BioChip J. (2011) 5(1): 32-38
DOI 10.1007/s13206-011-5106-3

Original Research

Protein profiling predicts the response to anthracycline and taxanes based neo-adjuvant chemotherapy in breast cancer

Shu Wang[1], Houpu Yang[1], Jiajia Guo[1], Miao Liu[1], Fuzhong Tong[1], Yingming Cao[1], Bo Zhou[1], Peng Liu[1], Hongjun Liu[1], Lin Cheng[1], Fei Xie[1], Deqi Yang[1] & Jiaqing Zhang[1]

Received: 7 October 2010 / Accepted: 29 November 2010 / Published online: 20 March 2011

Abstract Neo-adjuvant chemotherapy for breast cancer substantially benefits patients who achieve pathological response. However, clinical or pathological response information can only be obtained a period of time after chemotherapy. The identification of novel bio-markers or the application of new technique that can be used to predict treatment response before chemotherapy would allow therapy to be tailored on an individual patient basis. The purpose of this study is to identify the chemo-sensitivity and chemo-resistance related proteins using antibody microarray profiling, and to develop a multi-protein predictive model for breast cancer. Total protein was extracted from core needle biopsy samples obtained from 15 patients before treatment with neo-adjuvant TA (combination of taxanes and anthracycline) chemotherapy. Protein profiling was analyzed by antibody microarray. 10 patients were used as training set to develop the predictive model using the software PAM (prediction analysis of microarray). Another 5 patients were used as a validation set to test the model. In cross-validation, the molecular predictive model showed an accuracy of 90%, in independent validation, the model classified the cases with an accuracy of 80%. In conclusion, the proteomic predictive model has the potential to predict pathological response to neo-adjuvant TA chemotherapy.

Keywords: Breast cancer, Proteomic profiling, Antibody microarray, Neo-adjuvant chemotherapy, Predictive model

[1]Breast Disease Center, Peking University, People's Hospital, Beijing 100044, China
Correspondence and requests for materials should be addressed to S. Wang (✉wangshu@pkuph.edu.cn)

Introduction

Breast cancer is the most common female malignant tumor[1]. Systemic chemotherapy substantially decreases the risk of recurrence and death. Administration of chemotherapy before surgery defined as neo-adjuvant chemotherapy provides a unique opportunity for direct assessment of response to chemotherapy by monitoring changes in tumor size and histological changes in the first few months of treatment. Recent studies suggest neo-adjuvant chemotherapy benefits patients with locally advanced breast cancer who achieve pathological response[2], and pathological response to neo-adjuvant chemotherapy is a valid surrogate marker of long-term survival[3].

However, either clinical or pathological response evaluation could only be obtained a period of time after chemotherapy, we are unable to distinguish who would be more sensitive or more resistant than others in advance. Even though primary drug resistance will result in treatment failure in some patients, all patients eligible for neo-adjuvant chemotherapy empirically receive the same treatment in practice. This consideration has led to the hypothesis that identifying predictors for chemotherapy would be of value to decrease non-effective treatment and related toxic effects.

Some biological markers such as HER-2, p53, Topo II, Ki67 *et al*. have been evaluated as predictors of response to specific regimens, but the results are conflicting[4-9]. It has been reported that HER2 overexpression[4] or Ki67[6] is associated with higher rate of pCR to preoperative chemotherapy, co-overexpression of Topo II and HER-2 is related to anthracycline sensitivity, patients with p53-deficient tumors may benefit from

paclitaxel[8], and other studies showed that overexpression or amplification of HER2 or Topo II were not predictive of response[5,9]. Up to now, none of a single candidate marker could provide sufficient evidence to be transferred to clinical practice[10].

