

Gene expression profiles do not consistently predict the clinical treatment response in locally advanced breast cancer

Therese Sørlie,¹ Charles M. Perou,³ Cheng Fan,³ Stephanie Geisler,⁵ Turid Aas,⁵ Andrew Nobel,⁴ Gun Anker,⁵ Lars A. Akslen,⁶ David Botstein,⁷ Anne-Lise Børresen-Dale,^{1,2} and Per Eystein Lønning⁵

¹Department of Genetics, Institute for Cancer Research, Rikshospitalet-Radiumhospitalet Medical Center; ²Medical Faculty, University of Oslo, Oslo, Norway; ³Lineberger Comprehensive Cancer Center and Departments of Genetics and Pathology and Laboratory Medicine and ⁴Department of Statistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; ⁵Department of Medicine, Section of Oncology, Haukeland University Hospital; ⁶The Gade Institute, Section for Pathology, University of Bergen, Bergen, Norway; and ⁷Department of Molecular Biology, 140 Carl Icahn Laboratory, Princeton University, Princeton, New Jersey

Abstract

Neoadjuvant treatment offers an opportunity to correlate molecular variables to treatment response and to explore mechanisms of drug resistance *in vivo*. Here, we present a statistical analysis of large-scale gene expression patterns and their relationship to response following neoadjuvant chemotherapy in locally advanced breast cancers. We analyzed cDNA expression data from 81 tumors from two patient series, one treated with doxorubicin alone (51) and the other treated with 5-fluorouracil and mitomycin (30), and both were previously studied for correlations between *TP53* status and response to therapy. We observed a low frequency of progressive disease within the luminal A subtype from both series (2 of 36 versus 13 of 45 patients; $P = 0.0089$) and a high frequency of progressive disease

among patients with luminal B type tumors treated with doxorubicin (5 of 8 patients; $P = 0.0078$); however, aside from these two observations, no other consistent associations between response to chemotherapy and tumor subtype were observed. These specific associations could possibly be explained by covariance with *TP53* mutation status, which also correlated with tumor subtype. Using supervised analysis, we could not uncover a gene profile that could reliably (> 70% accuracy and specificity) predict response to either treatment regimen. [Mol Cancer Ther 2006;5(11):2914–8]

Introduction

Resistance to cytotoxic compounds is a main reason for therapy failure in most malignancies, including breast cancer. *In vitro* experiments as well as studies in animal models have shown that mutations in the *TP53* gene are associated with chemoresistance (1). Molecular studies of tumors from patients treated with neoadjuvant chemotherapy using either doxorubicin monotherapy or 5-fluorouracil and mitomycin (FUMI) in concert revealed that *TP53* mutations affecting the DNA-binding domain of the protein correlate with drug resistance (2–4). However, neither in these tumors nor in the studies reported by others (5) did mutations in *TP53* unequivocally predict drug resistance, suggesting that other interactions and genes must be involved (6).

By subjecting the same tumors characterized for *TP53* mutations in relation to chemotherapy response to DNA microarray analysis, we were able to classify tumors into five distinct subtypes based on their gene expression patterns (7). This classification showed prognostic effect with respect to relapse-free as well as overall survival in our cohort (8) and also in series of patients examined by other investigators (9). The prognostic significance of gene expression profiles has been well documented with respect to breast cancer (10–13) as well as other malignancies (14–17). Although these findings confirm the biological relevance of such genomic analyses, a prognostic factor provides no specific information about responsiveness to specific treatments and should be distinguished from a “predictive factor” (18, 19). Knowledge about the value of genome-wide expression analyses in predicting treatment response in breast cancer has resulted in at least to two studies correlating gene expression profiles with sensitivity to taxane monotherapy (20, 21) and three studies (22–24) reporting sensitivity to anthracycline combination regimens containing either cyclophosphamide or a taxane. However, the predictive powers achieved in any of these studies do not allow clinical implementation without further evaluations.

Received 3/7/06; revised 6/4/06; accepted 9/25/06.

