DOI:10.1093/jnci/dju324
First published online October 11, 2014
©The Author 2014. Published by Oxford University Press. All rights reserved.
For Permissions, please e-mail: journals.permissions@oup.com.

ARTICLE

Pluripotent Stem Cell miRNAs and Metastasis in Invasive Breast Cancer

Stefano Volinia, Gerard Nuovo, Alessandra Drusco, Stefan Costinean, Ramzey Abujarour, Caroline Desponts, Michela Garofalo, Raffaele Baffa, Rami Aeqilan, Kati Maharry, Maria Elena Sana Ramiro, Garzon, Gianpiero Di Leva, Pierluigi Gasparini, Paola Dama, Jlenia Marchesini, Marco Galasso, Marco Manfrini, Carlotta Zerbinati, Fabio Corrà, Timothy Wise, Sylwia E. Wojcik, Maurizio Previati, Flavia Pichiorri, Nicola Zanesi, Hansjuerg Alder, Jeff Palatini, Kay F. Huebner, Charles L. Shapiro, Massimo Negrini, Andrea Vecchione, Anne L. Rosenberg, Carlo M. Croce

Manuscript received May 12, 2014; revised June 4, 2014; accepted September 3, 2014.

Correspondence to: Stefano Volinia, PhD, Biosystems Analysis, LTTA, Department of Morphology, Surgery and Experimental Medicine, Università degli Studi, 44121 Ferrara, Italy (e-mail: s.volinia@unife.it); and Carlo M. Croce, MD, Department of Molecular Virology, Immunology, and Medical Genetics, Ohio State University, 460W12th Ave, Room 1082, Columbus, OH 43210 (e-mail: carlo.croce@osumc.edu).

Background The purpose of this study is to determine whether microRNA for pluripotent stem cells are also expressed in breast cancer and are associated with metastasis and outcome.

Methods We studied global microRNA profiles during differentiation of human embryonic stem cells (n =26) and in breast cancer patients (n = 33) and human cell lines (n = 35). Using in situ hybridization, we then investigated *MIR302* expression in 318 untreated breast cancer patients (test cohort, n = 22 and validation cohort, n = 296). In parallel, using next-generation sequencing data from breast cancer patients (n = 684), we assessed microRNA association with stem cell markers. All statistical tests were two-sided.

Results In healthy tissues, the *MIR302* (high)/*MIR203* (low) asymmetry was exclusive for pluripotent stem cells. *MIR302* was expressed in a small population of cancer cells within invasive ductal carcinoma, but not in normal breast ($P < .001$). Furthermore, *MIR302* was expressed in the tumor cells together with stem cell markers, such as CD44 and BMI1. Conversely, *MIR203* expression in 684 breast tumors negatively correlated with CD44 (Spearman correlation, Rho = -0.08, $P = .04$) and BMI1 (Rho = -0.11, $P = .004$), but positively correlated with differentiation marker CD24 (Rho = 0.15, $P < .001$). Primary tumors with lymph node metastasis had cancer cells showing scattered expression of *MIR302* and widespread repression of *MIR203.* Finally, overall survival was statistically significantly shorter in patients with *MIR302*-positive cancer cells ($P = .03$).

Conclusions In healthy tissues the *MIR302*(high)/*MIR203*(low) asymmetry was characteristic of embryonic and induced pluripotency. In invasive ductal carcinoma, the *MIR302*/*MIR203* asymmetry was associated with stem cell markers, metastasis, and shorter survival.

JNCI J Natl Cancer Inst (2014) 106(12): dju324 doi:10.1093/jnci/dju324

Several investigators have suggested that a small proportion of cancer cells within certain tumors might have the properties of cancer initiating or cancer stem cells (CSCs) (1). The CSC hypothesis provides an attractive mechanism to account for the therapeutic refractoriness and dormant behavior exhibited by tumors (2,3). Breast tumors are also thought to contain CSCs reminiscent of normal stem cells, and poorly differentiated breast cancers (BCs) display high content of prospectively isolated CSCs (4). Furthermore, the induction of epithelial-mesenchymal transition (EMT) in transformed mammary epithelial cells creates cells that appear to be enriched for CSCs, as gauged by tumor-seeding ability, mammosphere formation and cell-surface markers (5,6). Overall, the study of CSC biology is predicated on the ability to accurately assess the CSC representation within tumors (7).

MicroRNAs (miRNAs) are small noncoding RNAs that play important post-transcriptional roles by repressing messenger RNA activity. miRNAs are crucial for embryonic stem cells' (ESCs') self-renewal and differentiation; miRNAs from the *MIR302* cluster (hsa-miR-302a/b/c/d) predominate in human ESCs (8), and their promoter is turned off later in development (9). Oct4 and Sox2 are transcription factors required for pluripotency during early embryogenesis and for the maintenance of ESCs. Oct4 and Sox2 bind to

a conserved promoter region of *MIR302* and regulate its expression (10,11). It has been reported that *MIR302* can reprogram somatic and cancer cells into induced pluripotent stem cells (iPSCs) (11–14). Our hypothesis was that *MIR302* is expressed in CSCs within breast tumors, where it acts to induce pluripotency and eventually metastasis (15). Thus, we examined *MIR302* expression in normal breast and invasive ductal carcinoma (IDC).

Methods

Cell Cultures, Tissues, and Expression

All reagents for stem cell differentiation and induction were from Invitrogen/Gibco (Carlsbad, CA), except where mentioned otherwise. Prior to differentiation, H1 cells were cultured on irradiated mouse embryo fibroblasts in complete ESC media; DMEM/F12 (11330032), 20% knock out replacement serum (10828-028), 2 mM Glutamax (35050061), 0.11 mM β-mercaptoethanol (21985023), 10 ng/ml basic fibroblast growth factor. All differentiation experiments were performed in triplicate and are described in details in Supplemental Methods (available online). All tissues were obtained under the guidelines of approved protocols from the Ohio State University Internal Review Board (2009E0406, 2009C0004), and informed consent was obtained from each subject or from his or her guardian. LNA in situ hybridization (ISH) for hsa-miR-302a/b/c and d was performed as described in Supplemental Methods (available online). Twenty-two tumors were studied by ISH on excisional biopsies. Two hundred and ninety-six IDC cases and 68 normal breast controls were studied by in situ hybridization on tissue microarrays (TMAs). Thirty three primary IDCs and the respective matched metastases were analyzed by miRNA microarrays (ArrayExpress accession number E-TABM-971). Three pathologists analyzed the slides blinded to clinical data. ISH scores were the consensus of the pathologists' individual scores. Hybridization included no-probe for background assessment, scrambled probe as, and U6 detection as positive control. Global expression of miRNA was studied using oligonucleotide microarrays (16).