The improvement of high-throughput molecular technologies have caused a fundamental transformation in breast cancer research. Ayers *et al*. identified a set of key genetic markers (74 genes) that predicts response to paclitaxel followed by 5-fluorouracil, doxorubicin, and cyclophosphamide regimen with an accuracy of 75%[11]. Chang *et al*. identified a 92-genes expression pattern predicting response to docetaxel by DNA microarray analysis with an accuracy of 88%[12]. These small studies have shown that multiple gene profiling give more detailed portrait of breast cancer and should be more powerful to untangle the heterogeneity of breast cancer[11-17]. Since proteins are the actual executors of biologic functions and may describe the cancer characteristics more essentially than genes, direct analysis of protein profiling represents a complementary approach to gene profiling[18,19]. Two-dimensional gel electrophoresis, in combination with mass spectrometry, is a state-of-the-art method for detecting global variation of protein expression[18]. However, the technique has limitation to detect low abundance proteins and is not ready for high-throughput applications. In terms of easy use and rapid quantitative data generation, antibody microarray offering several advantages over traditional proteomic technologies, has successfully been applied for the discovery of potential biomarkers. Smith and colleagues[19] utilized an antibody microarray kit to analyze 224 proteins in MDA-MB-231 breast cancer cell line and a novel derivative displaying significant resistance to doxorubicin, and found decreased expression of p-ERK, CK18 and cyclinB1 was associated with doxorubicin resistance.

Currently, anthracyclines and taxanes are most widely used agents in neo-adjuvant chemotherapy in breast cancer[20], and rare research has investigated in vivo application of proteomic profiling in neo-adjuvant chemo-sensitivity prediction so far. Our study was designed to evaluate the feasibility of developing a multi-protein model by antibody microarray to predict the response to anthracyclines and taxanes based (TA) regimen in patients with primary breast cancer.

Results and Discussion

Development of multi-protein predictor

Antibody microarray is not widely used in neo-adjuvant chemotherapy prediction, one of the possible reason is that researchers are not sure if sufficient proteins could be obtained by core needle biopsy. In our study, all 15 samples provided sufficient protein (808-3,716 μg) for antibody microarray analysis, and therefore the method seems feasible for clinical application. Five sensitive and five resistant tumors were included in the discovery set, and another two sensitive tumors, two resistant tumors, and a moderately sensitive tumor were included in the independent validation set (Table 1). PAM was used to identify a minimal subset of proteins that succinctly characterized sensitive and resistant tumors, by using a threshold of 1.2 (Figure 1), a set of 96 proteins were selected.

If we compared the 96 protein list with previous

Table 1. Characteristics of the patients.

No.	Age	Menopausal status	Tumor size (cm × cm)	LN	ER	PR	HER-2	M&P Grade	Group
1	80	Post	3.5 × .3.5	+	+	+	−	1	Training
2	74	Post	2 × 2	+	−	+	++	1	Training
3	70	Post	4 × 4	+	+	+	−	1	Training
4	82	Post	2 × 3	+	+	+	−	1	Training
5	47	Pre	7.5 × 4	+	+	+	−	2	Training
6	52	Post	4 × 4	−	+	+	+++	4	Training
7	46	Pre	6 × 5	−	+	+	++	4	Training
8	50	Pre	4 × 2	−	+	+	−	4	Training
9	53	Post	3 × 2	−	−	−	UA	5	Training
10	45	Pre	3 × 2	−	−	+	+++	5	Training
11	71	Post	5 × 3	−	+	+	−	2	Validation
12	52	Post	3 × 3	−	−	+	+++	2	Validation
13	51	Post	2.5 × 2.5	+	−	−	+++	3	Validation
14	45	Pre	8 × 8	+	+	+	−	4	Validation
15	50	Pre	5 × 5	+	−	−	+++	5	Validation

LN, Lymph node; UA, Unavailable; +, Positive; −, Negative; M&P, Miller & Payne; ER, Estrogen Receptor; PR, Progesterone Receptor; HER-2, Human Epidermal Growth Factor Receptor 2.

