

## Short communication

# Gene expression profiling of breast cancer

Lajos Pusztai

Department of Breast Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, TX 77230-1439, USA

Corresponding author: Lajos Pusztai, [lpusztai@mdanderson.org](mailto:lpusztai@mdanderson.org)

Published: 18 December 2009

This article is online at

<http://breast-cancer-research.com/supplements/11/S3/S11>

*Breast Cancer Research* 2009, **11(Suppl 3):S11**

(doi:10.1186/bcr2430)

© 2009 BioMed Central Ltd

### Molecular types of breast cancer

Important differences in the clinical behaviour of oestrogen receptor (ER)-positive and ER-negative cancers have been recognised for a long time [1]. Nevertheless, breast cancer was regarded as a single disease with variable histology and clinical course. More recently, high-throughput analytical methods revealed unexpectedly large-scale molecular differences between ER-positive cancers and ER-negative cancers [2]. These results prompted a conceptual shift in the classification of breast cancer, which is increasingly viewed not as a single disease but as a collection of several biologically distinct neoplastic diseases that arise from the breast epithelium.

The different molecular types of breast cancer may originate from different epithelial precursors such as luminal (ER-positive cancers) or basal (ER-negative tumours) epithelial cells, or may represent different stages of arrest along a spectrum of differentiation [3,4]. ER-positive cancers can be further divided into good prognosis and poor prognosis subgroups. At the molecular level, this division corresponds to differences in proliferative activity [5]. Clinical heterogeneity also exists within ER-negative cancers but the molecular differences that are associated with prognosis or chemotherapy sensitivity are less easily identifiable with gene expression analysis or other high-throughput analytical methods. This may be due to fewer and more subtle molecular differences between subsets of ER-negative cancers or to a larger number of subsets that exist among these tumours, and therefore their delineation will require much larger sample sizes than currently available.

Four different molecular types of breast cancers can currently be defined with confidence: ER/progesterone receptor-negative and HER-2-negative (triple-negative or basal-like) breast cancer; highly proliferative ER-positive breast cancer (luminal B type or MammaPrint high risk or OncotypeDX high recurrence score cancers); low proliferation ER-positive breast cancers (luminal A type or MammaPrint low risk or OncotypeDX low recurrence score cancers); and HER-2

amplified breast cancer. The gene expression characteristics that define the HER-2 amplified cancers are smaller than those that define triple-negative breast cancers or separate the low-risk and high-risk ER-positive cancers, but considering HER-2-positive cancers as a separate molecular type is nevertheless justified on the grounds of its unique sensitivity to HER-2 targeted therapies and the characteristic HER2 gene amplification that defines this disease.

It is important to recognise that there is no standardised and uniformly accepted molecular assay to assign molecular class to breast cancer. The original intrinsic subtype predictor has undergone important methodological changes in each subsequent publication. The genes that are used for classification, the prediction algorithm, and the reference or training population vary from manuscript to manuscript. Clinicians adopted a simple, routine marker-based classification schema that includes triple-negative cancers, HER-2 amplified cancers and low-proliferation and high proliferation ER-positive cancers. These four clinical subsets correspond closely ( $\geq 80\%$  concordance) to groups defined by various molecular classification methods that rely on gene expression analysis or immunohistochemistry. The schema attests for the robustness of the original concept that all the different molecular classification methods tend to assign individual patients, with reasonably high concordance, to the same molecular subtype.

### Which ER-positive breast cancers require adjuvant chemotherapy in addition to endocrine treatment?

This question represents a central challenge in the management of stage I to stage II ER-positive breast cancers, and the answer often relies on a great deal of clinical judgment. Chemotherapy selection for these patients is largely based on the estimated risk of recurrence determined by tumour size and nodal status. Seasoned clinicians also consider the age of the patient, tumour grade and perhaps the semi-quantitative ER, progesterone receptor and Ki67 expression

values that each influence to a variable extent the risk of recurrence and endocrine sensitivity. Unfortunately, the human mind is not well suited for rapid and reproducible calculation of multivariate prediction models. It is difficult to estimate precisely whether a 47-year-old woman with a 1.8 cm, grade 2, 15% ER-positive, 10% Ki67 invasive cancer has better or worse prognosis than a 65-year-old woman with a 3 cm, 80% ER-positive and high nuclear grade tumour.

The most important practical implication of gene expression analysis and the new molecular classification schema of breast cancer has been the development of multigene-prognostic predictors that can assist in assessing the risk of recurrence in ER-positive cancers. Multigene predictors are conceptually similar to other multivariate outcome prediction models such as the AdjuvantOnline prognostic model or nomograms that predict complete response to neoadjuvant chemotherapy or the probability of finding additional positive axillary lymph nodes after a positive sentinel lymph node biopsy [6-8]. The multivariate models have two important advantages over the gestalt of physicians; they combine variables in a mathematically optimal manner by weighing the importance of each variable based on evidence (derived from the training data), and they provide a standardised tool for decision-making.

