

Molecular Portraits and 70-Gene Prognosis Signature Are Preserved throughout the Metastatic Process of Breast Cancer

Britta Weigelt,¹ Zhiyuan Hu,^{4,5} Xiaping He,^{4,5} Chad Livasy,⁶ Lisa A. Carey,⁷ Matthew G. Ewend,⁷ Annuska M. Glas,⁴ Charles M. Perou,^{4,5,6} and Laura J. van't Veer^{1,2,3}

Divisions of ¹Experimental Therapy and ²Diagnostic Oncology, The Netherlands Cancer Institute; ³Agendia B.V., Amsterdam, the Netherlands; ⁴Lineberger Comprehensive Cancer Center; and Departments of ⁵Genetics, ⁶Pathology and Laboratory Medicine, and ⁷Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

Abstract

Microarray analysis has been shown to improve risk stratification of breast cancer. Breast tumors analyzed by hierarchical clustering of expression patterns of “intrinsic” genes have been reported to subdivide into at least four molecular subtypes that are associated with distinct patient outcomes. Using a supervised method, a 70-gene expression profile has been identified that predicts the later appearance or absence of clinical metastasis in young breast cancer patients. Here, we show that distant metastases display both the same molecular breast cancer subtype as well as the 70-gene prognosis signature as their primary tumors. Our results suggest that the capacity to metastasize is an inherent feature of most breast cancers. Furthermore, our data imply that poor prognosis breast carcinomas classified either by the intrinsic gene set or the 70 prognosis genes represent distinct disease entities that seem sustained throughout the metastatic process. (Cancer Res 2005; 65(20): 9155-8)

Introduction

DNA microarray technology, which allows the analysis of the expression levels of a thousand of genes in a single experiment, offers great potential to improve our knowledge of tumor molecular biology but also for the discovery of new molecular markers. Over the past years, microarray analysis has been extensively used to improve the diagnosis and risk stratification of many cancers (1–6). Two major studies have described the use of this technology to assess the molecular classification of human breast cancer and have defined new subgroups that are relevant to patient management. Using hierarchical clustering, Perou et al. showed that breast tumors can be classified into specific subtypes based solely on differences in gene expression patterns named the molecular portraits of breast tumors (3). Three estrogen receptor (ER)–negative subtypes of breast carcinomas were identified (“basal like,” “HER2+,” and “normal breast like”), and later, at least two ER-positive tumor subtypes (“luminal A” and “luminal B”) were defined (4). Importantly, the breast cancer subtypes also represent clinically distinct subgroups of patients, as they show differences in metastasis-free and overall survival (4). Sørlie et al. identified the luminal A subgroup of ER-positive tumors to be

associated with the best outcome, whereas the basal-like and the HER2+ tumors have the worst outcome (4). The identification of the distinct subtypes was based upon the hierarchical clustering of an “intrinsic” gene set, which comprises genes that show little variance within repeated samplings of the same tumor but which show high variance across different tumors (3). Recently, a new intrinsic gene set of 1,300 genes (i.e., the “Intrinsic/UNC” gene set) was identified using a larger set of breast tumors and genes than in the original report and was validated on a test set of tumors from independent microarray studies.⁸ The breast tumor subtypes based on the new Intrinsic/UNC gene set were nearly identical to those previously identified and shown to be reproducible across independent data sets, across different microarray platforms, also with respect to clinical outcome.⁸ In complementary studies, van't Veer et al. used a supervised classification method to identify a gene expression signature in young breast cancer patients with lymph node–negative tumors that is associated with patient outcomes and prognosis (6). Using this signature consisting of 70 genes, primary breast tumors can be classified as having either a “poor prognosis” signature, which means they are likely to metastasize, or a “good prognosis” signature, meaning that the development of metastases is unlikely. This 70-gene signature was subsequently confirmed in a consecutive series of 295 patients with both lymph node–positive and lymph node–negative disease (7). The finding that the metastasis risk of a breast cancer patient can be predicted by the overall gene expression profile of its primary tumor challenged the widely accepted idea that metastatic potential is acquired relatively late during multistep tumorigenesis (8). The concept that the ability to metastasize to distant sites is an early and inherent genetic property of breast cancers seemed supported by our previous finding that the overall molecular program established in a human primary breast carcinoma is preserved in its distant metastasis (9). Here, we test whether the molecular subtypes of breast cancer as well as the 70-gene prognosis profiles are also maintained throughout the metastatic process.