Study Design

Study design, patient selection, RNA isolation from tumor material, histopathology analyses, clinical annotation, and clinical interpretation were carried out at the Ohio State University. RNA amplification and microarray hybridizations were carried out at the Comprehensive Cancer Center of Ohio State University. Figure 1 illustrates the strategy we used to derive and validate the miRNA microsignature associated with stem cell pluripotency and clinical outcome in breast cancer. The clinical end point in this analysis was overall survival.

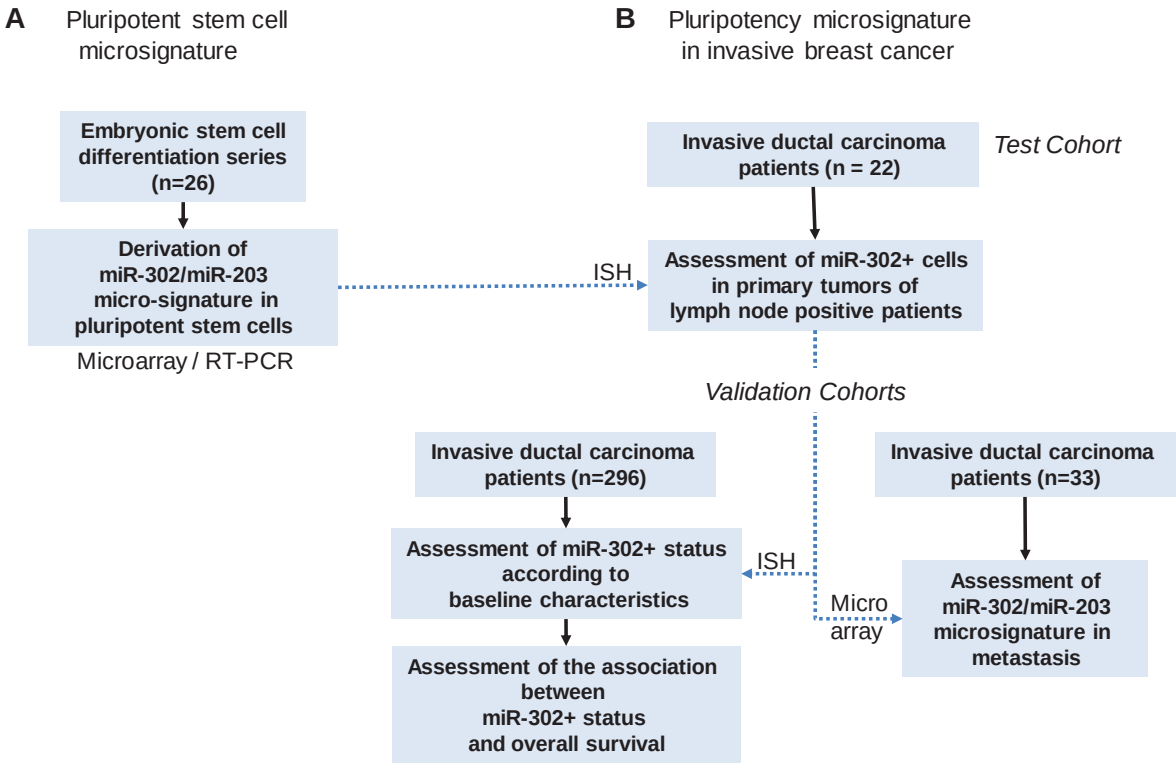

Figure 1. The experimental strategy to derive and validate the miRNA microsignature associated with stem cell pluripotency and clinical outcome in breast cancer. The **diagram** shows the strategy to derive and validate the minimal miRNA signature in Test and Validation Groups. **A)** Shows the approach to the development and validation of the miRNA signature associated with pluripotency in embryonic and induced pluripotent stem cells. Microarrays and real-time polymerase chain reaction were used to identify the key microRNAs associated to pluripotent stem cells. **B)** Shows the strategy to test and validate the pluripotency microsignature in breast cancer. In situ hybridization and microarrays were applied to microRNA detection and quantification. The clinical end point for this analysis was overall survival. ISH = in situ hybridization; RT-PCR = real-time polymerase chain reaction.

Bioinformatics and Statistical Analysis

Microarray analyses, including the miRNA-directed messenger RNA degradation (17) and real-time polymerase chain reaction (RT-PCR) detection of miRNA were performed as detailed in the Supplemental Methods (available online). The Mann-Whitney U was used to compare continuous variables; the Fisher's exact test and the Chi-square test were used to compare categorical variables. The Cochran-Armitage trend test was used to assess the association between *MIR302* status and increasing degree of lymph node involvement (N) or increasing disease stage (AJCC 7th Edition). We used two validated stem cell datasets, available from the Gene Expression Omnibus at the National Center for Biotechnology Information (datasets GSE14012 [18] and GSE12499 [19]) for the mRNA profiles of embryonic, induced pluripotent, and adult neural stem cells. Overall survival (OS) was measured from the date of study entry until the date of death and patients alive at the last follow-up were censored. Survival estimates were calculated using the Kaplan-Meier method, and groups were compared using the log-rank test. We used Cox regression with backward selection for the multivariable analysis of prognostic markers and overall survival. All tests were two-sided. A P value of less than .05 was considered statistically significant. To identify the genes whose expression was significantly related to an miRNA expression, we computed a significance level for each gene by testing the hypothesis that the Spearman's correlation between gene and miRNA expression was zero. These P values were then used in a multivariable permutation test, in which the miRNA levels were randomly permuted among samples. The false discovery rate was the proportion of genes claimed to be correlated to the miRNA in the permutation test. The multivariable permutation test is nonparametric and does not require the assumption of Gaussian distributions. Additional methods are described in the Supplementary Methods (available online).

