Four miRNAs associated with aggressiveness of lymph node-negative, estrogen receptor-positive human breast cancer

John A. Foekens*, Anieta M. Sieuwerts*, Marcel Smid*†, Maxime P. Look*, Vanja de Weerd*, Antonius W. M. Boersma*, Jan G. M. Klijn*, Erik A. C. Wiemer*, and John W. M. Martens*‡

*Department of Medical Oncology, Erasmus Medical Center Rotterdam, Josephine Nefkens Institute, 3000 CA Rotterdam, The Netherlands; and †Cancer Genomics Centre, 3000 CA, Rotterdam, The Netherlands

Edited by Carlo M. Croce, Ohio State University, Columbus, OH, and accepted by the Editorial Board June 3, 2008 (received for review April 4, 2008)

In this study, we quantified 249 mature micro-RNA (miRNA) transcripts in estrogen receptor-positive (ER^+) primary breast tumors of patients with lymph node-negative (LNN) disease to identify miRNAs associated with metastatic capability. In addition, the prognostic value of the candidate miRNAs was determined in ER^-/LNN breast cancer. Unsupervised analysis in a prescreening set of 38 patients identified three subgroups predominantly driven by three miRNA signatures: an ER-driven luminal B-associated miRNA signature, a stromal miRNA signature, and an overexpressed miRNA cluster located on chromosome 19q23, but these intrinsic miRNA signatures were not associated with tumor aggressiveness. Supervised analysis in the initial subset and subsequent analysis in additional tumors significantly linked four miRNAs (*miR-7*, *miR-128a*, *miR-210*, and *miR-516–3p*) to ER^+/LNN breast cancer aggressiveness ($n = 147$) and one miRNA (*miR-210*) to metastatic capability in ER^-/LNN breast cancer ($n = 114$) and in the clinically important triple-negative subgroup ($n = 69$) (all $P < 0.05$). Bioinformatic analysis coupled *miR-210* to hypoxia/VEGF signaling, *miR-7* and *miR-516–3p* to cell cycle progression and chromosomal instability, and *miR-128a* to cytokine signaling. In conclusion, our work connects four miRNAs to breast cancer progression and to several distinct biological processes involved therein.

micro-RNA | prognosis

Various prognostic mRNA expression signatures have been revealed (1–3), but the forces driving breast cancer aggressiveness are poorly understood. Micro-RNAs (miRNAs) are small nonprotein-coding RNAs (4, 5) that use the endogenous RNA interference pathway to modulate gene expression, thereby contributing to various normal cellular and developmental processes and to malignant transformation and tumor progression (6, 7). Currently, >500 human miRNA genes have been identified (http://microrna.sanger.ac.uk/sequences), and biochemical and *in silico* analyses have suggested that each of these miRNAs targets a plethora of mRNAs. Furthermore, it is likely that miRNAs can act as tumor suppressor genes and oncogenes (5, 7, 8) because they are located in genomic regions that show copy number alterations in a variety of cancer types (9). To make the picture more complete, next to miRNAs, a new class of nonprotein-coding RNAs encoded by transcribed ultraconserved regions exists that might interact with miRNAs and are also located in genomic regions that show gain or loss in human cancers (8). A current model proposes that coding and noncoding RNAs cooperate in the initiation and progression of malignancy (8). Focusing on breast cancer, levels of specific miRNAs vary between normal and malignant breast tissue, tumors of different grade, molecular subtype, and nodal and steroid hormone-receptor status (10, 11), and a few miRNAs have been causally linked to breast cancer proliferation and invasion (12–14). However, convincing data in extensive series of well-defined clinical specimens are currently lacking. In this study, we explored whether miRNAs are associated with prognosis in ER^+/LNN breast cancer. In addition, biological pathways were connected to the identified prognostic miRNAs using global gene expression data also available for the studied specimen. Finally, the prognostic significance of selected miRNAs was also studied in ER^-/LNN breast cancer and in triple-negative breast cancer (i.e., those tumors that are negative for both steroid receptors and do not show HER-2 amplication).