Grant support: Norwegian Cancer Society, Norwegian Research Council, “SalusAnsvar” Award (A-L. Børresen-Dale), National Cancer Institute Breast Specialized Program of Research Excellence program grant P50-CA58223-09A1, and Breast Cancer Research Foundation (C.M. Perou).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: The analyses of the data presented in this report are original work of the authors and have not been presented previously. Some molecular data on these cases have been used in two previous studies (see refs. 3 and 4), and the raw microarray data have been published in a different context (see ref. 9).

Requests for reprints: Per Eystein Lønning, Department of Medicine, Section of Oncology, Haukeland University Hospital, N-5021 Bergen, Norway. Phone: 47-55975000; Fax: 47-55973599. E-mail: per.lonning@helse-bergen.no

Copyright © 2006 American Association for Cancer Research.

doi:10.1158/1535-7163.MCT-06-0126

The aim of this study was to examine the potential of gene expression profiles as predictive factors of drug sensitivity in two uniformly treated breast cancer cohorts previously characterized for the predictive value of *TP53* mutations and for the prognostic importance of gene expression profiles. Similar to findings by others, we found gene expression profiles defined by response-guided supervised analysis to be limited with respect to predicting therapy response.

Materials and Methods

Patient and Treatment Information

The patients included in this study were part of two prospective studies evaluating predictive factors for response to chemotherapy in locally advanced breast cancer (T₃/T₄ and/or N₂). From one (doxorubicin series), we analyzed tumor samples from a subgroup consisting of 51 patients who were treated with doxorubicin monotherapy weekly in the neoadjuvant setting, scheduled for 16 weeks with 4 weekly assessments of clinical response (3). From the second similar study (FUMI series), we analyzed tumors from 30 patients who were treated with FUMI at 3-week intervals (4). Because these protocols were applied before implementation of the "Response Evaluation Criteria in Solid Tumors" criteria (25), for both studies, the response rates were classified according to the International Union Against Cancer criteria (26). Thus, responses were classified as partial response (reduction >50% in the sum of all tumor lesions, calculated for each as the product of the largest diameter and the one perpendicular to it), progressive disease (increase in the diameter product of any individual tumor lesion by >25%), or stable disease (anything between partial response and progressive disease). Therapy was terminated immediately in case progressive disease was revealed. An overview of patient characteristics is shown in Table 1, and a complete listing of all tumors and experiments is available in Supplementary Table S1.⁸

Microarray Analysis

Gene expression data were collected using cDNA arrays produced at the Stanford Functional Genomics Facility.⁹ The procedures used, including RNA extraction, hybridization, and data processing, have been described previously (7, 8) and are available at the Stanford Genomics Breast Cancer Consortium Portal Web site.¹⁰ The common set of genes used for the doxorubicin series totaled ~8,000, whereas for the FUMI series this number was ~30,000 due to more recent production lots of cDNA microarrays. Specifically, for these analyses, the back-

Table 1. Clinical characteristics of patients included in this study

	Doxorubicin (51)	FUMI (30)
Response		
PD	8	7
SD	22	13
PR	21	10
Histology		
Ductal	46	26
Other	5	4
<i>TP53</i> mutations	24	16
Median age	62 (32–85)	64 (37–82)
Median OS (mo)	33 (7–92)	24 (3–54)
No. relapses	26	17

Abbreviations: PD, progressive disease; SD, stable disease; PR, partial response; OS, overall survival.

ground-subtracted, lowess-normalized (27) log₂ ratio (Cy5/Cy3) intensity values were first filtered to select genes that had a signal intensity of at least 30 units above background in both channels. Only genes that met these criteria in at least 70% of the total data set were included for subsequent analysis, which totaled 4,424 probes for the entire data set. Next, missing values were imputed using the k-nearest neighbor imputation algorithm (28). Gene annotation from each data set was translated to UniGene Cluster IDs using the SOURCE database (29). Multiple occurrences of a UniGene Cluster IDs were collapsed by the median value for that ID within an experiment set.