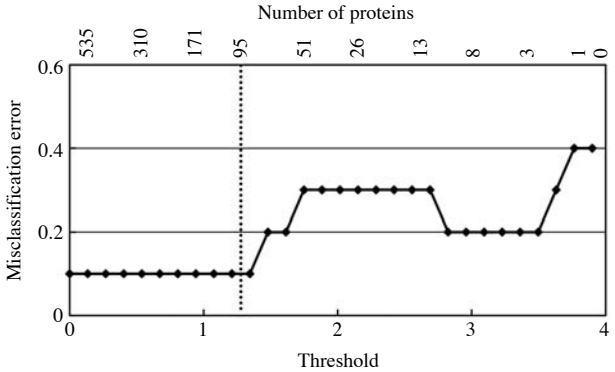

Figure 1. Cross validation error is shown as a function of the threshold parameter. The threshold value 1.2 is chosen and yields a subset of 96 selected proteins.

gene profiling results, there is little consistency in the specific molecular identities. There are two possible explanations. Firstly, our antibody microarray included 656 probes detecting known cancer related proteins, the detecting scale is much smaller than whole genome screening; secondly, the criteria of chemotherapy response is different. Most of the other studies generate the model comparing PCR and non-PCR, but our study generates the model comparing grade 1+2 and grade 4+5.

Hierarchical clustering analysis

The unsupervised hierarchical clustering analysis, based on 96 proteins profiling, generated a tree with a clear distinction of samples in two main groups, represented by sensitive tumors and resistant tumors (Figure 2). It seems that antibody microarray analysis is a promising strategy for chemotherapy response prediction.

Although we do not pay much attention to single markers, the classifier protein list indicates some possible mechanism of chemo-sensitivity. E.g. Topo II is a key enzyme in DNA replication and the target of anthracyclines. TA sensitive tumors have higher expressions of Topo II, which lends support to the idea that Topo II is a sensitive marker in anthracycline-containing regimen[21]. Metalmatrixproteinase is a key enzyme in metabolism of matrix. TA resistant tumors seem to have higher expressions of metalmatrixproteinases. This result is consistent with Liu's report in non-small cell lung cancer[22], suggesting tumor microenvironment could be of value in response to therapeutics. Furthermore, there are some proteins such as CD20 that are rarely studied in breast cancer. It is interesting that in the protein clustering tree, it is very close to MMP7. This gives us a clue that the potential relationship between leukocyte differentiation antigen family and metalmatrixproteinase family should be validated.

Cross validation and independent validation

In cross-validation (n=10), the 96-protein model classified the 10 training cases correctly with an accuracy of 90% (Figure 2), four sensitive and 5 resistant cases were classified correctly with probabilities at 76% to 100%, one resistant case was incorrectly classified as sensitive. In independent validation (n=5), the model classified the cases with an accuracy of 80%, two sensitive and two resistant cases were classified correctly with probabilities at 99% to 100%. The additional moderate sensitive case, which was evaluated as Miller & Payne 3, was classified as resistant (Figure 3).

Although there is little overlap in specific molecules between our study and previous ones, all of these studies achieved satisfactory prediction accuracy. It seems that the specific marker is not as important as the molecular sets, we hypothesize that different gene or protein sets could track common biological characteristics which lead to similar prediction results.

The enrichment of Gene Ontology Term and pathway of the 96 discriminatory proteins was analyzed by software MAS. As shown in Table 2, these 96 differentially expressed proteins mainly played a role in protein binding, metal ion binding, ATP binding, transcription factor activity, *et al*. As shown in Table 3, they were primarily involved in cell cycle, focal adhesion, neuroactive ligand-receptor interaction, MAPK signaling pathway and ECM-receptor interaction *et al*. These functions have been widely reported to have close relationship with chemotherapy sensitivity.

When we went over the literatures, we found the same situation also happened in gene profiling predicting models for relapse free survival and overall survival. Fan *et al*.[23] investigated five prognosis models for breast cancer (intrinsic subtypes, 70-gene profile, wound response, recurrence score, and the two-gene ratio) with little overlap in terms of gene identity in a single data set of 295 samples. Even though different gene sets were used for prognostication, four of the five models showed significant agreement in the outcome predictions for individual patients and are probably tracking a common set of biologic phenotypes.