Results

To identify clinically relevant miRNAs associated with time to distant metastasis (TDM) (i.e., tumor aggressiveness) in ER^+/LNN primary breast cancer, we analyzed the miRNA expression levels by real-time RT-PCR in an extensive cohort of breast cancers. None of these patients had received any (neo)adjuvant systemic therapy. We performed a prescreen on 38 samples to identify prognostic candidates among the 249 miRNAs quantitated by real-time PCR. Unsupervised Pearson correlation clustering using the 75% most variable miRNAs subdivided these 38 samples into three groups (Fig. 1*A*). Careful examination of the cluster arms showed that three "intrinsic" miRNA signatures drove this subdivision [supporting information (SI); Fig. S1*a*]. To further explore these intrinsic miRNA signatures, we re-clustered the samples using these three miRNA signatures separately and associated clinical and biological information with the generated subgroups. This revealed that miRNA signature 1 was dominated by a single co-expressed cluster of miRNAs [i.e., miR-515–522, located on chromosome 19q23 (Fig. 1*B*)]. The second miRNA signature (Fig. 1*C*) consisted of stroma-specific miRNAs because these miRNAs were predominantly expressed in stromal fibroblast and endothelial cells and marginally in the epithelial breast cancer cell lines (Fig. 1*D*). Furthermore, their expression correlated with the stromal content of the specimens (Fig. 1*E*; Mann–Whitney test, $P = 0.04$). The third intrinsic miRNA signature (Fig. 1*F*) associated with ER mRNA expression (Fig. 1*G*; Mann–Whitney test, $P = 0.0006$) separated luminal B from other intrinsic subtypes of breast cancer (Fig. 1*G*; Fisher's exact test, $P = 0.0002$) and specifically identified ER^+ breast cancer cell lines (Fig. 1*H*) among the epithelial breast cancer cell lines and stromal endo-

Author contributions: J.A.F., A.M.S., J.G.M.K., E.A.C.W., and J.W.M. designed research; J.A.F., A.M.S., M.S., M.P.L., V.d.W., A.W.M.B., E.A.C.W., and J.W.M.M. performed research; J.A.F. contributed new reagents/analytic tools; A.M.S., M.S., M.P.L., A.W.M.B., and J.W.M.M. analyzed data; and J.A.F., A.M.S., M.S., M.P.L., J.G.M.K., E.A.C.W., and J.W.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

‡To whom correspondence should be addressed at: Department of Medical Oncology, Josephine Nefkens Institute, Erasmus MC, PO Box 2040, 3000 CA Rotterdam, The Netherlands. E-mail: j.martens@erasmusmc.nl.

This article contains supporting information online at www.pnas.org/cgi/content/full/0803304105/DCSupplemental.

