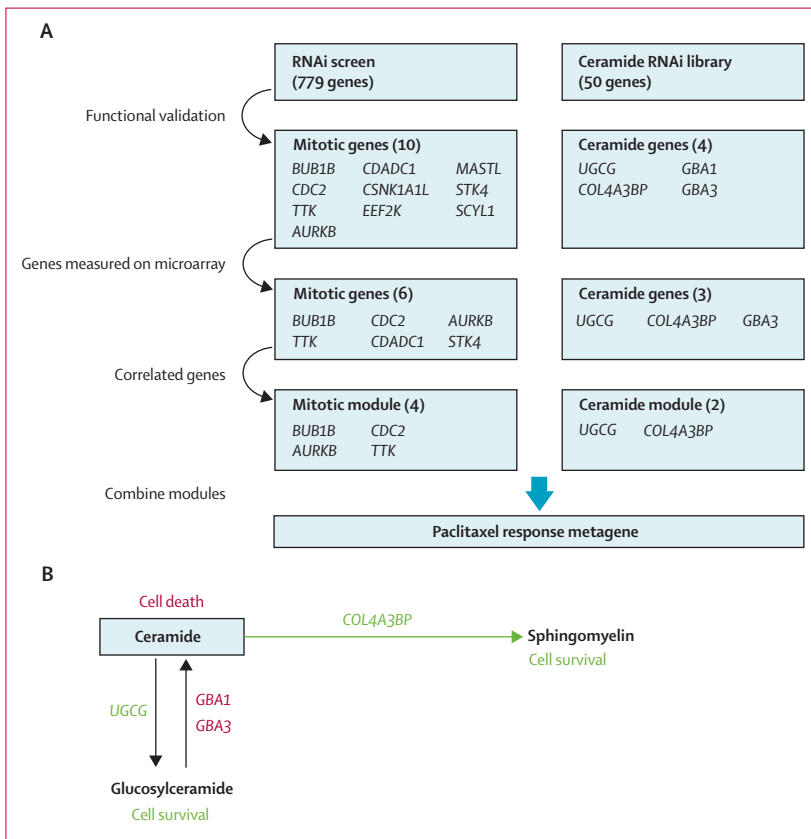

Figure 1: Process to derive the paclitaxel response metagene
(A) Flow chart describing the derivation of the gene modules and the paclitaxel response metagene. *UGCG* was identified from the literature as a gene in the same pathway that promotes paclitaxel sensitivity when repressed. (B) Proximal regulators of the proapoptotic lipid, ceramide, that alter paclitaxel sensitivity through conversion to sphingomyelin (via COL4A3BP ceramide transporter) or glucosylceramide (UGCG glucosylceramide synthase) and conversion of glucosylceramide to ceramide (GBA1 and GBA3 beta-glucosidase). Genes for which repression promotes paclitaxel sensitivity are shown in green and for which repression promotes paclitaxel resistance are shown in red. RNAi=RNA interference.

treated cohorts to increase statistical power for univariate and multivariate logistic regression analysis. Univariate and multivariate logistic regression was done using iteratively reweighted least squares. All covariates were converted to binary, using the median as a threshold to determine paclitaxel response metagene status and Ki67 status. Analyses compared stage T3 and T4 versus T1 and T2; grade 3 versus grade 1 and 2; and N0 versus N1+. A p value of 0·05 or lower was considered significant for all statistical tests. All statistical analyses were done in R version 2.9.2 using the software packages affy, genefilter, beeswarm, gdata, ROC, hgu133a.db, u133x3p.db, squash, boot, rmeta, and survplot (webappendix p 11).

Role of the funding source

The funding sources and sponsors of the trials had no role in the design of the study; collection, analysis, or interpretation of the data; or writing of this report. All authors had access to the raw data. The corresponding author had full access to all data and had the final responsibility to submit for publication.

Results

To develop an analytical framework based on experimentally established functional links instead of associative correlations, we revisited our previously published functional genomic RNAi screen that identified genes which either promote resistance or sensitivity to paclitaxel.[9] Based on this screen and previously published data, we identified two gene modules of tightly correlated genes: a four-gene mitotic module where higher expression predicted sensitivity, and a two-gene ceramide module where higher expression predicted resistance. We then used the mean expression of the genes within each

module as a single predictive value; an approach previously shown to improve reliability across many tumour samples.[8,13] Finally, we combined the mitotic and ceramide modules into a single functionally derived paclitaxel-response metagene by subtracting the mean expression of the genes in the ceramide module from the mean expression of the genes in the mitotic module (figure 1). Therefore, this summary measure reflects the difference in the mean expression of the two modules and is predicted to correlate with paclitaxel sensitivity.