Results

Expression of *MIR302* and *MIR203* in Pluripotent Stem Cells

Embryonic stem cells express high levels of *MIR302* miRNAs, not detected in any adult healthy tissue or cell type, including CD34+ hemopoietic stem cells (Supplementary Figure 1 and Supplementary Table 1, available online). We studied ESCs, embryoid bodies (EBs) at seven and 14 days, trophoblast and spontaneously differentiating monolayers (total n = 26) to determine the expression profile of *MIR302* and of the other miRNAs in development (Supplementary Table 2, available online). The miRNA profiles for the ESC-derived cell types demonstrated high-lineage specificity (Figure 2A). RT-PCR confirmed and strengthened the microarray results. In fact, the *MIR302* locus (8,9) was highly expressed in ESCs and also in induced pluripotent stem cells (Supplementary Table 3, available online). The expression of the *MIR302* locus (8,9) was downregulated in monolayers, trophoblast, and 14 days' embryoid bodies (Figure 2B). Conversely, stemness repressor *MIR203* (20), not expressed in both pluripotent ESCs and iPSCs, was expressed in all ESC-derived lineages, including the early seven-day EBs (Figure 2C). Hence, the ESC and iPSC pluripotency status was associated with the inverse molecular status of these two miRNAs, *MIR302* high and *MIR203* low (Figure 2D). In fact, none of the healthy adult tissues or cell types had a *MIR302/MIR203* asymmetry (measured by the expression ratio) as high as pluripotent stem cells (Supplementary Table 3, available online). Other miRNAs showed specific profiles along the ESC differentiation route, for example the EMT regulator hsa-miR-205 (21,22) was most expressed in spontaneously differentiated cells and in the trophoblast. miR-372, previously implicated in iPSC induction (14), was high in embryoid bodies (Figure 2A; Supplementary Tables 2 and 3, available online).

MicroRNAs control protein output and exert their activity by specifically repressing translation and by reducing targets' messenger RNA levels (23,24). Therefore, we compared the miRNA-directed messenger RNA degradation (17) in ESCs, iPSCs, and neural stem cells to that of adult cell types. Using this in silico assay, *MIR302* miRNAs were identified as the most active microRNAs in the degradation of ESC target mRNAs (P = .01) (Supplementary Tables 4 and 5, available online). *MIR302* repression activity was also statistically significant in iPSC mRNA profiles (18), but not in those of neural stem cells (19), confirming the lack of *MIR302* expression in nonpluripotent stem cells. The *MIR302* cluster was found to control nuclear proteins and transcription factors involved in embryonic development and differentiation, mitosis and apoptosis (Supplementary Figure 2 and Supplementary Table 6, available online).

MIR302 Expression in Cancer

High levels of the miRNAs from the *MIR302* locus had been reported in malignant germ cell tumors (8,25,26), such as seminoma and teratocarcinoma. Expanding on the CSC hypothesis, we decided to investigate *MIR302* in BC using in situ hybridization (ISH) for single-cell detection of *MIR302*. We standardized the ISH on mouse embryos and on a negative control: unremarkable adult breast. The embryos showed scattered *MIR302*-positive cells. Colocalization ISH with CD44 showed that the cells expressing hsa-miR-302a in embryos were also positive for the CD44 mesenchymal stem cell marker (Supplementary Figure 3, available online). No signal for *MIR302* was detected in the normal adult breast.

ISHs on excisional biopsies of BC showed that tumors did not have a basal and diffuse MIR302 expression, rather often contained some *MIR302*-positive cells (Figure 3). Colocalization experiments showed that many cells in the cancer sections were positive with CD44, but only a smaller fraction was positive to hsa-miR-302a (Supplementary Figure 4, available online). On the other hand, all cells expressing hsa-miR-302a were also positive (+) for CD44. In addition, the *MIR302*-positive cells had the cytological features of cancer and expressed cytokeratin AE 1/3, confirming that they were malignant epithelial cells (Figure 3).

The Stem Cell and EMT Markers vs miR-203 in Breast Cancer

We investigated the expression of the *MIR203* gene in tumors from The Cancer Genome Atlas (TCGA) BC cohort, where both miRNA and mRNA levels were measured by RNAseq in 684 patients (27). Such a large cohort was required to robustly assess the effect of a miRNA in vivo because: 1) BC is a highly heterogeneous disease; 2) tumors are complex tissues with a