© 2008 by The National Academy of Sciences of the USA

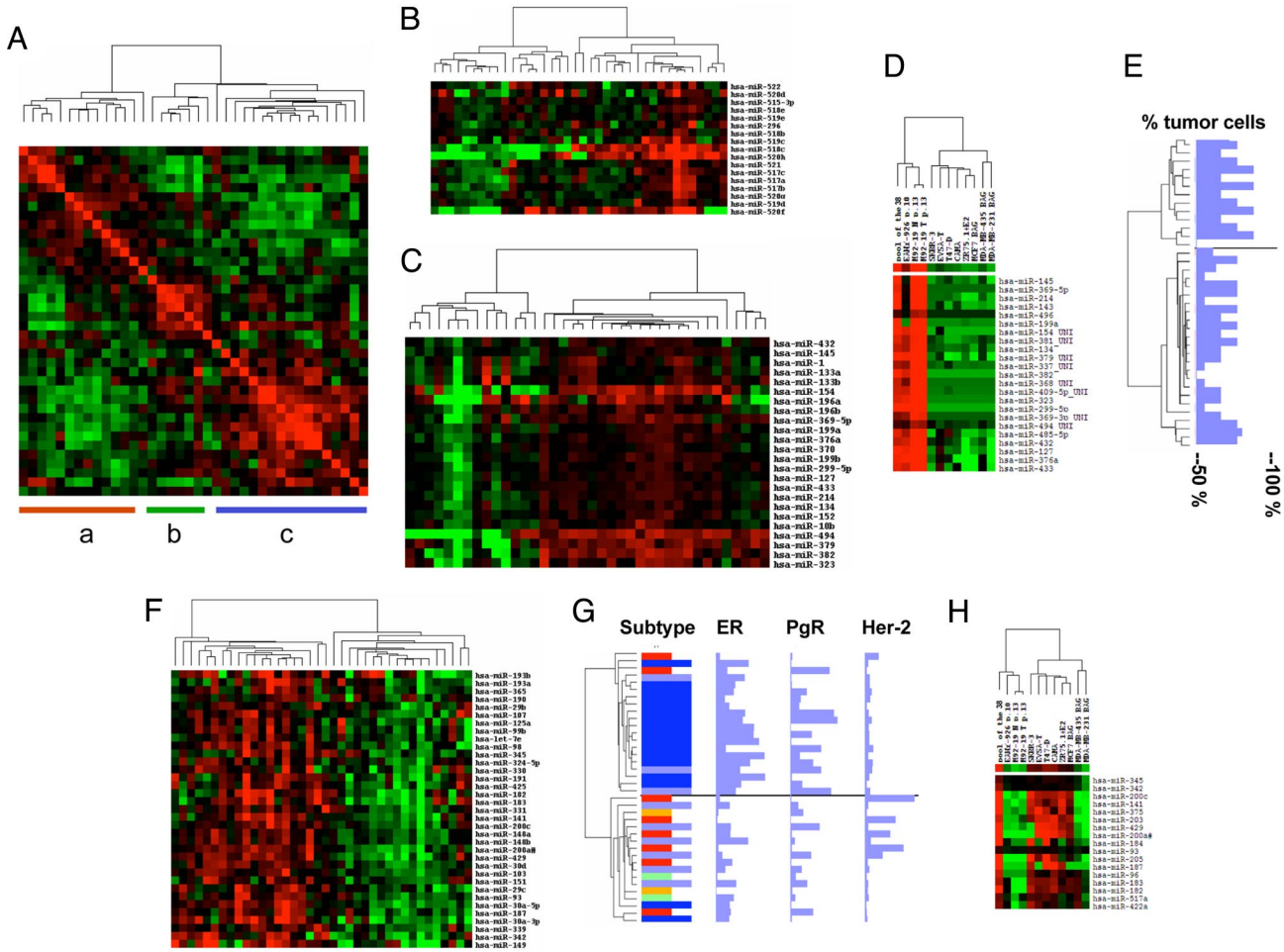

Fig. 1. Exploring the miRNA expression profile of 38 ER$^+$/LNN breast cancer specimens screened for the expression of 249 miRNAs. (*A*) Pearson correlation plot of the specimens based on the most variable miRNAs representing the top 75% of those measurable, revealed three main clusters marked a, b, and c. Red indicates a strong positive correlation, whereas green indicates a weak positive correlation between samples. Three miRNA signatures (1, 2, and 3) were identified driving the unsupervised hierarchical clustering (see Fig. S1*a*; colors of the subgroups a, b, and c match between Fig. 1*A* and Fig. S1*a*). (*B*, *C*, *F*) Hierarchical clustering of specimens based on the three dominant miRNA signatures separately. (*D*) MiRNAs specifically expressed in the stromal cell lines. (*E*) Tumor content of tumor samples with the sample groups after clustering using miRNA signature 2. (*F*) Cluster arm of *c* is presented transposed, with (*Right*) the tumor content (relative to stromal content) observed in the hematoxylin-eosin–stained sections of the samples. Shown is a part of the hierarchical clustering plot of breast cancer, endothelial and fibroblast cell lines that represent the miRNAs specifically expressed in the stromal cell lines. The complete clustering plot of the cell lines mentioned is shown in Fig. S1*b*. (*G*) Biological features (intrinsic subtypes and mRNA expression of ER, PgR, and Her-2) associated with the sample groups after clustering using miRNA signature 3. The cluster arm of *F* is presented transposed, with the subtype of the samples (*right*) and the quantitative mRNA expression levels of ER, PgR, and Her-2. For the subtypes, dark blue represents luminal B, light blue luminal A, red Her-2, green normal-like, and orange basal tumors. (*H*) MiRNAs specifically expressed in the ER$^+$ breast cancer cell lines. Shown is a part of hierarchical clustering plot that represents the miRNAs more abundantly expressed in the ER$^+$ cell lines. For the complete clustering plot, see Fig. S1*b*. Note the significant overlap (<60%) between the miRNAs present in *C* and *D* and in *F* and *H*. For *B*, *C*, *D*, *F*, and *H*, red color indicates high expression level and green color indicates low expression level of the indicated miRNA in a particular specimen.

thelial and fibroblast cells studied (see Fig. S1*b* for complete clustering diagram). Several miRNAs in this latter signature have previously been linked to ER and PgR status in breast cancer (10). This indicates that also within ER$^+$ breast tumors, miRNA expression is influenced by steroid hormones or their receptors, as is the case for many ordinary coding messengers (15). Although we found specific biological features connected to these three intrinsic miRNA signatures, they were not linked to breast cancer aggressiveness.