Conclusions

In summary, we identified candidate proteins linked to breast cancer chemo-sensitivity and developed a multi-protein sensitivity predicting model for TA regimen by antibody microarray. Even though the number

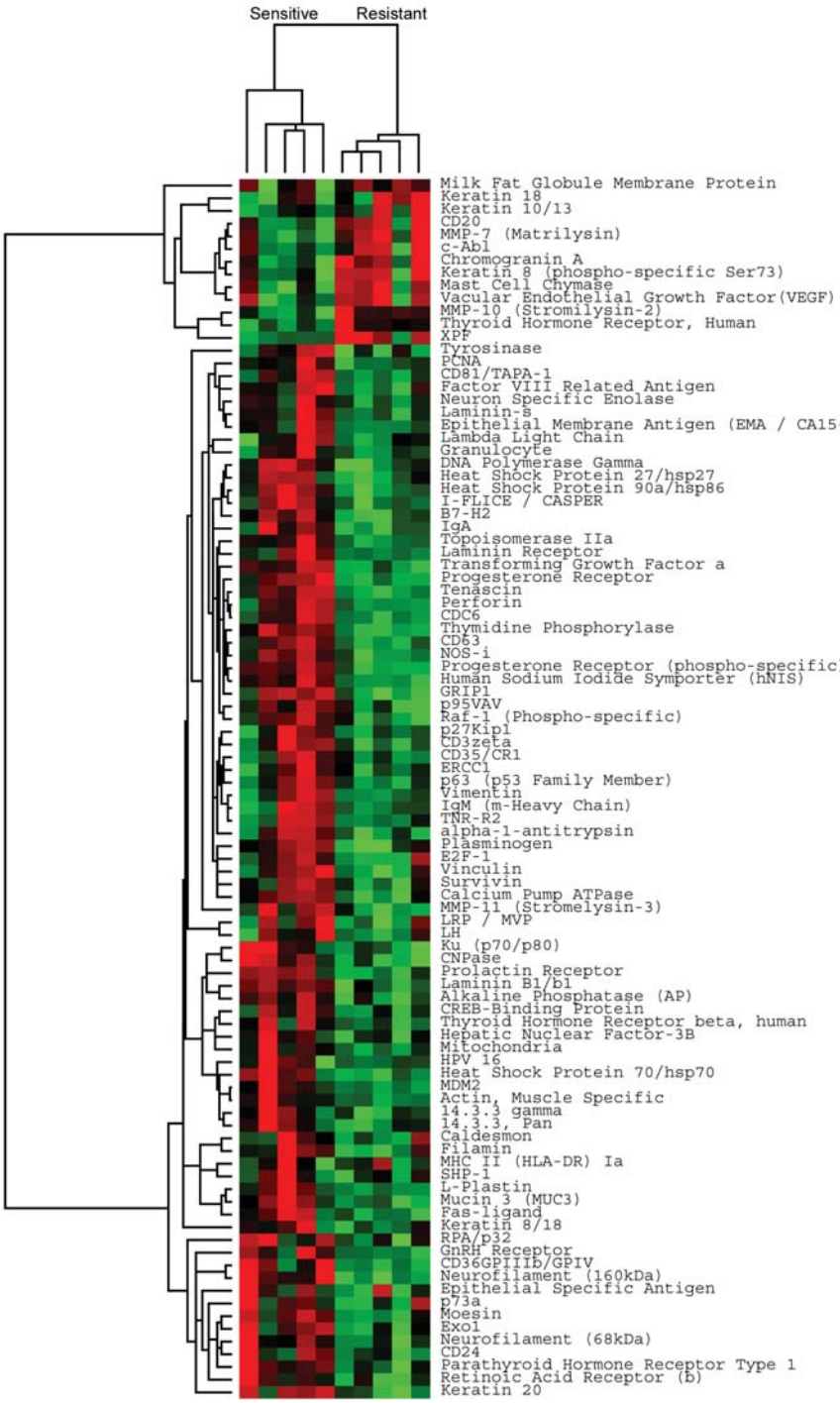

Figure 2. Heat map of the chosen 96 proteins. Within each of the horizontal partitions, we have ordered the proteins by hierarchical clustering, and similarly for the samples within each vertical partition.

of patients included in this study was relatively small, the predicting result seems promising. Further larger patient cohorts are needed to optimize the model. We hope chemotherapy could be carried on an individual patient basis with more efficiency, less cost and toxicity.