Materials and Methods

Detailed information on RNA isolation, amplification, labeling, hybridization, scanning, microarray analysis, and patient information has been described previously (6, 7, 9).⁸ The microarray data presented here have been deposited into the GEO under the series number of

Requests for reprints: Laura J. van't Veer, Plesmanlaan 121, 1066 CX Amsterdam, the Netherlands. Phone: 31-20-512-2754; Fax: 31-20-512-2759; E-mail: l.vt.veer@nki.nl.
©2005 American Association for Cancer Research.
doi:10.1158/0008-5472.CAN-05-2553

⁸ Z. Hu et al. A new breast tumor intrinsic gene list identifies novel characteristics that are conserved across microarray platforms, submitted for publication.

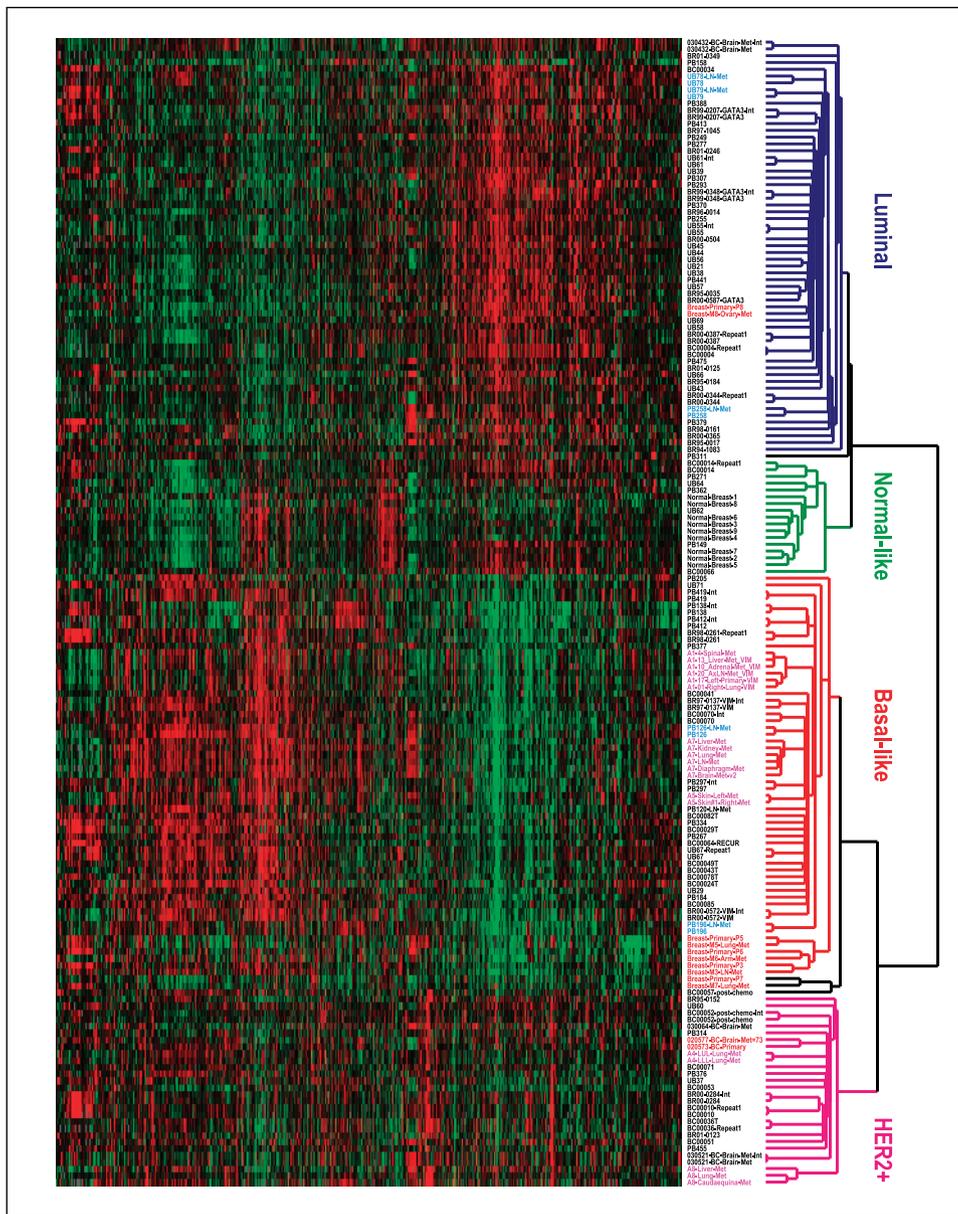


Figure 1. Determination of the molecular intrinsic subtypes of primary tumors and their metastatic counterparts. A hierarchical clustering analysis was performed using a 1,300-gene intrinsic gene list.⁸ Row, tumor; column, gene; red codes, six primary tumor-distant metastasis pairs; light blue codes, primary tumor-lymph node metastasis pairs; pink codes, local and distant metastasis samples from autopsy patients (which for patient A1 includes the primary).

GSE2741, with the 23 samples not described in Hu et al. being GSM52910 to GSM52932.

In brief, seven pairs of matching primary breast tumors and distant metastases of the Netherlands Cancer Institute described previously (ref. 9; pairs 1 and 3-8) were hybridized and analyzed for their 70-gene expression profile (6), and five pairs (pairs 3 and 5-8) analyzed by hierarchical clustering for their molecular breast cancer subtype. The hierarchical clustering analysis was done using the Intrinsic/UNC gene list comprising 1,410 microarray elements (representing 1,300 genes)⁸ using 156 arrays [Agilent Human 1A Oligo Microarray (V2)] representing 107 patients. The molecular subtypes of the samples were determined by the dendrogram that was associated with characteristic gene expression patterns. In addition, a set of five pairs of primary tumors and lymph node metastases, one pair of primary breast tumor and a brain metastasis, as well as multiple distant metastasis samples of five autopsy patients from the University of North Carolina (UNC) at Chapel Hill were analyzed for their molecular breast cancer subtype.

These studies were approved by the Medical Ethical Committee of the Netherlands Cancer Institute and the Institutional Review Board of UNC.