In the T-FAC treated triple-negative cohorts, the paclitaxel response metagene was highly predictive of pCR, with an AUC of 0·79 in MDA1 (p=0·0053; 95% CI 0·53–0·93; n=27), an AUC of 0·72 in MDA/MAQC (p=0·031; 0·48–0·90; n=30), and an AUC of 0·74 when the two cohorts were combined (p=0·0013; 0·58–0·86; n=57; table 2, figure 2 A and B, webappendix p 2). Consistent with these data, the paclitaxel response metagene was also predictive of response across all patients in each cohort: for MDA1, AUC was 0·79 (p<0·0001; 0·69–0·87; n=133); for MDA/MAQC-II, AUC was 0·76 (p=0·00061; 0·61–0·87; n=100); and for the combined cohort, AUC was 0·77 (p<10⁻⁸; 0·69–0·84; n=233; data not shown).

Measurements of gene expression in the cohorts treated with neoadjuvant paclitaxel were consistent with in-vitro functional observations; repression of genes in the mitotic module was associated with resistance to T-FAC therapy, and repression of genes in the ceramide module was associated with sensitivity to T-FAC. Moreover, genes whose suppression increased paclitaxel resistance across all three cancer cell lines of different tissue origin were enriched in a list of genes that are differentially expressed in patients who were sensitive or resistant to paclitaxel (a binomial test addressing the probability that four of six

	Mitotic module	Ceramide module	Paclitaxel response metagene	Paclitaxel response metagene including all module genes	Literature-based metagene
All patients					
MDA1	0·73; 0·63–0·82; <0·0001	0·74; 0·62–0·83; <0·0001	0·79 ;0·69–0·87; <0·0001	0·79; 0·68–0·85; <0·0001	0·52; 0·39–0·62; 0·39
MDA/MAQC-II	0·71; 0·52–0·83; 0·0059*	0·73; 0·58–0·85; 0·0023*†	0·76; 0·61–0·87; 0·00061*	0·75; 0·58–0·86; 0·00094*	0·56; 0·35–0·72; 0·77†
TOP	0·49; 0·34–0·65; 0·57	0·50; 0·33–0·69; 0·51†	0·50; 0·34–0·66; 0·52	0·50; 0·36–0·65; 0·49	0·56; 0·41–0·71; 0·80†
EORTC_FEC	0·57; 0·45–0·67; 0·11	0·59; 0·48–0·71; 0·061†	0·62; 0·49–0·72; 0·025*	0·62; 0·51–0·72; 0·023*	0·55; 0·43–0·66; 0·81†
EORTC_TET	0·53; 0·37–0·68; 0·64	0·54; 0·39–0·69; 0·71†	0·53; 0·38–0·67; 0·65	0·52; 0·37–0·70; 0·61†	0·55; 0·40–0·70; 0·76†
Triple-negative					
MDA1	0·70; 0·45–0·87; 0·038*	0·65; 0·41–0·86; 0·093†	0·79; 0·53–0·93; 0·0053*	0·74; 0·51–0·9; 0·017*	0·55; 0·29–0·79; 0·33
MDA/MAQC-II	0·63; 0·36–0·85; 0·14	0·67; 0·38–0·86; 0·075†	0·72; 0·48–0·9; 0·031*	0·71; 0·43–0·88; 0·035*	0·58; 0·31–0·78; 0·75†
TOP	0·50; 0·33–0·72; 0·51	0·55; 0·4–0·74; 0·31†	0·53; 0·31–0·77; 0·38	0·52; 0·3–0·71; 0·42	0·64; 0·38–0·83; 0·91†
EORTC_FEC	0·54; 0·36–0·72; 0·3	0·50; 0·33–0·68; 0·52†	0·53; 0·36–0·71; 0·35	0·54; 0·36–0·69; 0·33	0·59; 0·41–0·74; 0·87†
EORTC_TET	0·63; 0·44–0·80; 0·92†	0·64; 0·44–0·79; 0·93	0·64; 0·43–0·81; 0·93†	0·61; 0·40–0·78; 0·89†	0·52; 0·31–0·70; 0·59†
DFCI	0·57; 0·04–0·87; 0·33	0·56; 0·28–0·89; 0·35†	0·59; 0·22–0·82; 0·28	0·56; 0·20–0·94; 0·35	0·70; 0·0–0·98; 0·92†