Statistical Analysis

Relationships between gene expression profiles and response to chemotherapy were analyzed using "nearest shrunken centroid classifier" [prediction analysis for microarrays (PAM); ref. 30]. In addition, several other supervised prediction methods were used: recursive sample classification and gene selection with SVM for microarray data (r-SVM; ref. 31), Random Forest by Salford Systems (32), a k-nearest neighbor classifier with either Euclidean distance or one-minus-Spearman correlation as the distance function, and a class nearest centroid metric with either Euclidean distance or one-minus-Spearman correlation as the distance function (See Supplementary data for a more detailed description of the various methods; ref. 33).⁸ As discussed elsewhere (19), the terms partial response and stable disease are pragmatic terms that describe a status of tumor "growth arrest" with or without a certain degree of macroscopic reduction in tumor size; the discrimination between the two may often be arbitrary. However, progressive disease tumors are distinctive and easily discriminated from the other groups; therefore, our primary statistical analyses aimed at comparing progressive disease tumors versus the others. Finally, because these experiments were done across many different production lots of microarrays, we attempted to correct for systematic array batch bias by using "distance-weighted discrimination" (34).

⁸ Supplementary material for this article is available at Molecular Cancer Therapeutics Online (<http://mct.aacrjournals.org/>).

⁹ <http://www.microarray.org/sfgf/jsp/home.jsp>

¹⁰ http://genome-www.stanford.edu/breast_cancer/

Results and Discussion

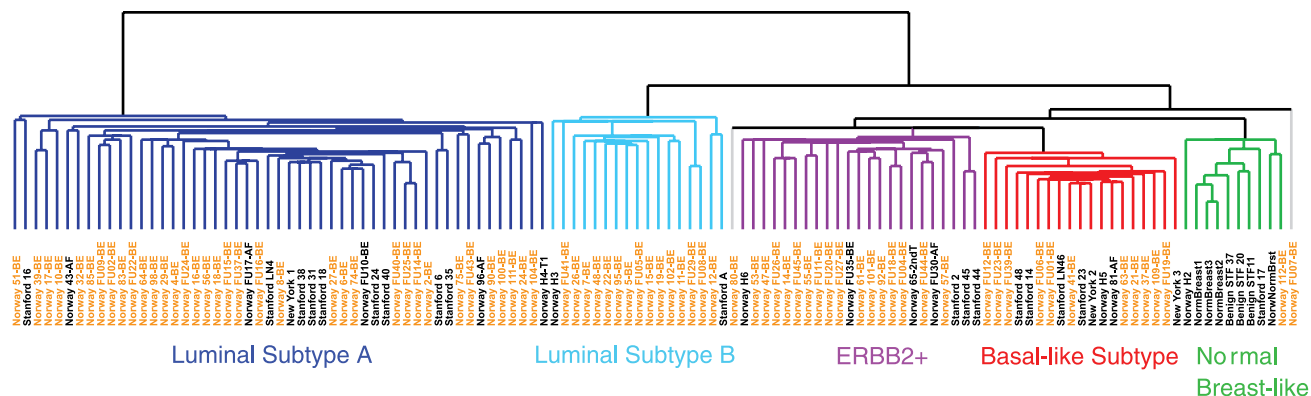
Prognostic versus Predictive Factors

Several gene expression-based classification schemes for various cancer types have emerged from DNA microarray studies over the last years. The biological importance of these classifications is highlighted by two significant observations: first, the possibility that, for each cancer type, including breast (9, 12, 35, 36), lung (15), lymphoma (14, 17), and head and neck (33), it is possible to classify individual tumors into groups characterized by distinct gene profiles, and second, the fact that these classifications provide prognostic information. These analyses, however, have thus far been of limited value for predicting therapeutic response in individual patients. A prognostic factor is traditionally associated with disease-free or overall survival (in the absence of systemic adjuvant therapy), whereas a predictive factor predicts response, or lack of, to a particular treatment (19). Although numerous prognostic factors have been identified in breast cancer, no predictive factors have been generally accepted thus far, with the exception of estrogen receptor- α and progesterone receptor

for endocrine therapy and HER2 for trastuzumab. In previous studies (3, 4), we found mutations in the *TP53* gene affecting the L2/L3 domains of the protein to be associated with nonresponse to treatment with doxorubicin or FUMI. However, such mutations were only predictive for nonresponse in 60% of the progressive disease tumors. Although our findings strongly advocate a role of the p53 pathway in response to these therapies in breast cancer, they also suggest that other genes must also be involved (6).