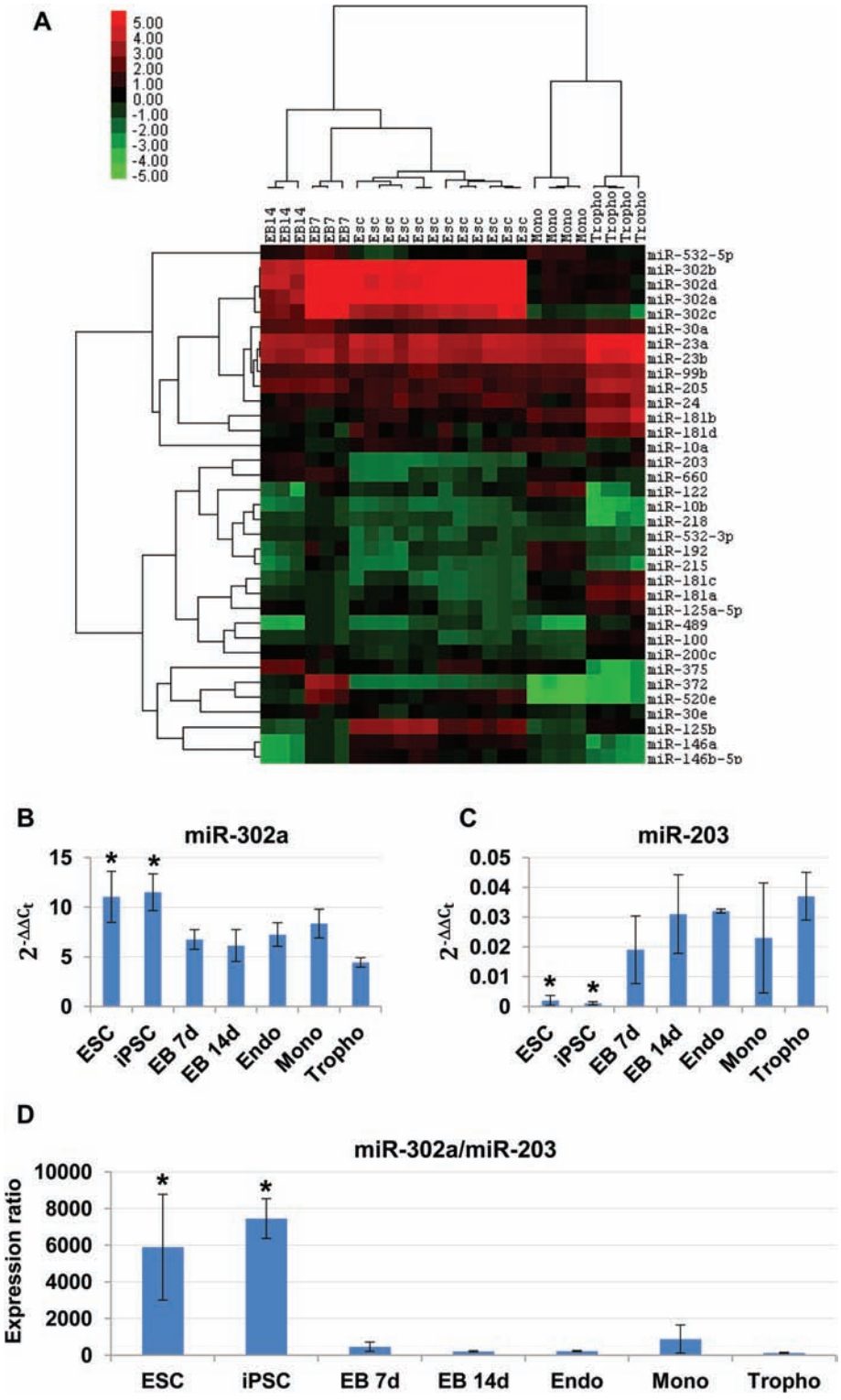

Figure 2. *MIR302* expression in a context of repressed *MIR203* locus indicates pluripotency in stem cells. **A)** The miRNA microsignature in differentiating human embryonic stem cells (16). Real-time polymerase chain reaction detection of mature microRNA in pluripotent stem cells. The **bars** indicate the relative expression for the hsa-miR-302a **(B)**, hsa-miR-203 **(C),** and hsa-miR-302a/hsa-miR-203 ratio **(D)**. Normalization was performed with RNU44. Values are averages of at least three independent experiments. The **asterisk** indicates *P* values under .001 (two-tailed *t* test was used to compare either embryonic stem cells or induced pluripotent stem cells to the nonpluripotent cell types). **Error bars** indicate standard deviation. EB 7d/EB 14d = embryoid bodies (7 and 14 days, respectively); Endo = definitive endoderm; ESC = embryonic stem cells; iPSC = induced pluripotent stem cells; Mono = spontaneously differentiating monolayer; Tropho = trophoblasts.

variable cancer cell proportion; and 3) miRNAs are only one of the components regulating protein output (23). In these tumors, miR-203 was inversely correlated with markers of BC stemness, EMT, and cancer progression (Supplementary Figure 5 and Supplementary Table 7, available online): BMI1 (Spearman correlation, Rho = -0.11, *P* = .004), CD44 (Rho = -0.08, *P* = .04), GATA3

(Rho = -0.15, $P < .001$), TWIST1 (Rho = -0.11, $P = .01$), ZEB1 (Rho = -0.12, $P = .003$), ZEB2 (Rho = -0.11, $P = .003$), SMAD3 (Rho = -0.17, $P < .001$), LEF1 (Rho = -0.2, $P < .001$), and PTEN (Rho = -0.12, $P = .002$). Conversely, miR-203 was positively correlated with markers of differentiation and MET, the reverse process to EMT (Supplementary Figure 5 and Supplementary Table 7, available online): CD24 (Spearman correlation, Rho = 0.15, $P < .001$), GRHL2 (Rho = 0.1, $P = .009$), ELF3 (Rho = 0.15, $P < .001$), ESRP1 (Rho = 0.11, $P = .005$), and ESRP2 (Rho = 0.14, $P < .001$) (28).These findings show that miR-203 levels are inversely correlated with CSC and EMT markers in vivo, and its expression is lost during cancer progression.

To further substantiate the findings from tumors, we studied the *MIR203* expression in 35 cell lines spanning all molecular subtypes of BC. These cell lines have been characterized for their PAM50 gene expression levels (29), stem cell content, and invasiveness (30,31). The cell lines with low *MIR203* expression were all classified as claudin-low (*t* test, $P = .006$) (Supplementary Figure 6, available online). In particular, the cell lines with the lowest miR-203 content, ie, MDA-MB-231, MDA-MB-436, and SUM-1315, were among those with the highest content in stem cells and/or with the highest invasive potential. These results are well in agreement with our results from the TCGA cohort and with Taube and colleagues (32), who recently showed that epigenetic silencing of *MIR203* was required for EMT and cancer stem cell properties.