Supervised analysis in this prescreening set of 38 ER$^+$/LNN breast tumors, on the other hand, suggested that several miRNAs were differentially expressed between patients with a short TDM versus those with a long TDM. Twelve candidate miRNAs most differentially expressed between the two groups, *miR-7*, *miR-22*, *miR-34b*, *miR-128a*, *miR-145*, *miR-151*, *miR-193b*, *miR-205*, *miR-210*, *miR-449*, *miR-489*, and *miR-516–3p*, were further explored by qRT-PCR in the remaining 147 ER$^+$ cases and in 114 available ER$^-$/LNN primary breast cancer cases. Patient and tumor characteristics in the complete cohort, including the 38 initial cases, and their associations with the levels of the 12 candidate miRNAs, are summarized in Table S1. Of note are the positive association of *miR-7*, *miR-34b*, and *miR-151* with tumor size ($P < 0.05$); the association of *miR-7*, *miR-210*, *miR-489*, and *miR-516–3p* with pathological grade ($P < 0.01$); and the positive association of *miR-34b*, *miR-145*, *miR-193b*, *miR-449*, and *miR-489* with each other ($n = 298$, Spearman rank correlation, $P < 0.001$; not shown) and with steroid hormone receptor levels ($P < 0.001$).

In the independent 147 ER$^+$/LNN samples, only levels of *miR-7*, *miR-128a*, *miR-210*, and *miR-516–3p* were significantly

associated with a shorter TDM in univariate and multivariate Cox regression analysis corrected for traditional prognostic factors (all $P < 0.05$ using a false discovery rate of 10%; see Table S2 and Kaplan-Meier survival curves in Fig. S2 *a–d*). Combining the information of these four prognostic miRNAs either by using hierarchical clustering (Fig. 2*A*; $P = 0.002$) or by building a Cox model (Fig. 2*B*; $P < 0.0001$) showed that this panel of miRNAs significantly identified subgroups with a very different TDM.

In the 114 ER$^-$/LNN breast cancer cases, of the 12 candidate prognostic miRNAs identified in ER$^+$ disease, only higher *miR-210* levels were in univariate analysis significantly related to shorter TDM (hazard ratio = 1.72; 95% confidence interval: 1.12–2.64; $P = 0.01$). Equally sized subgroups based on the *miR-210* expression levels showed a significant different prognosis in the whole ER$^-$ subgroup (see Fig. S2*e*; $P = 0.0026$) and in the triple-negative (ER$^-$/PgR$^-$/Her-2$^-$) subgroup (Fig. 2*C*; $P = 0.05$).

By using global testing (16), the availability of transcriptome information on these same specimens allowed us to associate biological pathways with the four prognostic miRNAs (see Table S3 for all top ranking pathways). We restricted the analysis to the ER$^+$ tumors, the primary focus of our study. *MiR-210* was most significantly coupled to "VEGF, hypoxia and angiogenesis" (comparative $P = 0.015$), in which *HRAS*, *PTK2*, *SHC1*, and *HIF1A* were overexpressed and *KDR* was under-expressed in tumors having high versus those having low *miR-210* expression (Fig. 3*A*). Furthermore, based on two published hypoxia gene expression signatures (17, 18), hypoxic tumors had higher *miR-210* expression levels than normoxic tumors (both $P < 0.001$; see Fig. S3). *MiR-7* and *miR-516–3p* were most significantly associated with cell cycle-related events (*miR-7* to "G2-to-M cell cycle checkpoint" [Fig. 3*B*, comparative $P = 0.012$]; *miR-516–3p* to "cyclins and cell cycle regulation" [Fig. 3*C*, comparative $P = 0.002$]). Cell cycle checkpoint deregulation in cancer has a link to chromosomal instability. Indeed, samples that had predicted chromosomal instability based on a published chromosomal instability gene signature (19) had high *miR-7* ($P < 0.001$) and high *miR-516–3p* ($P = 0.008$) expression (see Fig. S3). Finally, *miR-128a* was inversely related to various IL (IL-17 and IL-4), lymphocytic and inflammation signaling pathways (Fig. 3*D*; only the IL-17 pathway is shown; comparative $P < 0.05$ in all cases).

Discussion

miRNAs are an emerging group of ribonucleotides, and their clinical value in breast cancer is still largely unknown. The implications of our findings in the current study are twofold. Unsupervised analysis of miRNA expression data in 38 ER$^+$ clinical specimens screened for the entire panel of 249 miRNAs identified three miRNA signatures that we arbitrarily labeled intrinsic miRNA signatures. Interestingly, one signature was co-expressed in luminal B intrinsic subtype cancers, indicating, in line with suggestions of others (11), that intrinsic molecular subtypes based on mRNA expression (20, 21) have specific miRNAs co-expressed. The second miRNA signature was abundantly expressed in stromal cell lines, and is thus likely to be expressed in tumor stroma. These miRNAs might thus prove to be useful tumor-specific stromal markers. The third signature identified a single cluster of miRNAs located on chromosome 19q23, a region that does not show any amplification in the breast cancer specimen in our study cohort (unpublished data), suggesting these miRNAs are expressed in a co-regulated manner. Although the observations above are interesting, their clinical significance remains to be determined and requires additional validation.