Response to Therapy across Different Molecular Subtypes

In this study, we conducted statistical analyses of gene expression data from altogether 81 tumors, which represents one of the largest studies to explore the predictive value of gene expression profiles in breast cancer (23, 37). Response to therapy in tumors across the different, previously defined tumor subgroups is depicted in Fig. 1. One of 25 tumors belonging to the luminal A subgroup versus 7 of 26 tumors in all other groups were non-responding (progressive disease) to doxorubicin ($P = 0.0496$, two-sided Fisher's exact test); for the patients



Doxo (51)				
PD ($p < 0.01$)	1/25	5/11	1/8 (+1uc)	0/5
SD	14/25	2/11	2/8 (+1uc)	3/5
PR	10/25	4/11	5/8 (+1uc)	2/5
FUMI (30)				
PD ($P = 0.654$)	1/11	1/4	2/8	2/6
SD	6/11	1/4	5/8	1/6
PR	4/11	2/4	1/8	3/6
<i>TP53</i> mut ($P < 0.001$)	6/36 (17%)	11/15 (73%)	13/16 (81%)	8/11 (75%)

Figure 1. Distribution of progressive disease (PD), stable disease (SD), and partial response (PR) tumors and *TP53* mutation frequencies across the breast tumor intrinsic subtypes for two different neoadjuvant treatment regimens: doxorubicin monotherapy (*Doxo*) and FUMI. Tumors are ordered according to the subtypes as presented in Fig. 1 in Sorlie et al. (9). *Orange*, tumor samples included in this study. *TP53* mutation status is shown as percentage of tumors from both series combined (the normal breast-like subtype is excluded). Note that one progressive disease tumor from the FUMI series (Norway FU07-BE) and one stable disease tumor from the doxorubicin series (Norway 80-BE) were unclassified (uc). Source: PNAS, July 8, 2003, vol. 100, no. 14, 8418-8423. Copyright (2003) National Academy of Sciences, U.S.A.

treated with FUMI, 1 of 11 of the luminal A tumors versus 6 of 19 tumors from the other groups experienced progressive disease ($P = 0.2146$; both data sets combined: $P = 0.0089$). Interestingly, both of the luminal A type tumors expressing chemoresistance were wild-type for *TP53*. Although a luminal B profile was associated with resistance to doxorubicin (five of eight progressive disease; $P = 0.0078$), a similar finding was not identified among the tumors treated with 5-fluorouracil/mitomycin. A differential variation in response across subtypes has also been shown in a recently published study on breast cancer treated with preoperative chemotherapy (35, 37). Our finding that progressive disease was a rare event among tumors expressing the luminal A gene profile is interesting; however, these tumors rarely harbor mutations in the *TP53* gene (6 of 36 versus 32 of 44 among the other tumors; $P < 0.0001$), so this may simply reflect *TP53* status among these tumors. Nevertheless, this points to the importance of considering the molecular heterogeneity of tumors when assessing predictive as well as prognostic markers.

Prediction of Therapeutic Response Classes Using Supervised Analyses

To search the gene expression data for patterns associated with response (progressive disease, stable disease, or partial response; no complete response was recorded in these two series) and to explore the feasibility of using such patterns as predictors, PAM was used on all tumor samples obtained before therapy for each of three treatment groups: doxorubicin and FUMI separately and combined.

Doxorubicin. Training a predictor for progressive disease versus partial response (for which significant differences in gene expression might be expected) resulted in overall accuracy of 70% but with only three of eight progressive disease tumors correctly classified. When combining the response groups partial response and stable disease into one class, training of a predictor resulted in a similar accuracy (73%), now with five of eight progressive disease tumors correctly classified.