Materials and Methods

Chemotherapy protocol

This study was conducted at Peking University People's Hospital Breast Disease Center. From April 2007 to September 2007, 15 patients with locally advanced

breast cancer were recruited. All the clinical backgrounds of the patients are presented in Table 1. Samples were obtained by core needle biopsy (16G, Bard Ltd, USA) from the primary tumors before chemotherapy. Half of each tumor sample was subjected to histological observation to confirm the precise sampling of the cancer tissues in each patient. The other half was snap frozen in liquid nitrogen and kept at −70°C until extraction use. All patients received 4 cycles of TA chemotherapy (paclitaxel 175 mg/m^2 iv in 3 hours d1, epirubicin 60 mg/m^2 or perarubicin 45 mg/m^2 d1, cycled every 21 days). Every patient underwent modified radical mastectomy or breast conserving surgery after completion of chemotherapy. The study was approved by the Ethics Committee of Peking University People's Hospital.

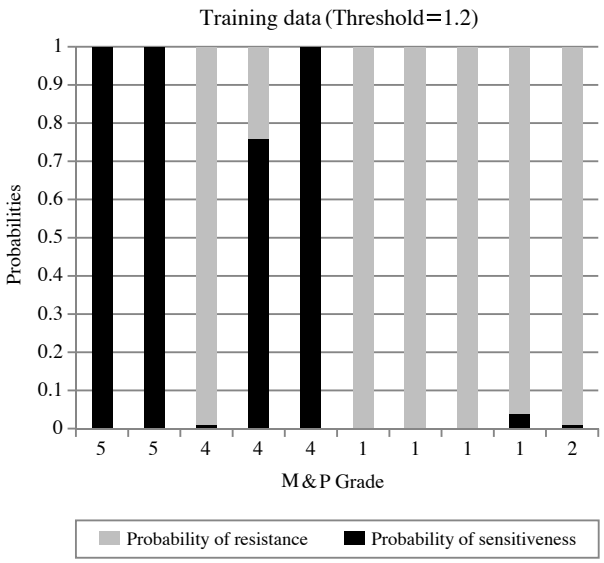

Figure 3. Estimated probabilities for the training data. 9 out of 10 in the training samples are correctly classified.

Evaluation of the pathological response

Miller & Payne histological grading system[3] which focuses on the reduction in tumor cellularity was introduced to assess the histological response.

Grade 1: No change or no reduction in overall cellularity;
Grade 2: A minor loss of tumor cells but overall cellularity still high or up to 30% loss;
Grade 3: Between an estimated 30% and 90% reduction in tumor cells;
Grade 4: more than 90% reduction of tumor cells;

Table 2. Enrichment of proteins in different funtions.