Data are AUC; 95% CI; p value. Histopathological subtypes were based on pathology review by ER/PR/HER2 immunohistochemistry or HER2 FISH, supplemented by expression if histopathological data were unavailable. AUC=area under the curve. FEC=fluorouracil, epirubicin, and cyclophosphamide. TET=docetaxel followed by epirubicin and docetaxel. ER=oestrogen receptor. PR=progesterone receptor. *Significant association between module or metagene and treatment response. †These values predict resistance to paclitaxel, all others predict sensitivity.

Table 2: Association between gene sets and treatment response for all patients and for triple-negative patients for each cohort
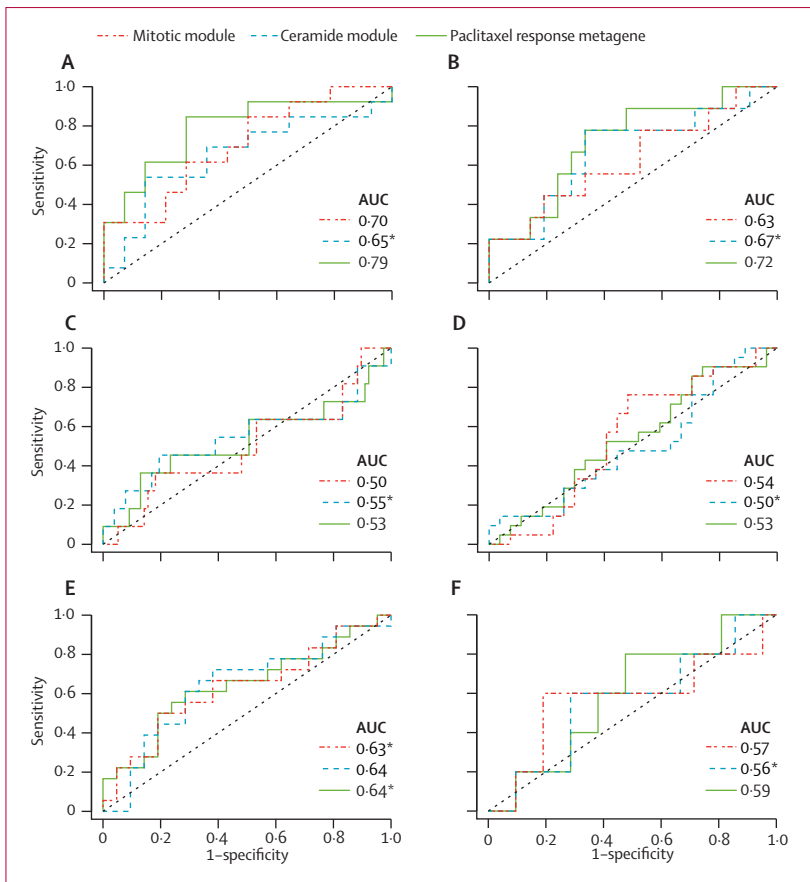
Figure 2: Receiver operating characteristic (ROC) curves for the gene modules and the paclitaxel response metagene

ROC curve in triple-negative patients in the T-FAC-treated cohorts MDA1 (A) and MDA/MAQC-II (B), the epirubicin-treated TOP cohort (C), the FEC-treated EORTC cohort (D), the TET-treated EORTC cohort (E), and in the cisplatin-treated DFCI cohort (F). AUC=area under the curve. FEC=fluorouracil, epirubicin, and cyclophosphamide. TET=docetaxel followed by epirubicin and docetaxel. *These values predict paclitaxel resistance; all others predict sensitivity.