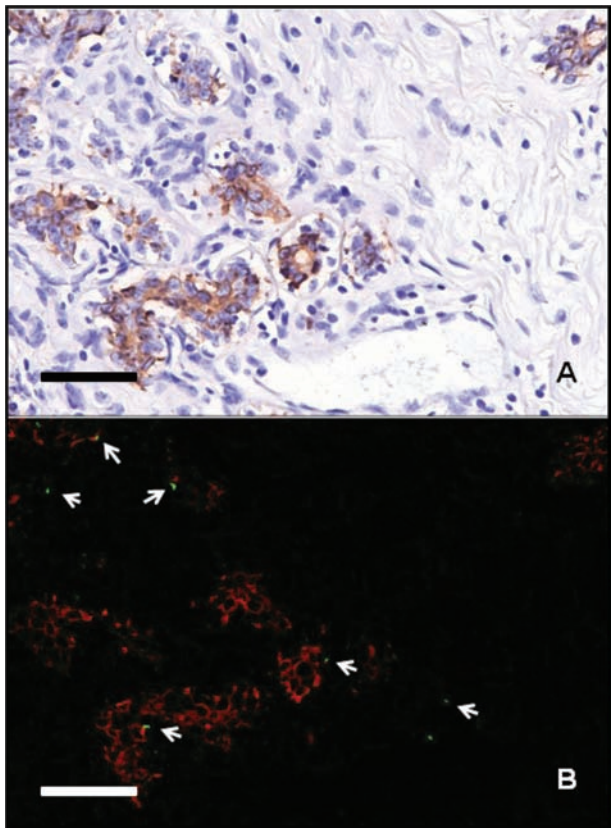

Figure 3. The hsa-miR-302a positive cells in breast cancer biopsies are carcinoma cells. Colocalization of pluripotency-specific hsa-miR-302a and of the carcinoma marker AE1/3 was performed on breast cancer excisional biopsies **(A)**; 400X magnification. Analysis with the Nuance system demonstrated the AE1/3 marker as **fluorescence-red** and the miR-302a/AE1/3 double-positive cells as **green (B, white arrows)**; 400X magnification. To confirm our in situ hybridizations, serial section assays with the different hsa-miR-302a/b/c probes showed overlapping labeling patterns. The **scale bar** in each panel is 75 micrometers.

MIR302 Association With Nodal or Distant Metastasis and Survival in Invasive Ductal Carcinoma

We initially investigated the association between clinical data and microRNA, by using ISH, in a test cohort of 22 women with IDC (10 N0 and 12 N-positive patients). We concentrated on the ISH detection of *MIR302*-positive cells to identify candidate pluripotent cells within the tumors (Figure 4). Primary IDC with nodal metastasis had higher concentration of *MIR302*-positive cancer cells than primary IDC with no lymph node metastasis (median = 2.75 vs 0 cells/cm^2, Mann-Whitney U test, $P = .02$). Tissue microarrays (TMAs) allowed us to extend the initial ISH study on a validation cohort of 296 IDC cases (Figure 5). TMAs confirmed that *MIR302*-positive cancer cells were present in IDC (55 out of 296 samples), but not in normal breasts (0 out of 68, $P < .001$, Fisher's exact test). To identify potential breast CSCs, we used antibodies against BMI1, a protein that plays an important role in regulating self-renewal of tumorigenic mammary stem cells (33,34). Most of the cells expressing miR-302a were also positive to BMI1 (Figure 5D), while only a minority of the BMI1-positive cells also coexpressed miR-302a (Figure 5E). Table 1 summarizes the association between *MIR302* and clinical variables in the TMA patients. In this larger cohort, there was again an association between *MIR302* and lymph node metastasis (trend test, $P = .005$; N0 vs N+, Fisher's exact test, $P = .03$). Furthermore, *MIR302*-positive tumors were more frequent among patients with higher disease stage (trend test, $P = .03$). We then investigated the

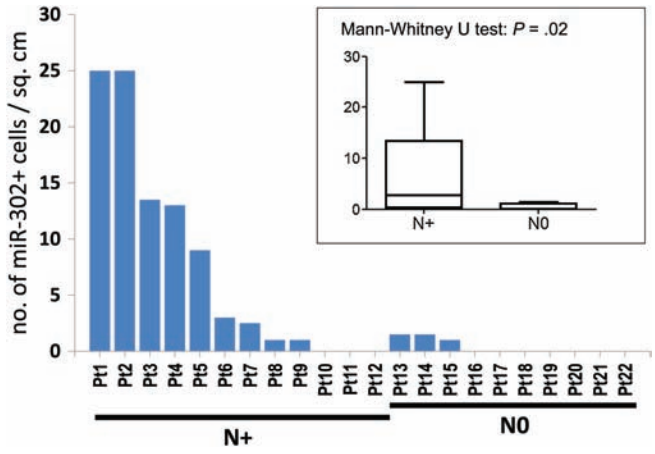

Figure 4. *MIR302* expression is higher in primary invasive ductal carcinoma with lymph node metastasis (test cohort). Primary tumors with lymph node metastasis contain larger amount of hsa-miR-302a positive cells, median = 2.75 positive cells/cm^2, than those with no lymph node metastasis, median = 0 cells/cm^2 ($P = .02$, two-tailed Mann-Whitney U test). Cancer tissues from 22 women with a history of invasive breast cancer were used. N+ on the abscissa indicate primary tumors with lymph node metastasis, N0 indicate those without metastasis. LNA in situ for hsa-miR-302a was performed on excisional biopsies. For multiple slides from the same biopsies the median scores were used. A patient who developed contralateral breast cancer was removed from the study. The **insert box plot** displays the number of hsa-miR-302a positive cells/cm^2 for the lymph node positive and negative patients (**error bars** represent standard deviation).