Results

To test whether the molecular subtype of a primary breast tumor is preserved in its metastasis, five of the seven pairs of primary tumors and matching metastases described in Weigelt et al. (9) were retested and analyzed using hierarchical clustering and 107 additional breast tumors using a new breast "intrinsic gene list."⁸ All primary breast tumors paired with their matching metastases, even when added to the large data set of breast tumors from 107 patients (Fig. 1, red codes). Two matching pairs were classified into the HER2+ group, one pair into the basal-like group, and one pair was "ER-negative unclassified" and clustered near the basal-like and the HER2+ group. These ER-negative breast cancer subtypes are associated with the shortest relapse-free and overall survival times (4).⁸ Pair 8 was classified into the luminal group of breast cancer subtypes. The finding that the individual portraits of tumors are maintained in their metastases was further confirmed by the paired clustering of an independent set of five

pairs of primary tumors and simultaneous lymph node metastases (Fig. 1, *light blue codes*) and one primary and metastatic brain tumor pair (Fig. 1, *red code*) of UNC. In a few additional cases from autopsy patients, we were able to sample and compare multiple metastasis sites from the same individual. The primary tumor and metastases of the spinal cord, liver, adrenal gland, lymph node, and lung of autopsy patient A1 were analyzed for their molecular subtype; of autopsy patient A7, metastases of the liver, kidney, lung, lymph node, diaphragm, and brain were tested. Remarkably, all metastasis samples obtained from one breast cancer patient cluster together and show the same molecular breast cancer subtype (Fig. 1, *pink codes*).

In the next step, the maintenance of the 70-gene prognostic signature throughout the metastatic process was tested for seven pairs of primary tumor and matching metastases (6, 9). Five primary breast carcinomas had a 70-gene signature associated with poor prognosis, as did their distant metastases (Fig. 2). Interestingly, pair 8 that was classified into the good prognosis luminal group of breast tumors also has a good prognosis 70-gene signature in both its primary and recurrence that was separated in time by 15 years. Only the primary tumor of pair 4 has a different 70-gene signature than its metastasis.