Owing to the substantial molecular differences that exist between poor prognosis and good prognosis ER-positive cancers that are driven by the large number of genes involved in regulating and executing cell proliferation, it is relatively easy to develop gene expression-based prognostic predictors for ER-positive cancers. Indeed, a number of such markers have been proposed and newer ones are constantly reported. Not all of these, however, are equally ready for clinical use. Reproducibility and robustness of the analytical method must be established along with potential sources of variation. The assay procedure has to be standardised and the decision thresholds have to be fully defined before clinical validation. The predictive performance of the proposed assay needs to be validated in independent patient cohorts that meet the characteristics of the intended use of the assay (for example, a prognostic assay cannot be meaningfully tested on patients who receive different types of systemic adjuvant therapies). The validation sample size has to be large enough to yield sufficiently narrow confidence intervals around the point estimates of test performance.

The generation of such supporting data is expensive, time consuming and it is often painstakingly difficult to obtain the appropriate clinical samples. Not surprisingly, therefore, few tests have gone through these validation steps. The best evidence currently exists for the use of the OncotypeDX assay as a risk stratification tool for ER-positive breast cancer. There are also convincing data to support the use of MammaPrint for a similar purpose. The performance characteristics and the data that support their use have extensively

been reviewed in several recent manuscripts [2,9]. It is important to recognise that OncotypeDX is not helpful to predict prognosis of ER-negative patients because all of these patients are categorised as high risk for recurrence. Similarly, MammaPrint is also not very efficient in identifying good prognosis ER-negative cancers – it categorises almost all ER-negative cancers (>90%) as high risk. It is certain that new assays will enter the market shortly with similar function (for example, the genomic grade index, PAM50) [10,11].

An important and consistently observed phenomenon in genomic prognostic marker research is that these predictors provide independent prognostic information when considered together with routine clinical characteristics (for example, tumour size, nodal status, histological grade) in multivariate analysis [2,9]. Equally importantly, however, the tumour size and nodal status also provide prognostic information independent of the genomic assay results. This suggests that integration of molecular predictor results with tumour size and nodal status could further increase predictive accuracy and should be considered together in the decision-making process. It is clear in several clinical scenarios that appropriate decisions can be made without the use of any molecular prognostic test; in many circumstances, however, these tests can lead to more informed decision-making.

### **One study for all breast cancers or separate studies for each molecular subtype?**

If breast cancer is a collection of several distinct neoplastic diseases, then it may not be prudent to conduct a single therapeutic trial or prognostic marker discovery study for all breast cancers together. Indeed, some of the conflicting results from various randomised therapeutic trials and biomarker studies may be explained by the unrecognised molecular heterogeneity and unintended systematic differences in trial populations. For example, different adjuvant clinical trials can accrue substantially different proportions of OncotypeDX low-risk and high-risk ER-positive cancers [12]. These differences can have a profound effect on the power of these studies to measure chemotherapy effect. A modest increase in the proportion of low-risk patients (who benefit little or none from adjuvant chemotherapy) from 40 to 60% can decrease the power of a trial from 80 to 60% to detect benefit from adjuvant chemotherapy [12].

It is also plausible to assume that different prognostic and treatment response markers may be optimal for the different molecular types of breast cancer. Indeed, emerging data suggest that even traditional markers such as proliferation carry different predictive values in ER-positive cancers and in ER-negative cancers [5]. Proliferative activity is generally higher in ER-negative breast cancers compared with ER-positive disease, but its prognostic and predictive value is less within the ER-negative subset than among ER-positive cancers. Different thresholds can be optimal to separate cases into low or high marker groups among ER-positive and

ER-negative cancers, respectively. Systematically different distribution of the genomic grade index (a 90-gene signature that distinguishes low-grade and high-grade cancers) between ER-positive and ER-negative cancers was also observed and translated into different chemotherapy response predictive values in these two different types of cancers [13].

A further potential confounder in combined analysis of all breast cancers in a biomarker study is that the same biomarker can interact to a variable, and sometimes opposing, extent with several different biological functions that determine survival. For example, microtubule binding protein Tau is an oestrogen-regulated gene that has low expression in ER-negative cancers and in luminal B, highly proliferative, ER-positive cancers. Combined analysis of all breast cancers can easily identify low expression of this molecule as a marker of increased chemotherapy sensitivity since ER-negative cancers and highly proliferative ER-positive cancers are more sensitive to cytotoxic drugs. Because of its association with the molecular and clinical phenotype, however, high expression of Tau can also be considered a marker of good prognosis, endocrine-sensitive, ER-positive cancers [14]. When overall survival is plotted for all patients treated with endocrine and chemotherapies by Tau status, high Tau expression may be associated with better survival (if most patients in the study are ER-positive and the sensitivity of high Tau to predict endocrine responsiveness is greater than its sensitivity to predict chemotherapy resistance) or it may be associated with worse survival (if most patients are ER-negative and high Tau has higher sensitivity for predicting chemotherapy resistance than endocrine sensitivity) [15].