statistical power. We found that the paclitaxel response metagene was the covariate most significantly associated with pCR (p=0·0039; odds ratio 19·92; 95% CI 2·62–151·57) in T-FAC-treated patients with triple-negative breast cancer, more than nodal status, T stage, tumour grade, and Ki67 (table 3). The paclitaxel response metagene also did better than the genomic grade index and the stroma signature in both univariate and multivariate analysis of the combined T-FAC clinical trials, when all patients were considered and in the triple-negative cohorts (table 4).[8,20] These data support the conclusion that the paclitaxel response metagene derived from an RNAi screen has predictive power in two paclitaxel clinical trial cohorts that had no role in the discovery of the genes included in the metagene. The response metagene was not significantly associated with recurrence-free survival in two untreated triple-negative breast cancer cohorts, indicating the predictive rather than prognostic power of the metagene (webappendix p 4).[16,18]

To further validate the functional importance of the genes contained within the metagene and limit the possibility that the correlation step would artificially enrich for genes predictive of pCR, we tested the ability to predict for pCR of 10 000 random nine-gene sets subjected to the same correlative approach. None of the 10 000 combinations predicted for pCR better than the paclitaxel response metagene (p<0·0001). We addressed the relevance of the RNAi screening process to select genes for inclusion in the metagene by performing the same analysis with 24 genes reported to be associated with paclitaxel or docetaxel resistance. The literature metagene did not show any predictive value in the MDA1 or MDA/MAQC-II cohorts. In a further analysis, we eliminated the correlation step from the derivation of the paclitaxel response metagene. Although the nine-gene set (figure 1) selected before the expression correlation step was still predictive of pCR with T-FAC, this gene set did not do as well as the paclitaxel response metagene. These results suggest that unbiased selection of correlated genes with consistent phenotypes across an RNAi screen improves the performance of the predictive metagene (table 2).

To assess the paclitaxel specificity of the paclitaxel response metagene, we assessed its predictive power in four cohorts that did not receive paclitaxel (table 1, webappendix p 5): the EORTC 10994 FEC trial cohort, EORTC 10994 TET trial cohort, the TOP epirubicin trial cohort, and the triple-negative DFCI cisplatin-treated cohort. The paclitaxel response metagene did not predict response in the triple-negative subtype (p>0·05, figure 2 C–F). Notably, the paclitaxel response metagene did not predict a significant response to the TET regimen. The triple-negative tumours analysed were somewhat homogeneous (reflecting the neoadjuvant setting of these trials), with similar nodal status, T stage, and were higher grade tumours (webappendix p 5), indicating that tumour heterogeneity is unlikely to account for this result.

mitotic genes from the RNAi screen were significantly predictive of pCR in the MDA1 cohort when compared with the background probability of 3802/26 138 that a gene in the MDA1 cohort was significantly predictive of pCR [OR 11·40, 95% CI 2·088–62·29; p=0·0052] and that four of six mitotic genes from the RNAi screen were significantly predictive of pCR in the MDA/MAQC-II cohort compared with the background probability of 1986/26 138 that a gene in the MDA1 cohort was significantly predictive of pCR [OR 23·55, 95% CI 4·31–128·69; p=0·00044], and that four of six mitotic genes from the RNAi screen were significantly predictive of pCR in both the MDA1 and the MDA/MAQC-II cohorts compared with the background probability of 1063/26 138 that a gene in both cohorts were significantly predictive of pCR [OR 45·61, 95% CI 8·34–249·30; p=0·000039]. These data suggest that the concordance of the functional genomic and clinical trial genomics datasets is unlikely to result from a chance association.

For multivariate analysis, we combined the two T-FAC-treated MDA1 and MDA/MAQC-II cohorts to increase

Using a meta-analysis, we combined the odds ratios of the paclitaxel response metagene to pCR in the two paclitaxel-treated cohorts and in the four non-paclitaxel treated cohorts. We found that the summary odds ratio of the paclitaxel treated cohorts was 5·65 (95% CI 1·67–19·11; p=0·0053), whereas the summary odds ratio of the non-paclitaxel treated cohorts was 0·87 (95% CI 0·44 to 1·67; p=0·67), consistent with improved predictive power of the paclitaxel response metagene in paclitaxel treated cohorts (figure 3).

Finally, we combined paclitaxel-treated with non-paclitaxel-treated triple-negative cohorts and did logistic-regression analysis using a mixed effects model, considering paclitaxel treatment, binary metagene status, and their interaction. We observed a significant interaction term between paclitaxel treatment and binary metagene status (OR 5·9, 95% CI 1·61–23·18; p=0·0089), indicating that the paclitaxel response metagene has paclitaxel-specific predictive power.