Discussion

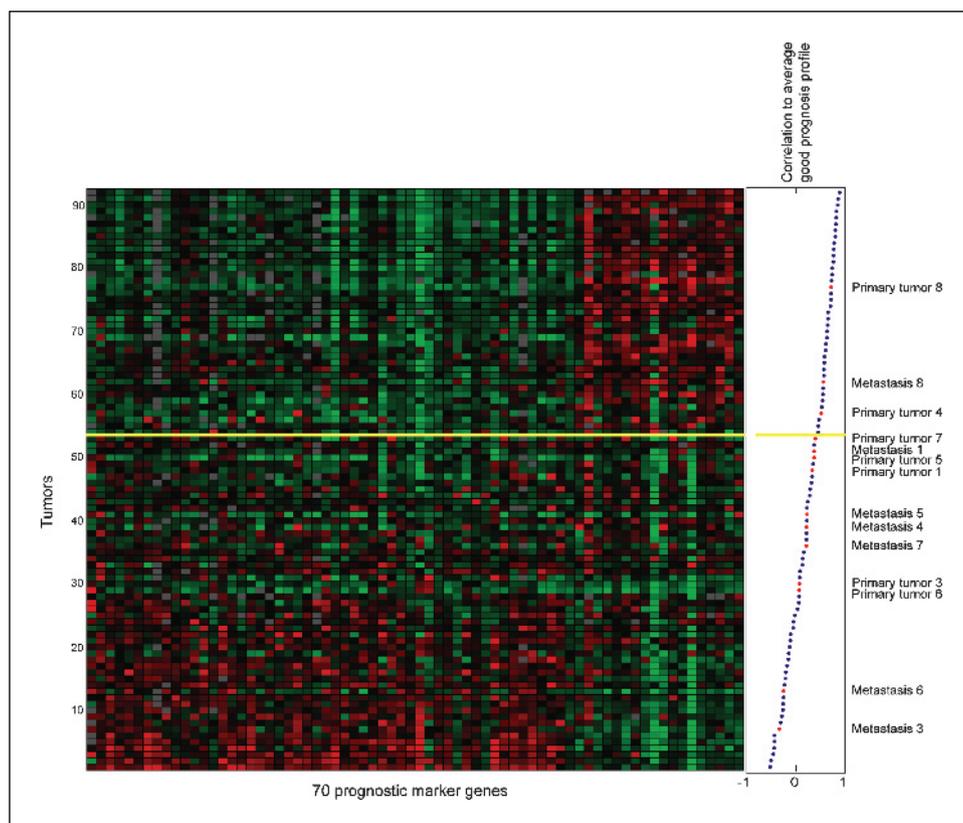
Gene expression profiling confirms that breast cancer is a heterogeneous disease, both biologically and clinically (3, 4, 6, 10, 11).⁸ Nevertheless, the genome-wide expression analysis of breast cancers has made it possible to identify signature patterns in primary breast carcinomas that are associated with patient outcomes and prognoses. Our data show that distant metastases

mirror the specific prognostic profiles, the molecular breast cancer subtypes and 70-gene prognostic signatures, of their primary breast tumors. Remarkably, multiple metastases from one patient all display the same molecular breast cancer subtype independent of the organ in which they developed and still maintain the unique molecular identity of the primary that they arose from. Our findings support the hypothesis that the molecular subtypes might originate from different cell types within the breast and therefore reflect different biological entities (12), which are maintained throughout the multistep metastatic process.

Almost all pairs are either of the poor prognosis molecular subtype HER2+ or basal-like or have a poor prognosis signature of 70 genes, which is what might be expected because all tumors metastasized. Primary tumor 8 and its distant metastasis sample in the ovary, however, had a good prognosis signature of 70 genes and a luminal breast cancer subtype, both of which are associated with a low metastasis risk. Of note, the metastasis of this primary tumor developed 15 years after primary diagnosis, compared with a mean of 4.5 years (range, 1.6-6.3) in the other six patients. Only one primary breast carcinoma did not maintain the 70-gene signature in its distant metastasis (pair 4), although these two samples were relatively close as observed by their Pearson correlation coefficient. This pair was due to limited amounts of RNA not tested for its breast cancer subtype. For the samples analyzed from UNC, all primary tumors and their associated metastases were collected at the same time.