Gene ontology term	Enrichment number	Protein name
protein binding	31	ACTA1;SERPINA1;CD24;CD81;CDC6;POLG;E2F1;EXO1;GRIP1; CFLAR;IGHA1;KRT10;KRT18;KRT20;KRT8;LAMB1;LAMB2;MVP; IMMT;CDKN1B;TP73;PCNA;PRLR;RAF1;RPA2;PTPN6;THRB; THRA;TNFRSF1B;TYR;VIM
zinc ion binding	15	ALPL;CREBBP;MDM2;MMP10;MMP11;MMP7;NOS2;TP63;TP73; PGR;RAF1;RARB;BIRC5;THRB;THRA
protein homodimerization activity	11	CD247;VWF;FLNA;HSP90AA1;NOS2;PRLR;BIRC5;THRB;TOP2A; TYR;VEGFA
metal ion binding	11	CREBBP;MDM2;TP63;TP73;PGR;RAF1;RARB;BIRC5;THRB; THRA;TYR
calcium ion binding	9	ATP2C1;CHGA;LCP1;MMP10;MMP11;MMP7;NOS2;PRF1;PLG
transcription factor activity	8	CREBBP;E2F1;TP63;TP73;PGR;RARB;THRB;THRA
ATP binding	7	ACTA1;ATP2C1;CDC6;HSPA1B;HSP90AA1;RAF1;TOP2A
nucleotide binding	7	ACTA1;ATP2C1;CDC6;HSPA1B;HSP90AA1;RAF1;TOP2A
hydrolase activity	7	ALPL;ATP2C1;CNP;ERCC1;EXO1;PTPN6;ERCC4
structural constituent of cytoskeleton	6	ACTA1;KRT20;MSN;NEFM;NEFL;VIM
sequence-specific DNA binding	6	ERCC1;FOXA2;PGR;RARB;THRB;THRA
peptidase activity	6	SERPINA1;CMA1;MMP10;MMP11;MMP7;PLG
protein C-terminus binding	5	ERCC1;XRCC6;NEFL;TOP2A;ERCC4
magnesium ion binding	5	ALPL;ATP2C1;POLG;ENO2;ERCC4
identical protein binding	5	HSPB1;LCP1;MDM2;NEFL;TP63
receptor activity	5	MS4A1;CR1;GNRHR;PRLR;TNFRSF1B
signal transducer activity	5	ATP2C1;CD24;CREBBP;FLNA;TGFA
steroid hormone receptor activity	4	PGR;RARB;THRB;THRA
protein N-terminus binding	4	YWHAQ;VWF;PGR;ERCC4
structural molecule activity	4	KRT18;KRT8;LAMB2;VCL
DNA binding	4	POLG;EXO1;XRCC6;PCNA

Table 3. Enrichment of proteins in different pathways.

KEGG pathway	Enrichment number	Protein name
Cell cycle	9	YWHAG;YWHAQ;ABL1;CDC6;CREBBP;E2F1;MDM2;CDKN1B;PCNA
Focal adhesion	8	VWF;FLNA;LAMB1;LAMB2;RAF1;TNC;VEGFA;VCL
Neuroactive ligand-receptor interaction	6	GNRHR;PTH1R;PLG;PRLR;THRB;THRA
MAPK signaling pathway	6	FASLG;FLNA;HSPB1;HSPA1B;HSPA1A;RAF1
ECM-receptor interaction	5	CD36;VWF;LAMB1;LAMB2;TNC
Natural killer cell mediated cytotoxicity	5	CD247;FASLG;PRF1;RAF1;PTPN6
Nucleotide excision repair	4	ERCC1;PCNA;RPA2;ERCC4
Complement and coagulation cascades	4	SERPINA1;CR1;VWF;PLG
ErbB signaling pathway	4	ABL1;CDKN1B;RAF1;TGFA
Hematopoietic cell lineage	4	MS4A1;CR1;CD36;HLA-DRB1
Antigen processing and presentation	4	HSPA1B;HSPA1A;HSP90AA1;HLA-DRB1
Cytokine-cytokine receptor interaction	4	FASLG;PRLR;TNFRSF1B;VEGFA