Similar opposing associations exist between proliferative activity, tumour grade, OncotypeDX score, and chemotherapy and endocrine therapy sensitivity. Such associations lead to apparent paradoxes such as high-grade cancers being more sensitive to chemotherapy yet having worse overall survival. Predictive and prognostic marker research would benefit from refocusing efforts to develop markers separately for different clinical and molecular types of breast cancers.

## Competing interests

The author declares that they have no competing interests.

## Acknowledgement

This article has been published as part of *Breast Cancer Research* Volume 11 Suppl 3 2009: Controversies in Breast Cancer 2009. The full contents of the supplement are available online at <http://breast-cancer-research.com/supplements/11/S3>.

## References

- Hess KR, Pusztai L, Buzdar AU, Hortobagyi GN: **Estrogen receptors and distinct patterns of breast cancer relapse.** *Breast Cancer Res Treat* 2003, **78**:105-118.
- Sotiriou C, Pusztai L: **Gene-expression signatures in breast cancer.** *N Engl J Med* 2009, **360**:790-800.
- Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Børresen-Dale AL, Brown PO, Botstein D: **Molecular portraits of human breast tumours.** *Nature* 2000, **406**:747-752.
- Prat A, Perou CM: **Mammary development meets cancer genomics.** *Nat Med* 2009, **15**:842-844.
- Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, Desmedt C, Ignatiadis M, Sengstag T, Schütz F, Goldstein DR, Piccart M, Delorenzi M: **Meta-analysis of gene-expression profiles in breast cancer: toward a unified understanding of breast cancer sub-typing and prognosis signatures.** *Breast Cancer Res* 2008, **10**:R65.
- Olivetto IA, Bajdik CD, Ravdin PM, Speers CH, Coldman AJ, Norris BD, Davis GJ, Chia SK, Gelmon KA: **Population-based validation of the prognostic model ADJUVANT! for early breast cancer.** *J Clin Oncol* 2005, **23**:2716-2725.
- Rouzier R, Pusztai L, Delaloge S, Gonzalez-Angulo AM, Andre F, Hess KR, Buzdar AU, Garbay JR, Spielmann M, Mathieu MC, Symmans WF, Wagner P, Atallah D, Valero V, Berry DA, Hortobagyi GN: **Nomograms to predict pathologic complete response and metastasis-free survival after preoperative chemotherapy for breast cancer.** *J Clin Oncol* 2005, **23**:8331-8339.
- Coutant C, Olivier C, Lambaudie E, Fondrinier E, Marchal F, Guillemin F, Seince N, Thomas V, Levêque J, Barranger E, Darai E, Uzan S, Houvenaeghel G, Rouzier R: **Comparison of models to predict nonsentinel lymph node status in breast cancer patients with metastatic sentinel lymph nodes: a prospective multicenter study.** *J Clin Oncol* 2009, **27**:2800-2808.
- Marchionni L, Wilson RF, Wolff AC, Marinopoulos S, Parmigiani G, Bass EB, Goodman SN: **Systematic review: gene expression profiling assays in early-stage breast cancer.** *Annals Internal Med* 2008, **148**:358-369.
- Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, Gillet C, Ellis P, Harris A, Bergh J, Foekens JA, Klijn JG, Larsimont D, Buyse M, Bontempi G, Delorenzi M, Piccart MJ, Sotiriou C: **Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade.** *J Clin Oncol* 2007, **25**:1239-1246.
- Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM, Bernard PS: **Supervised risk predictor of breast cancer based on intrinsic subtypes.** *J Clin Oncol* 2009, **27**:1160-1167.
- Pusztai L, Broglio K, Andre F, Symmans WF, Hess KR, Hortobagyi GN: **Effect of molecular disease subsets on disease-free survival in randomized adjuvant chemotherapy trials for estrogen-receptor positive breast cancer.** *J Clin Oncol* 2008, **26**:4679-4683.
- Liedtke C, Hatzis C, Symmans WF, Desmedt C, Haibe-Kains B, Valero V, Kuerer H, Hortobagyi GN, Piccart-Gebhart M, Sotiriou C, Pusztai L: **Genomic grade index is associated with response to chemotherapy in patients with breast cancer.** *J Clin Oncol* 2009, **27**:3185-3191.
- Andre F, Hatzis C, Anderson K, Sotiriou C, Mazouni C, Mejia J, Wang B, Hortobagyi GN, Symmans WF, Pusztai L: **Microtubule-associated protein-tau is a bifunctional predictor of endocrine sensitivity and chemotherapy resistance in estrogen receptor-positive breast cancer.** *Clin Cancer Res* 2007, **13**:2061-2067.
- Pusztai L, Jeong JH, Gong Y, Ross JS, Kim C, Paik S, Rouzier R, Andre F, Hortobagyi GN, Wolmark N, Symmans WF: **Evaluation of microtubule associated protein-tau expression as prognostic and predictive marker in the NSABP-B 28 randomized clinical trial.** *J Clin Oncol* 2009, **27**:4287-4292.