Letter

Expression profiling predicts outcome in breast cancer

Laura J van 't Veer[1,2], Hongyue Dai[3], Marc J van de Vijver[1], Yudong D He[3], Augustinus AM Hart[4], René Bernards[2,5] and Stephen H Friend[3]

[1]Division of Diagnostic Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[2]Center for Biomedical Genetics, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[3]Rosetta Inpharmatics, Inc.*, Kirkland, Washington, USA
[4]Division of Radiotherapy, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[5]Division of Molecular Carcinogenesis, The Netherlands Cancer Institute, Amsterdam, The Netherlands

Corresponding author: Laura J van 't Veer (e-mail: lveer@nki.nl)

Published: 4 December 2002

Breast Cancer Res 2003, **5**:57-58 (DOI 10.1186/bcr562)
© 2003 BioMed Central Ltd (Print ISSN 1465-5411; Online ISSN 1465-542X)

See commentary, http://breast-cancer-research.com/content/5/1/23

Gruvberger *et al*. postulate, in their commentary [1] published in this issue of *Breast Cancer Research*, that our "prognostic gene set may not be broadly applicable to other breast tumor cohorts", and they suggest that "it may be important to define prognostic expression profiles separately in estrogen receptor (ER) positive and negative tumors". This is based on two observations derived from our gene expression profiling data in breast cancer [2]: the overlap between reporter genes for prognosis and ER status, and Gruvberger *et al*.'s inability to confirm the prognosis prediction using a nonoptimal selection of 58 of our 231 prognosis reporter genes.

The overlap between our prognosis reporter genes and the ER status genes is certainly very large, mainly because ~10% of all genes on our microarray contain information on the ER status (2460 out of 24,479). However, the overlap between the 70 optimal prognosis genes and the 550 optimal ER status genes is only 17% (12 out of 70). We therefore believe that there is a different subset of genes that reports prognosis as compared with ER status. Our prognosis classifier strongly predicts the risk of distant metastases (odds ratio = 15, 95% confidence interval = 4–56, $P < 0.0001$). Adjusted for associations with known clinical prognosticators, including ER status, in a multivariable analysis, this odds ratio slightly increases to 18 (95% confidence interval = 3.3–94, $P = 0.0002$) [2]. This indicates that the predictive capacity of our prognosis classifier cannot be explained by its association with, among other factors, ER status as suggested.

There are a few points of concern relating to Gruvberger *et al*.'s inability to develop an outcome classifier on their data set. It should be pointed out that the prognosis signature is subtle both in the number of genes and the magnitude of differential regulation as compared with the ER status signature. The microarray platform sensitivity and reproducibility is a key issue in generating high quality data that ultimately determine the power of discovery. The reproducibility of our platform is clearly demonstrated by the uniform patterns related to ER status (Fig. 3b in [2]) as compared with the ER status expression patterns published by Gruvberger *et al*. (Fig. 2b in [3]). Our experiences with cDNA arrays and oligonucleotide arrays showed that the cDNA arrays are less sensitive in detecting small differential regulations because of the high chance of nonspecific binding.

Another point concerns the reference sample used in the two-color assay and its effect on the sensitivity. A nearby reference (a tumor sample pool in our case) makes the differences between the samples easily detectable, whereas a distant reference (one cell line [1]) makes this harder to detect since the fluctuations of the large offset interferes with the measurement of small differential signals. Moreover, we always repeat the same measurement twice in fluor-reversal pairs to minimise the potential labeling biases.

It is possible that the aforementioned reasons and the fact that most of our 70 optimal genes were not on their array

* A wholly-owned subsidiary of Merck & Co., Inc.

ER = estrogen receptor.

prevented Gruvberger *et al*. from revealing the prognosis signature.

Finally, it is of importance to note that we derived our prognosis classifier from breast cancer patients of whom the majority (93%) did not receive adjuvant systemic therapy, whereas all Gruvberger *et al*.'s patients received adjuvant tamoxifen treatment. So, a relatively better outcome (approximately 30%) can be expected within Gruvberger *et al*.'s ER-positive subgroup because of the tamoxifen treatment. The predictive power of our prognosis reporters may be reduced in an adjuvantly treated patient group.

A confirmation on a large unselected 'cohort' of breast cancer patients is required to validate our findings. We have recently completed DNA microarray analyses on a cohort of 295 breast cancer samples. The results of this will be published in the near future.

References

1. Gruvberger SK, Ringner M, Eden P, Borg A, Ferno M, Peterson C, Meltzer P: **Expression profiling to predict outcome in breast cancer: the influence of sample selection.** *Breast Cancer Res* 2002, **5**:23-26.
2. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, Peterse JL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH: **Gene expression profiling predicts clinical outcome of breast cancer.** *Nature* 2002, **415**:530-536.
3. Gruvberger S, Ringner M, Chen Y, Panavally S, Saal LH, Borg A, Ferno M, Peterson C, Meltzer PS: **Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns.** *Cancer Res* 2001, **61**:5979-5984.

Correspondence

Laura J van 't Veer, Division of Diagnostic Oncology, The Netherlands Cancer Institute, 121 Plesmanlaan, 1066 CX Amsterdam.
E-mail: lveer@nki.nl