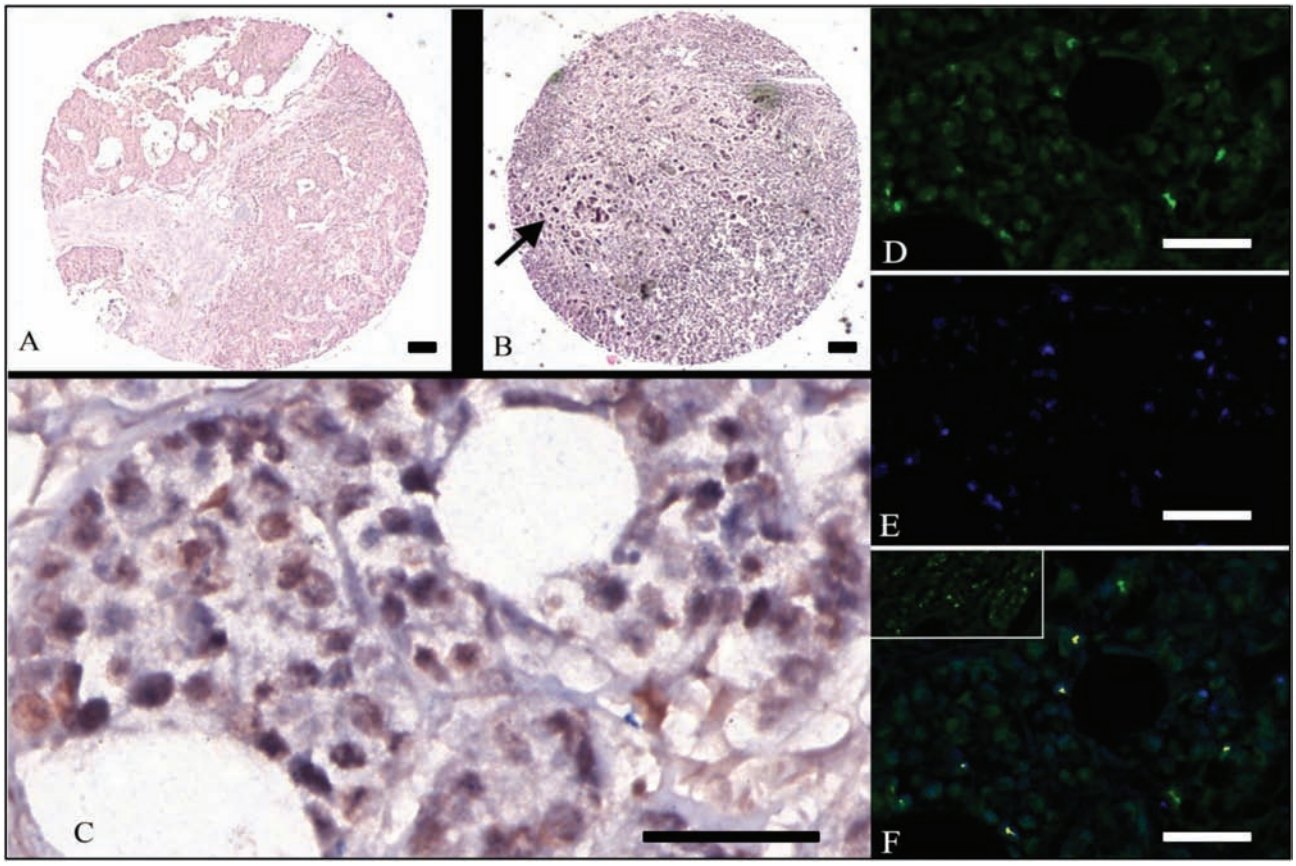

Figure 5. *MIR302a* microRNAs and BMI1 expression in invasive ductal carcinoma with metastasis (validation cohort). **Panels A and B** show cores of a TMA from patients with breast cancer and no metastases (**panel A**) or with lymph node metastases (**panel B**). Note that the former has no detectable hsa-miR-302a positive cells, whereas the latter core has multiple cancer cells with a high level of hsa-miR-302a (**arrow**). **Panel C** shows a breast cancer after codetection of hsa-miR-302a and BMI1. The Nuance system converted the images to **green** (BMI1, **panel D**), **blue** (hsa-miR-302a, **panel E**) and then merged, such that cells coexpressing both are in **yellow false color** (**panel F**); 400X magnification. Note that several of the hsa-miR-302a–positive cells also coexpress BMI1. The small insert on the **top left corner** of **panel F** is a negative control where BMI1 antibody was used together with the scrambled microRNA probe; note the absence of yellow. The **scale bar** in each panel is 75 micrometers.

MIR302/MIR203 asymmetry in a third cohort of 33 BC patients, with either nodal (n = 21) or distant metastasis (n = 12), using microarrays (ArrayExpress accession number E-TABM-971). When compared with the respective primary tumors, both nodal (Supplementary Figure 7A, available online) and distant metastases (Supplementary Figure 7B, available online) had the excess of *MIR302* over *MIR203* typical of pluripotent stem cells (paired *t* test, $P < .001$ and $P = .007$, respectively).

Finally, we evaluated the association between *MIR302* and overall survival in 96 IDC patients. For this group, the median follow-up of patients alive at last follow-up was 75 months. Patients with primary tumor positive to *MIR302* by ISH had statistically significantly shorter OS (Log-rank test, $P = .03$) (Figure 6). The estimated five-year survival rate was 76% for patients with negative *MIR302* status and 57% for those with positive status. As expected, disease and N stages were also associated with OS ($P < .001$, for both). We then performed a multivariable analysis of disease stage, N stage, and miR-302 as prognostic markers for outcome. In these 96 patients, disease stage ($P = .001$) and miR-302 (hazard ratio = .47, 95% confidence interval = .22 to .98, $P = .05$) were left in the final model (Supplementary Table 8, available online). Thus, miR-302 was an independent prognostic marker for OS.

Discussion

The focus of our work was on the possible role of *MIR302* and *MIR203* noncoding genes in breast cancer and in metastasis. The rationale was driven by the hypothesis that a miRNA pluripotency program could become active in some cancer cells within the primary tumor and eventually contribute to metastatic development. To test this hypothesis we initially studied the differentiation of healthy pluripotent stem cells, then the progression of human breast cancer.

In normal tissues, *MIR302* expression was restricted to pluripotent ESC or their closely derived embryonic cell types, and not present in multipotent elements, such as neural stem cells or CD34+ cells. In addition, we determined that pluripotent stem cells, ESCs or IPSCs, were characterized by the concurrent low expression of the *MIR203*. We then showed that invasive breast cancer contained *MIR302*-positive cells. Overexpression of *MIR302* miRNAs in human cancer was previously described only in malignant germ cell tumors (25,26). We speculate that *MIR302* in germ cell tumors is related to its corresponding levels in the healthy embryonic counterpart. Conversely, since there is no *MIR302* expression in normal breast, we propose that the *MIR302*-positive cells are related to breast cancer pathogenesis, and not to its histological origin.

