Circulating tumor cell detection and transcriptomic profiles in early breast cancer patients

In the early 2000s, studies using high-throughput transcriptomic analyses revealed different aspects of breast cancer biology. Several outcome-based predictors and biology-based classifiers have been proposed: invasiveness gene signature (IGS), intrinsic molecular subtype, wound-response signature, etc. Recently, circulating tumor cell (CTC) cytological detection has been associated with metastasis-free and overall survival in the REMAGUS02 neoadjuvant trial [1]. In this article, we report, for the first time, the transcriptomic analysis of 60 nonmetastatic breast cancers according to CTC detection.

The REMAGUS02 trial included breast cancer patients with locally advanced or large breast cancer. CTC detection was carried out at baseline and at the end of chemotherapy using the CellSearch® system (Veridex, Raritan, NJ). Total RNA extracted from the pretreatment cancer biopsy was hybridized on the GeneChip Human Genome U1331 2.0 Array (Affymetrix, Santa Clara, CA). Conditions of the sampling procedures and RNA quality control are detailed elsewhere [2]. The genechip robust multi-array average procedure [3] was used to normalize the gene expression data. A hierarchical clustering was then carried out using the 10 000 probe sets that showed the highest values in the interquartile range (IQR). We applied the intrinsic gene set on the complete dataset of 60 samples to define the molecular subtypes and the IGS. We used the defined and validated centroids of 506 genes to discriminate between previously identified molecular breast cancer subtypes. We matched the probe list UniGene ID (Build#204) to the GeneChip® Human Genome U133A 2.0 Array, resulting in a list of 294 unique probe sets. Each sample was assigned to the nearest subtype/centroid as determined by the highest Spearman rank order correlation between the gene expression values of the 294 probe sets and the 5 subtype centroids. The IGS score was determined by calculating the Pearson correlation between the probe set expression values of each sample and the 110 reference expression levels of the same genes defining the IGS signature. Single probe set analyses were carried out using the Wilcoxon rank sum test. Finally, because half of the probe sets showed the highest IQR values, a significance analysis of microarray (SAM) was proposed to detect differentially associated genes (DEG) between CTC-positive and -negative patients.

At first, unsupervised clustering revealed three major tumor clusters; they unsurprisingly corresponded to triple-negative, human epithelial growth factor receptor 2 (HER2)-positive/estrogen receptor (ER)-negative, and ER-positive breast cancer immunohistochemical phenotypes. CTC detection was not statistically different among these three clusters (Figure 1). Breast cancers were then classified according to selected biology-based signatures: the intrinsic subtype classifier and IGS, which is stemness related. CTC detection was not statistically different among subgroups: basal, n = 6/18 (33%); HER2, n = 3/10 (33%); luminal A, n = 2/15 (13%); luminal B, n = 2/10 (20%); normal like, n = 2/7 (29%); low-risk IGS, n = 9/27 (33%); and high-risk IGS, n = 6/33 (18%). Single probe set analyses were then carried out on candidate genes that are directly involved in CTC detection by the CellSearch® system (cytokeratin 8, 18, 19, and EpCAM) or that are surrogate markers for breast cancer stem cells (CD24, CD44, ALDH1A1): messenger RNA levels were not correlated with CTC positivity. In addition, at the high false discovery rate of 30%, only 18 DEG were found using a SAM procedure.

This is the first study, to date, to correlate CTC detection in nonmetastatic breast cancer with the gene expression profile of the primary tumor. We did not confirm previously published in vitro experiments that suggested that normal-like
breast cancers were not detected by the CellSearch® system [4]. Also, CTC detection was not related to EpCAM expression. Usually, low sample sizes are generally responsible for numerous associations, which are, in the vast majority of cases, artifacts. Here, the opposite results were obtained as no significant statistical association with any gene or profile was observed. Beyond contamination and tumor heterogeneity, two reasons may explain how CTC detection strikingly escaped the power of such large-scale analyses. First, CTC concentration is low in the neoadjuvant setting and some CTC-positive patients may have been missed, lowering the signal–noise ratio of our analysis. Second, whereas high-throughput microarrays successfully predict activation of major biological processes (e.g. inflammation, proliferation), they mostly fail to identify DEG associated with cancer cell dissemination. In early breast cancer patients, there is currently no predictor of lymph node metastasis despite unpublished attempts, whereas for bone marrow disseminated tumor cells, a previously published study reported only 20 DEG, with a high 20% false discovery rate [5]. Our results are, therefore, in line with these studies, suggesting that cellular dissemination is a complex process that may rely on several independent molecular mechanisms, which cannot be directly unraveled by these techniques.

F. Reyal†1, F. Valet†2, P. de Cremoux3, C. Mathiot4, C. Decraene3, B. Asselin2, E. Brain6, S. Delaloge7, S. Giacchetti8, M. Marty8, J. Y. Pierga9,10 & F. C. Bidard9,10*

Departments of 1Surgery, 2Biostatistics, 3Molecular Biology, 4Hematology, 5Translational Research, Institut Curie, Paris; 6Department of Medical Oncology, Hospital René Huguenin, Institut Curie, Saint Cloud; 7Department of Medicine, Institut Gustave Roussy, Villejuif; 8Department of Medical Oncology, Hôpital Saint Louis, Paris; 9Department of Medical Oncology, Institut Curie, Paris; 10University Paris Descartes, Paris, France (*E-mail: fcbidard@curie.fr)

acknowledgements

We thank Mustapha Khazour and Isabelle Vaucher for technical support. This study was presented, in part, at the 2009 San Antonio Breast Cancer Symposium.

funding

French Programme Hospitalier de Recherche Clinique ISRCTN10059974 (PHRC AOM0211).

disclosure

The authors declare no conflict of interest.

references


doi:10.1093/annonc/mdr144
Published online 27 April 2011

Figure 1. Clustering of CTC+ (n = 15) and CTC− (n = 45) patients according to the 10 000 probe sets that have the highest variability. CTC, circulating tumor cell. Estrogen receptor (ER), HER2 status and tumor grade (Gde) were defined by pathological analysis on the initial tumor biopsy. pCR, pathological complete response.