Discussion

This study supports the use of high-throughput RNAi functional genomics screening to accelerate discovery of predictive biomarkers in cancer medicine. By filtering for common paclitaxel resistance pathways through RNAi screening across three cell lines of different tumour origin, and selecting genes which correlate across independent cohorts, we derived a paclitaxel response metagene that is predictive of T-FAC response in two clinical trial datasets, but not in cohorts treated without paclitaxel. Our results show the usefulness of this approach to identify drug-specific response predictors. Data reported here support a model whereby expression of genes that regulate mitotic arrest and chromosomal stability, mediated through spindle assembly checkpoint signalling, and genes that influence ceramide conversion to sphingomyelin or glucosylceramide, are associated with altered response to T-FAC therapy in vivo (figure 4).[9,21]

Since the metagene was derived from an RNAi screen in a triple-negative breast-cancer cell line, we investigated the power of the metagene to predict response to paclitaxel in clinical trials that include patients with this histological subtype. The patient demographics in the four non-paclitaxel trials included in this study are typical of the neoadjuvant setting and similar to MDA1 and MDA/MAQC-II, with a bias towards node positive, higher grade (G2–G3) and larger tumours (T2 and above), indicating that tumour heterogeneity across the trial cohorts is unlikely to contribute to the observed results. However, although the paclitaxel response metagene performs better than the genomic grade index and the stromal signature in multivariate analysis, both of which have predictive value in patients treated with T-FAC,[8,20] the AUCs reported in the MDA1 cohort result in false negatives and denial of active therapy in two of 13 (15%) of the responding patients, to spare suboptimal treatment in ten of 14 patients with resistant disease. A false negative proportion of one in

	Univariate analysis	Multivariate analysis
All patients		
Oestrogen receptor	0·084; 0·04–0·19; <0·0001*	0·18; 0·07–0·47; 0·00039*
HER2	3·84; 1·84–8·01; 0·00037*	2·35; 1·01–5·47; 0·048*
T stage	1·07; 0·55–2·08; 0·84	0·67; 0·29–1·52; 0·34
Ki67	3·60; 1·79–7·23; 0·00032*	1·17; 0·45–3·07; 0·75
Grade	6·24; 2·77–14·04; <0·0001*	1·99; 0·73–5·44; 0·18
Node	2·42; 1·07–5·48; 0·034*	1·98; 0·77–5·06; 0·15
Metagene	7·28; 3·23–16·41; <0·0001*	1·79; 0·56–5·70; 0·32
Triple-negative		
T stage	0·79; 0·26–2·44; 0·68	0·39; 0·09–1·60; 0·19
Ki67	0·79; 0·27–2·29; 0·66	0·10; 0·01–0·80; 0·030*
Grade	3·46; 0·67–17·83; 0·14	6·59; 0·76–57·36; 0·088
Node	3·46; 0·67–17·83; 0·14	9·62; 1·02–90·28; 0·048*
Metagene	4·51; 1·41–14·43; 0·011*	19·92; 2·62–151·57; 0·0039*

Data are odds ratio; 95% CI; p value. All covariates are binary. The binary metagene and binary Ki67 is defined as the upper vs lower 50th percentile. T stage is T3–4 vs T1–2. Grade is 3 vs 1 and 2. Node is N+ versus N0. *Significant values.

Table 3: Univariate and multivariate analysis for response to treatment (combined analysis)

	Univariate analysis	Multivariate analysis
All patients		
Stroma metagene	0·45; 0·23–0·86; 0·016*	0·55; 0·27–1·11; 0·095
Genomic grade index	4·11; 1·98–8·54; 0·00015*	1·59; 0·66–3·80; 0·30
Paclitaxel response metagene	7·28; 3·23–16·41; <0·0001*	5·37; 2·12–13·64; 0·00040*
Triple-negative		
Stroma metagene	0·28; 0·09–0·85; 0·025*	0·27; 0·07–1·03; 0·055
Genomic grade index	2·35; 0·65–8·52; 0·19	0·41; 0·06–2·68; 0·35
Paclitaxel response metagene	4·51; 1·41–14·43; 0·011*	5·47; 1·21–24·76; 0·027*

Data are odds ratio; 95% CI; p value. *Significant values.