5-Fluorouracil and Mitomycin. Prediction of progressive disease versus partial response showed an accuracy of 78% with five of seven progressive disease tumors correctly classified. Next, training a predictor for progressive disease versus the combined groups partial response/stable disease resulted in an accuracy of 63% with only two of the progressive disease tumors correctly classified by cross-validation.

Doxorubicin and FUMI Combined. Similar PAM analyses for the two series combined (81 patients of whom 15 experienced progressive disease) showed similar accuracy rates; 62% for progressive disease versus partial response (6 of 15 progressive disease correctly classified) and 62% for progressive disease versus the partial response/stable disease combined group, with 7 progressive disease tumors correctly classified.

In addition to PAM, several additional statistical methods were used to determine if the less than optimal prediction accuracies were due to a particular analysis method (i.e., PAM). In particular, Random Forest, which is a multitree

method for classification, and predictive modeling using a support vector machine method, termed r-SVM, which implements recursive gene ranking and selection steps, were both tested. All methods gave similar results, and thus, these results cannot be attributed to the statistical method used. A complete listing of the different methods and the prediction accuracies, sensitivities, and specificities resulting from the altogether seven prediction methods is presented in Supplementary Table S2.⁸ The values varied to some degree in magnitude, depending on the analysis method used and the sorting of the response groups. In particular, all predictors tended to do poorly in identifying the progressive disease tumors and often classified non-progressive disease samples correctly. This finding is a critical feature for the objective assessment of predictive profiles because, when a minor class is compared with a major class, a given "accurate" predictor could be developed that simply predicts most of the samples to be the major class.

The expression data used in this study were generated using several different production batches of cDNA arrays, and inconsistencies in such data that arose from process errors have been detected (38). Thus, we analyzed separately data from patients with progressive disease versus those with partial response using samples that had been hybridized on microarrays from the same batch only. Although this improved prediction accuracy up to 80%, only half of the progressive disease tumors were correctly classified (Supplementary Table S2).⁸ This finding of an inability to accurately identify most progressive disease tumors was true for all the predictors developed using the seven different methods, suggesting that this is inherent in the data and not due to the analysis method.

Conclusions

The aim of this study was to explore whether an analysis of gene expression data in a breast cancer cohort previously shown to yield prognostic gene profiles could identify gene signatures associated with response or resistance to chemotherapy. If so, this could add to the predictive value of *TP53* mutations previously reported in the same tumors. However, we could not identify a gene profile using multiple diverse supervised analysis methods, which was highly accurate at identifying either drug-sensitive or drug-resistant tumors. Molecular tumor subtype was modestly correlated with response with luminal A tumors showing a low rate of progressive disease and luminal B tumors treated with doxorubicin showing a high rate of progressive disease. In conclusion, we were not able to show that gene expression profiles can be used to accurately predict chemotherapy response in this data set. Similar to other studies (20–22, 24, 39), these results indicate that supervised analyses of relatively small sample sizes and with incomplete validation may not reveal a gene profile of sufficient predictive power to be of clinical use and suggest that genomic analyses using microarrays may need a different approach that incorporates functional hypotheses (40) to predict therapy sensitivity.

Acknowledgments

We thank the Stanford Microarray Database and the Stanford Functional Genomics Facility for contributing to this project.