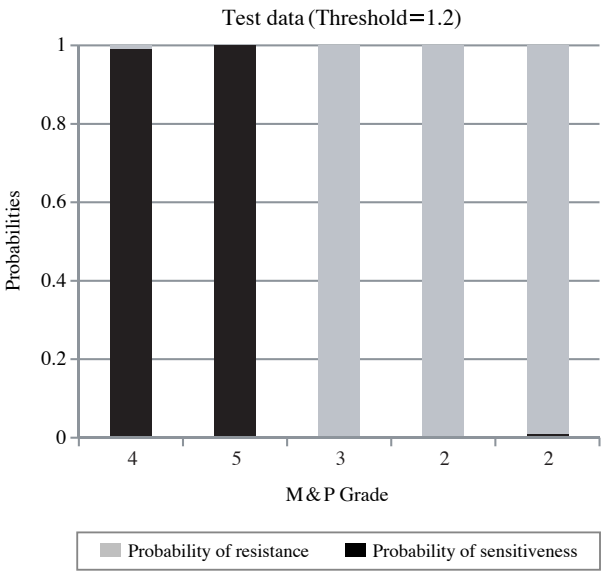

Figure 4. Estimated probabilities for the test data. 4 out of 5 in the test samples are correctly classified. One sample whose actual M&P grade is 3 is classified as resistant.

Grade 5: No malignant cells identifiable and DCIS may be present.

In this study, Grade 1 and Grade 2 were defined as resistant to chemotherapy, while Grade 4 and Grade 5 were defined as sensitive to chemotherapy. Additionally, Grade 3 was defined as moderately sensitive to chemotherapy.

Antibody microarray hybridization

Protein was extracted from a pretreatment biopsy sample using tissue protein extraction reagent (Kangchen, Shanghai, China). The protein amount and quality was assessed with BCA Protein Assay Kit (Kangchen, Shanghai, China). The total protein was labeled by biotin (biotin:protein 1 : 7). Biotin labeled protein (200 μg) was hybridized to an antibody microarray (Springbio Master-656, Springbio Ltd, USA) for 2 hours. The antibody microarray contains a unique set of 656 cancer relevant markers, such as progesterone receptor related to breast cancer, neurofilament related to tumor signal transduction, and so on. Each slide is printed with these probes in duplicates. After the hybridized slides incubated with 1,000 μL cy3-streptavidin for 2 hours, TIFF images with fluorescent signals were generated through scanning the slides with a GenePix 4000B microarray scanner (excitation wavelength at 532 nm, Axon Ltd, USA). The fluorescence images were analyzed using Genepix Pro 6.0.

Statistical analysis

We averaged the fluorescence of the duplicates of each protein, and then deducted the background fluorescence. After eliminating proteins that were not present in all of the sample, we calculated the protein/tubulin fluorescence ratio as the relative expression level to be analyzed. The protein profiling data was exported to Excel version 2003 for further filtering and analysis. The first 10 patients were used as "training set" for predictive marker discovery, and 5 additional patients were used as the independent "validation set".

We developed the classifier using the PAM analysis package (Prediction Analysis of Microarray, Stanford University, http://www-stat.stanford.edu/tibs/PAM). The developers of the package[24] devised an approach to cancer class prediction from expression profiling based on an enhancement of the simple nearest prototype (centroid) classifier. They shrunk the prototypes and hence obtained a classifier that was often more accurate than competing methods. The method of

"nearest shrunken centroids" identified subsets of genes or proteins that best characterized each class. We used the tool to shrink the proteins to an optimized subset, then conducted clustering analysis with Cluster 3.0 (Stanford University), and visualized the results with Java Treeview 1.1.3 (AIOK). The PAM package estimated the class probabilities to determine the neoadjuvant chemotherapy response classification for the discovery and validation cases to accomplish the cross and independent validations.

For discriminatory protein function analysis, a web-based molecular annotation tool package, Microarray Annotation System (MAS, http://bioinfo.capitalbio.com/mas) was applied.

Acknowledgements This project was supported by National Science Foundation of China (NSFC 30600742).