Table 4: Performance of the paclitaxel response metagene compared with published gene-expression signatures in univariate and multivariate logistic regression (combined analysis)
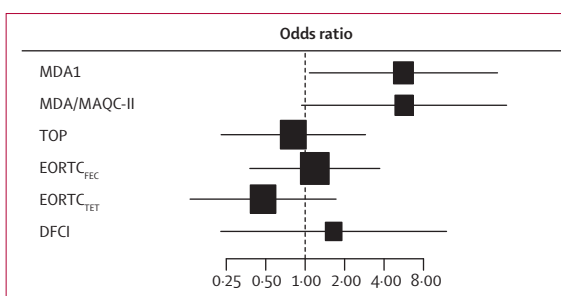

Figure 3: Odds ratio for pathological complete response in patients with high paclitaxel response metagene score, compared with a low score
FEC=fluorouracil, epirubicin, and cyclophosphamide. TET=docetaxel followed by epirubicin and docetaxel.

13 (8%) would be possible by sacrificing specificity, resulting in the sparing of suboptimal therapy in seven of 14 patients with resistant disease. A false negative proportion of 8% would compare with the false negative proportion of ER analysis in breast cancer (Sotiriou C, unpublished data). We plan to assess whether the
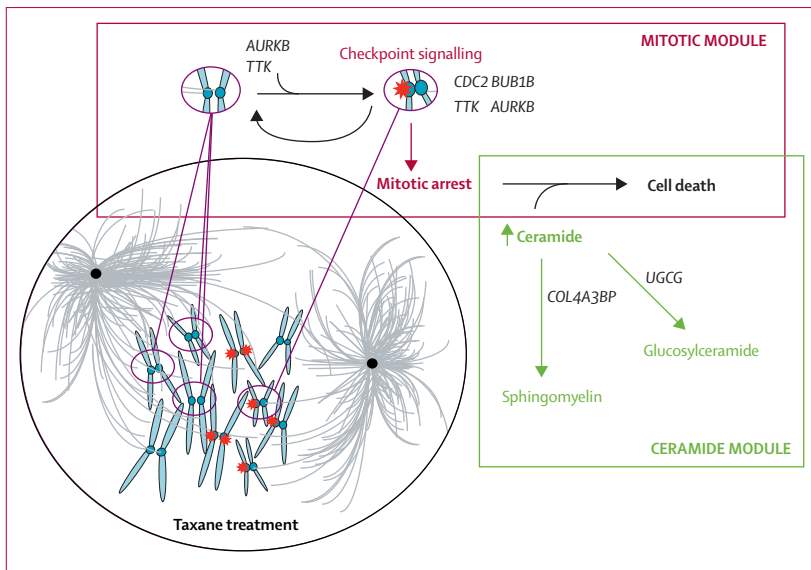
Figure 4: Model for the role of the metagene in paclitaxel response
Paclitaxel stabilises microtubules (green) and causes mitotic arrest, due to spindle assembly checkpoint signalling (red) generated by unattached kinetochores (yellow). BUB1B, TTK, and AURKB all have roles in checkpoint signalling, which maintains the activity of CDC2 to keep cells arrested in mitosis. AURKB may maintain checkpoint activation through detachment of syntelic attachments (purple rings and magnified section) which may be promoted during paclitaxel exposure. Following a mitotic arrest, sensitive tumour cells undergo apoptosis. The mechanisms for this are currently unclear. Ceramide is a proapoptotic lipid that may contribute to paclitaxel-induced cell death following a mitotic arrest. Decreased levels of the mitotic module may attenuate a mitotic arrest, and consequently, cell death. Conversely, decreased levels of the ceramide transporter, COL4A3BP (CERT), and glucosylceramide synthase, UGCG, may lead to an increase in the ceramide pool and potentiate paclitaxel-induced cytotoxicity.

performance of the paclitaxel response metagene approach might be improved through gene selection from genome-wide RNAi screening approaches targeting more than 21 000 genes across multiple cell lines, by contrast with the 829 genes assessed in this study.

Sources of random and systematic error should be considered when interpreting these data. Notably, the triple-negative breast-cancer cohorts are likely to be molecularly heterogeneous. Although no patients enrolled in the T-FAC trials were known to have germline *BRCA* mutations, the same DNA repair-pathway mechanisms may be disrupted in sporadic breast cancers that have also been implicated in taxane resistance in vitro.[22,23] Also, it should be noted that the T-FAC clinical trial datasets were acquired from fine needle aspirations whereas the non-paclitaxel datasets were acquired from core biopsies. RNA yield and expression profiling are similar using both techniques,[24] but we cannot exclude the possibility that the enrichment of stromal elements in the core biopsy datasets contribute to the lack of predictive power of the paclitaxel response metagene. There is heterogeneity between the clinical studies including the timing and measurement of pCR by different pathological centres. The five trials examined here varied in chemotherapy exposure from 12 to 24 weeks, and two trials used monotherapy schedules that might affect the proportion of patients achieving a pCR.