References

- Wallace-Brodeur RR, Lowe SW. Clinical implications of p53 mutations. *Cell Mol Life Sci* 1999;55:64–75.
- Aas T, Borresen AL, Geisler S, et al. Specific P53 mutations are associated with *de novo* resistance to doxorubicin in breast cancer patients. *Nat Med* 1996;2:811–4.
- Geisler S, Lonning PE, Aas T, et al. Influence of TP53 gene alterations and c-erbB-2 expression on the response to treatment with doxorubicin in locally advanced breast cancer. *Cancer Res* 2001;61:2505–12.
- 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.
- Kandioler-Eckersberger D, Ludwig C, Rudas M, et al. TP53 mutation and p53 overexpression for prediction of response to neoadjuvant treatment in breast cancer patients. *Clin Cancer Res* 2000;6:50–6.
- Lonning PE. Genes causing inherited cancer as beacons to identify the mechanisms of chemoresistance. *Trends Mol Med* 2004;10:113–8.
- Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747–52.
- Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869–74.
- Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 2003;100:8418–23.
- Huang E, Cheng SH, Dressman H, et al. Gene expression predictors of breast cancer outcomes. *Lancet* 2003;361:1590–6.
- Ma XJ, Wang Z, Ryan PD, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell* 2004;5:607–16.
- 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.
- 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.
- Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling [see comments]. *Nature* 2000;403:503–11.
- Garber ME, Troyanskaya OG, Schluens K, et al. Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci U S A* 2001;98:13784–9.
- Pomeroy SL, Tamayo P, Gaasenbeek M, et al. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* 2002;415:436–42.
- Shipp MA, Ross KN, Tamayo P, et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med* 2002;8:68–74.
- Henderson IC, Patek AJ. The relationship between prognostic and predictive factors in the management of breast cancer. *Breast Cancer Res Treat* 1998;52:261–88.
- Lonning PE. Study of suboptimum treatment response: lessons from breast cancer. *Lancet Oncol* 2003;4:177–85.
- 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.
- Iwao-Koizumi K, Matoba R, Ueno N, et al. Prediction of docetaxel response in human breast cancer by gene expression profiling. *J Clin Oncol* 2005;23:422–31.
- Ayers M, Symmans WF, Stec J, et al. Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide chemotherapy in breast cancer. *J Clin Oncol* 2004;22:2284–93.
- Gianni L, Zambetti M, Clark K, 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* 2005;23:7265–77.
- Hannemann J, Oosterkamp HM, Bosch CA, et al. Changes in gene expression associated with response to neoadjuvant chemotherapy in breast cancer. *J Clin Oncol* 2005;23:3331–42.
- Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000;92:205–16.
- Hayward JL, Carbone PP, Heusen JC, Kumaoka S, Segaloff A, Rubens RD. Assessment of response to therapy in advanced breast cancer. *Br J Cancer* 1977;35:292–8.
- 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.
- Troyanskaya O, Cantor M, Sherlock G, et al. Missing value estimation methods for DNA microarrays. *Bioinformatics* 2001;17:520–5.
- Diehn M, Sherlock G, Binkley G, et al. SOURCE: a unified genomic resource of functional annotations, ontologies, and gene expression data. *Nucleic Acids Res* 2003;31:219–23.
- Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* 2002;99:6567–72.
- Zhang X, Wong WH. Recursive sample classification and gene selection based on SVM: method and software description. Technical report. Boston (MA): Department of Biostatistics, Harvard School of Public Health; 2001. p. 1–5. [<http://www.hsph.harvard.edu/bioinfocore/r-svm.pdf>].
- Breiman L. Random forests. Technical report. Berkeley (CA): Department of Statistics, University of California; 2001.
- Chung CH, Parker JS, Karaca G, et al. Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer Cell* 2004;5:489–500.
- Benito M, Parker J, Du Q, et al. Adjustment of systematic microarray data biases. *Bioinformatics* 2004;20:105–14.
- Bertucci F, Finetti P, Rougemont J, et al. Gene expression profiling identifies molecular subtypes of inflammatory breast cancer. *Cancer Res* 2005;65:2170–8.
- 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.
- Rouzier R, Perou CM, Symmans WF, et al. Breast cancer molecular subtypes respond differently to preoperative chemotherapy. *Clin Cancer Res* 2005;11:5678–85.
- Tu IP, Schaner M, Diehn M, et al. A method for detecting and correcting feature misidentification on expression microarrays. *BMC Genomics* 2004;5:64.
- Jansen MP, Foekens JA, van Staveren IL, et al. Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. *J Clin Oncol* 2005;23:732–40.
- Lonning PE, Sorlie T, Borresen-Dale A-L. Genomics in breast cancer—therapeutic implications. *Nat Clin Pract Oncol* 2005;2:26–33.