The analysis of the 70-gene signature revealed that in six of seven pairs, the metastasis has a higher correlation with the poor prognosis signature than its primary tumor (see Fig. 2). Genes correlated with the good prognosis signature of 70 genes remain virtually unchanged between primary and metastatic tumors

Figure 2. Expression data matrix of 70 prognostic marker genes from 78 primary breast tumors plus seven pairs of matching pairs of primary tumors and distant metastases (6, 9). Row, tumor; column, gene. Genes are ordered according to their correlation coefficient with the two prognostic groups; tumors are ordered by the correlation to the average profile of the good prognosis group (*right*). Above the yellow line, patients have a good prognosis signature, below the line, a poor prognosis signature.



within a patient. A small number of genes correlated with the poor prognosis signature is, however, up-regulated in the metastases compared with their matching primary breast carcinomas (data not shown), which leads to the “poorer” 70-gene signature of the metastases. The genes up-regulated in the metastatic tumors are among others involved in DNA replication (RFC4 and ORC6L) and signal transduction (IGFBP5 and PRC1; data not shown). Interestingly, only matrix metalloproteinase 9, which plays an important role in the proteolysis of the extracellular matrix (13), is down-regulated in five of the six metastases showing a poorer profile than their primary tumors (data not shown).

Our results presented here emphasize that the metastatic nature of poor prognosis breast carcinomas, which are depicted by the poor prognosis 70-gene profile or the luminal B, HER2+, or basal-like molecular subtype, is an inherent feature of breast cancers that remains stable with time and across distinct tumor outgrowth locations within the same individual. Because both the molecular

breast cancer subtype and prognostic expression profile of a primary breast tumor are maintained throughout the metastatic process, future treatment decisions based on the expression profile of a primary tumor is a rational approach towards preventing the outgrowth of metastases.

Acknowledgments

Received 7/20/2005; revised 8/23/2005; accepted 8/26/2005.

Grant support: Dutch Cancer Society (B. Weigelt and L.J. van't Veer), Cancer Genomics Initiative (B. Weigelt and L.J. van't Veer), National Cancer Institute Breast Specialized Programs of Research Excellence/UNC at Chapel Hill grant P50-CA58223-09A1 (C.M. Perou), National Institute of Environmental Health Sciences grant U19-ES11391-03 (C.M. Perou), National Cancer Institute grant RO1-CA-101227-01 (C.M. Perou), and NIH General Clinical Research Centers grant M01RR00046 (L.A. Carey).

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.

We thank Drs. Phil Bernard, Juan Palazzo, and Olufunmilayo I. Olopade for their contributions towards the 107-patient sample data set.

References

1. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286:531–7.
2. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403:503–11.
3. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747–52.
4. 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.
5. 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.
6. 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.
7. 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.
8. Fidler IJ, Kripke ML. Metastasis results from preexisting variant cells within a malignant tumor. *Science* 1977;197:893–5.
9. Weigelt B, Glas AM, Wessels LF, Witteveen AT, Peterse JL, van't Veer LJ. Gene expression profiles of primary breast tumors maintained in distant metastases. *Proc Natl Acad Sci U S A* 2003;100:15901–5.
10. Gruvberger S, Ringner M, Chen Y, et al. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. *Cancer Res* 2001;61:5979–84.
11. West M, Blanchette C, Dressman H, et al. Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc Natl Acad Sci U S A* 2001;98:11462–7.
12. Sorlie T. Molecular portraits of breast cancer: tumor subtypes as distinct disease entities. *Eur J Cancer* 2004;40:2667–75.
13. Bodey B, Bodey B, Jr., Siegel SE, Kaiser HE. Matrix metalloproteinases in neoplasm-induced extracellular matrix remodeling in breast carcinomas. *Anticancer Res* 2001;21:2021–8.