A metagene derived from reports of genes implicated in taxane resistance did not show T-FAC predictive power.

Furthermore, none of 10 000 random nine-gene sets performed better than the metagene. While these results support the RNAi approach to biomarker discovery and argue against the role of chance in these findings, the modest cohort sizes and the heterogeneity of the non-paclitaxel trials require replication in larger prospective studies to confirm the relevance of this method with a clinically applicable gene-expression assay. Experience with the Oncotype DX assay has shown that RNA-based expression measurements (real-time PCR) from paraffin-fixed tumour material can inform clinical decision making. A similar assay to assess the expression of the paclitaxel response metagene could be developed that, we estimate, would cost less than €30 per patient. With this assay, exact thresholds of mitotic and ceramide module expression should be defined retrospectively with tumours from the T-FAC trial before testing the defined threshold in a prospective trial. A randomised clinical trial comparing a paclitaxel with a non-paclitaxel regimen will be required to formally support the paclitaxel-specificity of the metagene and the relevance of RNAi to the biomarker discovery process. The usefulness of this approach in routine clinical practice should then be assessed, since patients enrolled in clinical trials may not accurately reflect the demographics and clinical stage of patients diagnosed with primary breast cancer.

Notably, the paclitaxel response metagene was not predictive of pCR in the EORTC$_{TET}$ cohort (docetaxel then epirubicin and docetaxel), which may be explained by the non-overlapping pathways of drug resistance to the two taxanes. The metagene was derived from a screen to identify mediators of paclitaxel not docetaxel resistance across three cell lines, and preclinical data has shown that docetaxel binds to β-tubulin with greater affinity than paclitaxel and has increased interphase (G1/S/G2) cell-cycle activity, mediating cell death through the induction of BCL2 phosphorylation.[25,26] Finally, 19–31% of patients respond to paclitaxel having progressed on or after docetaxel,[27,28] supporting the divergent drug-resistance mechanisms of these two taxanes.

We cannot be certain that the paclitaxel response metagene is paclitaxel specific, despite the test for interaction, since silencing of *COL4A3BP* and *UGCG* has been shown to promote doxorubicin sensitisation.[9,29] Consistent with this hypothesis, we note that the paclitaxel response metagene is weakly predictive of pCR in the EORTC$_{FEC}$ dataset across all patients (AUC 0·62; 95% CI 0·49–0·72; p=0·025). The ceramide module is likely the main contributor, since *COL4A3BP* expression alone predicts for pCR in all patients (0·59; 0·48–0·71; p=0·061); this would support the role of the ceramide pathway in the regulation of multidrug sensitivity in vivo.

In summary, we used in-vitro functional genomics analyses to guide the development of a metagene to predict paclitaxel response in patients with breast cancer. Although we cannot conclude that our approach is better than associative predictive strategies, the latter strategy

requires training of two large cohorts and validation of such signatures. The functional genomics approach used in this study could be an efficient method to accelerate biomarker development for experimental therapies in single-cohort early phase clinical trials, where stratification of response according to tumour expression of a functional metagene could be considered.