Table 1. Association of *MIR302*-positive ISH with clinical and molecular characteristics in invasive ductal carcinoma (validationTMA cohort)

Characteristic	*MIR302*-positive cases, No. (%) (n = 55)	*MIR302*-negative cases, No. (%) (n = 241)	*P*
Age, y			.83*
Median	48	48	
Range	27–84	28–81	
Female sex	54 (98.2)	241 (100)	.19†
Stage			.03‡
I	4 (7.3)	39 (16.2)	
IIA	19 (34.5)	100 (41.5)	
IIB	17 (30.9)	54 (22.4)	
IIIA	7 (12.7)	28 (11.6)	
IIIB	3 (5.5)	16 (6.6)	
IIIC	5 (9.1)	3 (1.2)	
IV	0 (0)	1 (0.5)	
Grade			.22‡
Low	5 (9.3)	18 (7.5)	
Intermediate	24 (44.4)	141 (59.0)	
High	25 (46.3)	80 (33.5)	
N (lymph node)			.005‡
0	26 (47.3)	154 (63.9)	
1	13 (23.6)	52 (21.6)	
2	14 (25.5)	33 (13.7)	
3	2 (3.6)	2 (0.8)	
Receptor status			
ER-positive	9 (20.9)	44 (28.9)	.44†
PR-positive	10 (23.2)	33 (21.0)	.83†
HER2/neu-positive	19 (52.8)	86 (59.3)	.57†
Triple-negative§	15 (42.1)	54 (37.6)	.70†
P53 positive	17 (51.5)	78 (63.9)	.23†

* Two tailed t test. ER = estrogen receptor; ISH = in situ hybridization; PR = progesterone receptor; TMA = tissue microarray.

† Two-tailed Fisher exact test.

‡ Two-sided Cochran-Armitage test for trend.

§ Cases that are negative for estrogen, progesterone, and HER2/neu receptors.

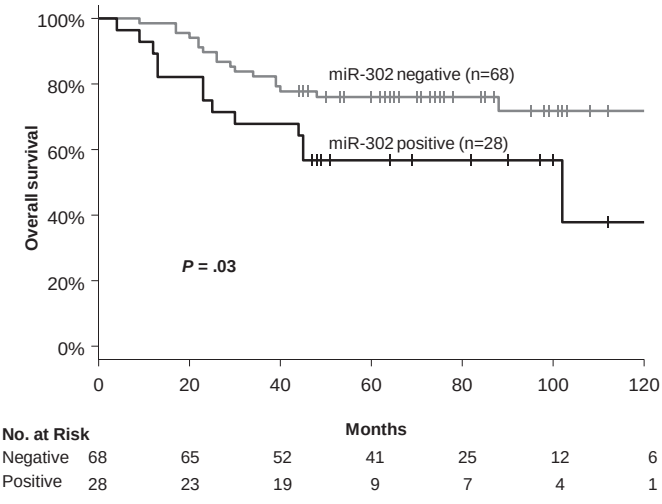

Figure 6. Overall survival in the invasive ductal carcinoma group, according to the *MIR302* expression. Kaplan-Meier curves of overall survival in patients with invasive ductal carcinoma based on the expression of *MIR302* in primary tumors. The patients were divided into hsa-miR-302a–positive and hsa-miR-302a–negative groups. Overall survival (OS) was analyzed by Kaplan-Meier method and two-sided log-rank test.

The *MIR302*-positive cancer cells were also positive for the CD44 stem cell–like mesenchymal cell marker (5,35) and for BMI1, characteristic of breast cancer stem cells (34).

miR-203 is one of the most downregulated miRNAs in solid cancers (36), alongside miR-200c, miR-192 and miR-215, other miRNAs which are also downregulated in ESCs. We showed here that loss of *MIR203* expression in breast tumors and in cancer cell lines is correlated to cancer stem cell and EMT markers, in agreement with what shown by Taube et al. (32).

Most importantly, the *MIR302/MIR203* expression asymmetry typical of pluripotency was present in primary BC associated to lymph node metastasis, and it was further increased in nodal and distant metastases themselves. The activity of *MIR302* in the induction of pluripotency (12,14) and that of *MIR203* as a repressor of p63, a regulator of stem-cell maintenance (20) and an important factor in metastasis (37,38), are both in agreement with the association between the *MIR302/203* asymmetry and metastasis.

Some limitations of our study merit attention. First, the detection of *MIR302* expression in few cells within the primary tumor can be currently performed only by in situ hybridization. We validated the ISH results in an independent cohort using TMA, but this technique might yield false negatives, due to the small size of the tissue cores. Second, since the treatments received by the patients were of various types, it was not possible to determine whether *MIR302* and *MIR203* expression were prognostic markers relating to the natural history of disease. Nevertheless, our results strongly suggest a novel molecular mechanism involved in the metastasis of breast cancer.

In conclusion, cancer cells with *MIR302* expression, within a context of *MIR203* downregulation, were associated with metastasis and shorter survival in breast cancer. The same *MIR302 (high)/MIR203(low)* asymmetry was highly specific of pluripotent stem cells.