References

1. Parkin, D.M., Bray, F., Ferlay, J. & Pisani, P. Global cancer statistics, 2002. *CA. Cancer J. Clin.* **55**, 74-108 (2005).
2. Fisher, E.R. *et al.* Pathobiology of preoperative chemotherapy: findings from the National Surgical Adjuvant Breast and Bowel (NSABP) protocol B-18. *Cancer* **95**, 681-695 (2002).
3. Ogston, K.N. *et al.* A new histological grading system to assess response of breast cancers to primary chemotherapy: prognostic significance and survival. *Breast* **12**, 320-327 (2003).
4. Andre, F. *et al.* HER2 expression and efficacy of preoperative paclitaxel/FAC chemotherapy in breast cancer. *Breast Cancer Res. Treat.* **108**, 183-190 (2008).
5. Zhang, F. *et al.* Correlation between HER-2 expression and response to neoadjuvant chemotherapy with 5-fluorouracil, doxorubicin, and cyclophosphamide in patients with breast carcinoma. *Cancer* **97**, 1758-1765 (2003).
6. Bozzetti, C. *et al.* Evaluation of HER-2/Neu amplification and other biological markers as predictors of response to neoadjuvant anthracycline-based chemotherapy in primary breast cancer: the role of anthracycline dose intensity. *Am. J. Clin. Oncol.* **29**, 171-177 (2006).
7. Colleoni, M. *et al.* Expression of ER, PgR, HER1, HER2, and response: a study of preoperative chemotherapy. *Ann. Oncol.* **19**, 465-472 (2008).
8. Kandioler-Eckersberger, D. *et al.* TP53 Mutation and p53 overexpression for prediction of response to neoadjuvant treatment in breast cancer patients. *Clin. Cancer Res.* **6**, 50-56 (2000).
9. Petit, T. *et al.* Comparative value of tumour grade, hormonal receptors, Ki-67, HER-2 and topoisomerase II alpha status as predictive markers in breast cancer patients treated with neoadjuvant anthracycline-based chemotherapy. *Eur. J. Cancer* **40**, 205-211 (2004).
10. Harris, L. *et al.* American society of clinical oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J. Clin. Oncol.* **25**, 5287-5312 (2007).
11. Ayers, M. *et al.* Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J. Clin. Oncol.* **22**, 2284-2293 (2004).
12. Chang, J.C. *et al.* Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* **362**, 362-369 (2003).
13. Dressman, H.K. *et al.* Gene expression profiles of multiple breast cancer phenotypes and response to neoadjuvant chemotherapy. *Clin. Cancer Res.* **12**, 819-826 (2006).
14. Folgueira, K. *et al.* Gene expression profile associated with response to doxorubicin-based therapy in breast cancer. *Clin. Cancer Res.* **11**, 7434-7443 (2005).
15. Gianni, L. *et al.* Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. *J. Clin. Oncol.* **23**, 7265-7277 (2005).
16. Hess, K.R. *et al.* Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J. Clin. Oncol.* **24**, 4236-4244 (2006).
17. Iwao-Koizumi, K. *et al.* Prediction of docetaxel response in human breast cancer by gene expression profiling. *J. Clin. Oncol.* **23**, 422-431 (2005).
18. Smith, L. *et al.* The proteomic analysis of cisplatin resistance in breast cancer cells. *Oncol. Res.* **16**, 497-506 (2007).
19. Smith, L. *et al.* The analysis of doxorubicin resistance in human breast cancer cells using antibody microarrays. *Mol. Cancer Ther.* **5**, 2115-2120 (2006).
20. Rastogi, P. *et al.* Preoperative chemotherapy: updates of national surgical adjuvant breast and bowel project protocols B-18 and B-27. *J. Clin. Oncol.* **26**, 778-785 (2008).
21. Coon, J.S. *et al.* Amplification and overexpression of topoisomerase IIα predict response to anthracycline-based therapy in locally advanced breast cancer. *Clin. Cancer Res.* **8**, 1061-1067 (2002).
22. Liu, H. *et al.* Predictive value of MMP-7 expression for response to chemotherapy and survival in patients with non-small cell lung cancer. *Cancer Sci.* **99**, 2185-2192 (2008).
23. Fan, C. *et al.* Concordance among gene-expression-based predictors for breast cancer. *N. Engl. J. Med.* **355**, 560-569 (2006).
24. Tibshirani, R., Hastie, T., Narasimhan, B. & Chu, G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc. Natl. Acad. Sci. USA* **99**, 6567-6572 (2002).