References
1   Gralow JR, Burstein HJ, Wood W, et al. Preoperative therapy in invasive breast cancer: pathologic assessment and systemic therapy issues in operable disease. *J Clin Oncol* 2008; **26:** 814–19.
2   Liedtke C, Mazouni C, Hess KR, et al. Response to neo-adjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol* 2008; **26:** 1275–81.
3   Rouzier R, Perou CM, Symmans WF, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin Cancer Res* 2005; **11:** 5678–85.
4   Carey LA, Dees EC, Sawyer L, et al. The triple negative paradox: primary tumour chemosensitivity of breast cancer subtypes. *Clin Cancer Res* 2007; **13:** 2329–34.
5   Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005; **365:** 488–92.
6   Ein-Dor L, Zuk O, Domany E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. *Proc Natl Acad Sci USA* 2006; **103:** 5923–28.
7   Bonnefoi H, Potti A, Delorenzi M, et al. Validation of gene signatures that predict the response of breast cancer to neoadjuvant chemotherapy: a substudy of the EORTC 10994/BIG 00-01 clinical trial. *Lancet Oncol* 2007; **8:** 1071–78.
8   Farmer P, Bonnefoi H, Anderle P, et al. A stroma-related gene signature predicts resistance to neo-adjuvant chemotherapy in breast cancer. *Nat Med* 2009; **15:** 68–74.
9   Swanton C, Marani M, Pardo O, et al. Regulators of mitotic arrest and ceramide metabolism are determinants of sensitivity to paclitaxel and other chemotherapeutic drugs. *Cancer Cell* 2007; **11:** 498–512.
10   Liu YY, Han TY, Giuliano AE, Cabot MC. Ceramide glycosylation potentiates cellular multidrug resistance. *Faseb J* 2001; **15:** 719–30.
11   Gouaze V, Liu YY, Prickett CS, Yu JY, Giuliano AE, Cabot MC. Glucosylceramide synthase blockade down-regulates P-glycoprotein and resensitizes multidrug-resistant breast cancer cells to anticancer drugs. *Cancer Res* 2005; **65:** 3861–67.
12   Lee AJ, Kolesnick R, Swanton C. RNAi-mediated functional analysis of pathways influencing cancer cell drug resistance. *Expert Rev Mol Med* 2009; **11:** 15.
13   Segal E, Friedman N, Koller D, Regev A. A module map showing conditional activity of expression modules in cancer. *Nat Genet* 2004; **36:** 1090–98.
14   Hess KR, Anderson K, Symmans WF, et al. Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J Clin Oncol* 2006; **24:** 4236–44.
15   Silver DP, Richardson AL, Eklund AC, et al. Efficacy of neoadjuvant cisplatin in triple-negative breast cancer. *J Clin Oncol* 2010; published online Jan 25. DOI: 0: JCO.2009.22.4725v1.
16   Desmedt C, Piette F, Loi S, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin Cancer Res* 2007; **13:** 3207–14.
17   Pawitan Y, Bjohle J, Amler L, et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res* 2005; **7:** 953–64.
18   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:** 671–79.
19   Chin K, DeVries S, Fridlyand J, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell* 2006; **10:** 529–41.
20   Liedtke C, Hatzis C, Symmans WF, et al. Genomic grade index is associated with response to chemotherapy in patients with breast cancer. *J Clin Oncol* 2009; **27:** 3185–91.
21   Swanton C, Nicke B, Schuett M, et al. Chromosomal instability determines taxane response. *Proc Natl Acad Sci USA* 2009; **106:** 8671–76.
22   Turner N, Tutt A, Ashworth A. Hallmarks of 'BRCAness' in sporadic cancers. *Nat Rev Cancer* 2004; **4:** 814–19.
23   Quinn JE, Kennedy RD, Mullan PB, et al. BRCA1 functions as a differential modulator of chemotherapy-induced apoptosis. *Cancer Res* 2003; **63:** 6221–28.
24   Symmans WF, Ayers M, Clark EA, et al. Total RNA yield and microarray gene expression profiles from fine-needle aspiration biopsy and core-needle biopsy samples of breast carcinoma. *Cancer* 2003; **97:** 2960–71.
25   Gligorov J, Lotz JP. Preclinical pharmacology of the taxanes: implications of the differences. *Oncologist* 2004; **9** (suppl 2)**:** 3–8.
26   Haldar S, Basu A, Croce CM. Bcl2 is the guardian of microtubule integrity. *Cancer Res* 1997; **57:** 229–33.
27   Yonemori K, Katsumata N, Uno H, et al. Efficacy of weekly paclitaxel in patients with docetaxel-resistant metastatic breast cancer. *Breast Cancer Res Treat* 2005; **89:** 237–41.
28   Sawaki M, Ito Y, Hashimoto D, et al. Paclitaxel administered weekly in patients with docetaxel-resistant metastatic breast cancer: a single-center study. *Tumori* 2004; **90:** 36–39.
29   Liu YY, Han TY, Giuliano AE, Hansen N, Cabot MC. Uncoupling ceramide glycosylation by transfection of glucosylceramide synthase antisense reverses adriamycin resistance. *J Biol Chem* 2000; **275:** 7138–43.