References

1. Visvader JE, Lindeman GJ. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat Rev Cancer*. 2008;8(10):755–768.
2. Clevers H. The cancer stem cell: premises, promises and challenges. *Nat Med*. 2011;17(3):313–319.
3. Chiang AC, Massague J. Molecular basis of metastasis. *N Engl J Med*. 2008;359(26):2814–2823.
4. Pece S, Tosoni D, Confalonieri S, et al. Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell*. 2010;140(1):62–73.
5. Mani SA, Guo W, Liao MJ, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*. 2008;133(4):704–715.
6. Morel AP, Lievre M, Thomas C, et al. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS ONE*. 2008;3(8):e2888.
7. Gupta PB, Chaffer CL, Weinberg RA. Cancer stem cells: mirage or reality? *Nat Med*. 2009;15(9):1010–1012.
8. Landgraf P, Rusu M, Sheridan R, et al. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell*. 2007;129(7):1401–1414.
9. Barroso-del Jesus A, Lucena-Aguilar G, Menendez P. The miR-302–367 cluster as a potential stemness regulator in ESCs. *Cell Cycle*. 2009;8(3):394–398.
10. Card DA, Hebbar PB, Li L, et al. *Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. Mol Cell Biol*. 2008;28(20):6426–6438.
11. Rosa A, Brivanlou AH. A regulatory circuitry comprised of miR-302 and the transcription factors OCT4 and NR2F2 regulates human embryonic stem cell differentiation. *EMBO J*. 2011;30(2):237–248.
12. Lin SL, Chang DC, Chang-Lin S, et al. Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. *RNA*. 2008;14(10):2115–2124.
13. Barroso-Deljesus A, Lucena-Aguilar G, Sanchez L, et al. The Nodal inhibitor Lefty is negatively modulated by the microRNA miR-302 in human embryonic stem cells. *FASEB J*. 2011;25(5):1497–1508.
14. Subramanyam D, Lamouille S, Judson RL, et al. Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nat Biotechnol*. 2011;29(5):443–448.
15. Nguyen DX, Bos PD, Massague J. Metastasis: from dissemination to organ-specific colonization. *Nat Rev Cancer*. 2009;9(4):274–284.
16. Liu CG, Calin GA, Volinia S, et al. MicroRNA expression profiling using microarrays. *Nat Protoc*. 2008;3(4):563–578.
17. Volinia S, Visone R, Galasso M, et al. Identification of microRNA activity by Targets' Reverse EXpression. *Bioinformatics*. 2010;26(1):91–97.
18. Sridharan R, Tchieu J, Mason MJ, et al. Role of the murine reprogramming factors in the induction of pluripotency. *Cell*. 2009;136(2):364–377.
19. Kim JB, Sebastiano V, Wu G, et al. Oct4-induced pluripotency in adult neural stem cells. *Cell*. 2009;136(3):411–419.
20. Yi R, Poy MN, Stoffel M, et al. A skin microRNA promotes differentiation by repressing 'stemness.' *Nature*. 2008;452(7184):225–229.
21. Gregory PA, Bert AG, Paterson EL, et al. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol*. 2008;10(5):593–601.
22. Polyak K, Weinberg RA. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer*. 2009;9(4):265–273.
23. Baek D, Villen J, Shin C, et al. The impact of microRNAs on protein output. *Nature*. 2008;455(7209):64–71.
24. Selbach M, Schwanhausser B, Thierfelder N, et al. Widespread changes in protein synthesis induced by microRNAs. *Nature*. 2008;455(7209):58–63.
25. Palmer RD, Murray MJ, Saini HK, et al. Malignant germ cell tumors display common microRNA profiles resulting in global changes in expression of messenger RNA targets. *Cancer Res*. 2010;70(7):2911–2923.
26. Murray MJ, Saini HK, van Dongen S, et al. The two most common histological subtypes of malignant germ cell tumour are distinguished by global microRNA profiles, associated with differential transcription factor expression. *Mol Cancer*. 2010;9:290.
27. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490(7418):61–70.
28. De Craene B, Berx G. Regulatory networks defining EMT during cancer initiation and progression. *Nature Reviews Cancer*. 2013;13(2):97–110.
29. Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res*. 2010;12(5):R68.
30. Marotta LL, Almendro V, Marusyk A, et al. The JAK2/STAT3 signaling pathway is required for growth of CD44(+)CD24(-) stem cell-like breast cancer cells in human tumors. *J Clin Invest*. 2011;121(7):2723–2735.
31. Sheridan C, Kishimoto H, Fuchs RK, et al. CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. *Breast Cancer Res*. 2006;8(5):R59.
32. Taube JH, Malouf GG, Lu E, et al. Epigenetic silencing of microRNA-203 is required for EMT and cancer stem cell properties. *Scientific Reports*. 2013;3:2687.
33. Liu S, Dontu G, Mantle ID, et al. Hedgehog signaling and Bmi-1 regulate self-renewal of normal and malignant human mammary stem cells. *Cancer Res*. 2006;66(12):6063–6071.
34. Shimono Y, Zabala M, Cho RW, et al. Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. *Cell*. 2009;138(3):592–603.
35. Shipitsin M, Campbell LL, Argani P, et al. Molecular definition of breast tumor heterogeneity. *Cancer Cell*. 2007;11(3):259–273.
36. Volinia S, Galasso M, Costinean S, et al. Reprogramming of miRNA networks in cancer and leukemia. *Genome Res*. 2010;20(5):589–599.
37. Adorno M, Cordenonsi M, Montagner M, et al. A Mutant-p53/Smad complex opposes p63 to empower TGFbeta-induced metastasis. *Cell*. 2009;137:87–98.
38. Wellner U, Schubert J, Burk UC, et al. The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat Cell Biol*. 2009;11(12):1487–1495.

Funding

SV was supported by Associazione Italiana Ricerca sul Cancro (IG 13585), Consiglio Nazionale delle Ricerche (EPIGEN), and Ministero della Istruzione Università e Ricerca (PRIN 2010). CMC was funded by National Institutes of Health (NIH) Grant U01 CA152758. KFH was supported by NIH Grant U01 CA154200. AV is supported by Associazione Italiana Ricerca sul Cancro AIRC (IG 11561). MN was supported by the Italian Ministry of University and Research FIRB 2011 (RBAP11BYNP). The preparation and characterization of TMA were supported by U01 CA154200 from the National Cancer Institute and by the Stephanie Spielman Foundation for Breast Cancer Research.

Notes

The study funders had no role in the design of the study, the collection, analysis, or interpretation of the data, the writing of the manuscript, nor the decision to submit the manuscript for publication.

The authors report no potential conflicts of interest.

Additional Contributions: Microarrays and other statistical analysis were performed using R, BioConductor, and BRB-ArrayTools developed by Dr. Richard Simon and BRB-ArrayTools Development Team.

Affiliations of authors: Department of Molecular Virology, Immunology and Molecular Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH (SV, GN, AD, SC, MiG, RAe, RG, GDL, PG, PD, TW, SEW, FP, NZ, HA, KFH, CMC); Biosystems Analysis, LTTA, Deptartment of Morphology, Surgery and Experimental Medicine, University of Ferrara, Ferrara, Italy, (SV, MES, JM, MaG, MM, CZ, FC, MP, JP, MN); Fate Therapeutics, San Diego, CA (RAb); Department of Chemistry, The Scripps Research Institute, La Jolla, CA (CD); Department of Urology, Thomas Jefferson University, Kimmel Cancer Center, Philadelphia, PA (RB); Comprehensive Cancer Center, The Ohio State University, Columbus, OH (KM); Deptartment of Internal Medicine, James Cancer Hospital and Ohio State University Comprehensive Cancer Center, The Ohio State University (CLS); Division of Pathology, II University of Rome "La Sapienza," Ospedale Santo Andrea, Rome, Italy (AV); Department of Surgery, Thomas Jefferson University Medical College, Philadelphia, PA (ALR).