The supervised approach, however, identified among 249 miRNAs several candidate prognostic miRNAs. These were further studied in a significant number of well characterized clinical specimens, leading to the identification of four miRNAs

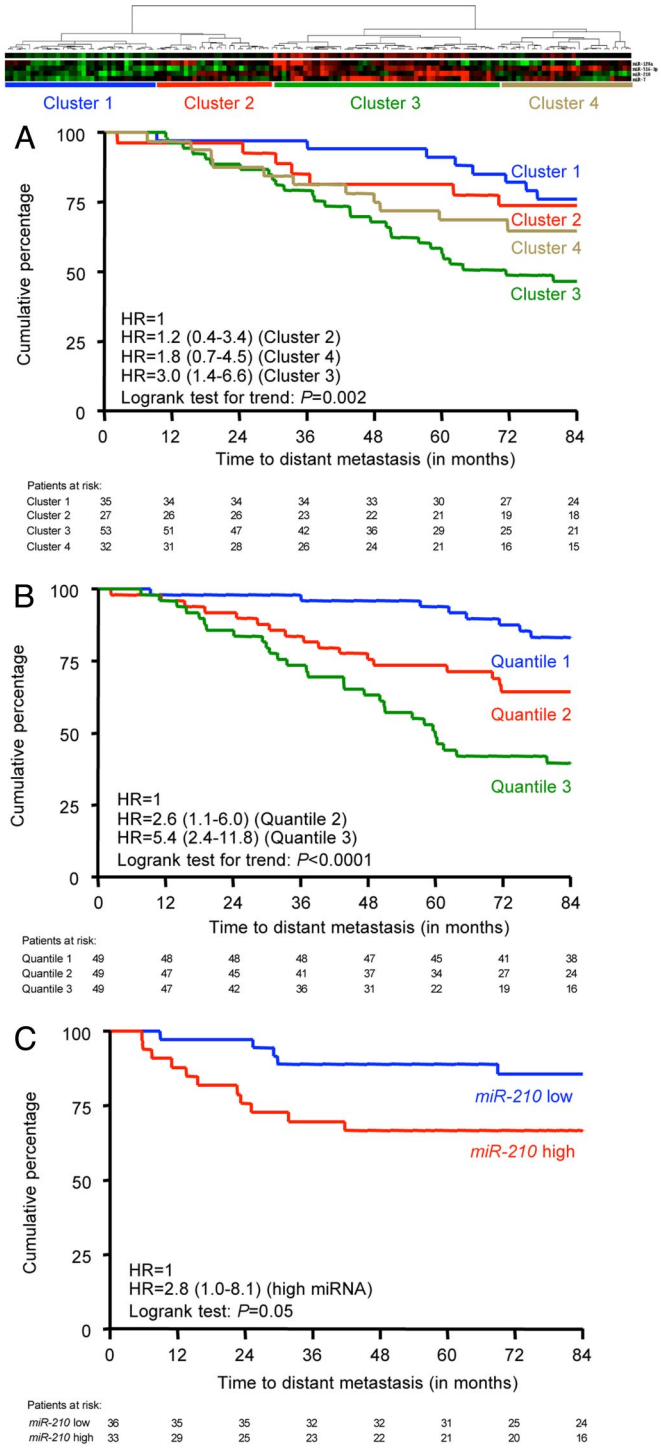

Fig. 2. Kaplan–Meier TDM curves of primary breast cancer patients based on miRNA expression levels. Kaplan–Meier TDM curves for 147 ER$^+$/LNN cases as a function of four subgroups (clusters 1–4) revealed by hierarchical clustering of the samples based on the levels of *miR-7*, *miR-128a*, *miR-210*, and *miR-516-3p* (clustering plot is shown above the Kaplan–Meier curve) (*A*) and as a function of three equally sized subgroups (quantiles 1–3) based on a score of the coefficients of the multivariable Cox regression model of *miR-7*, *miR-128a*, *miR-210*, and *miR-516-3p* as continuous variables (*B*). (*C*) Kaplan–Meier TDM curves for 69 triple-negative (ER$^-$/PgR$^-$/Her-2$^-$) breast cancer cases as a function of subgroups dichotomized at the median *miR-210* expression level. Patients at risk at the indicated time intervals are shown at the bottom of each graph.

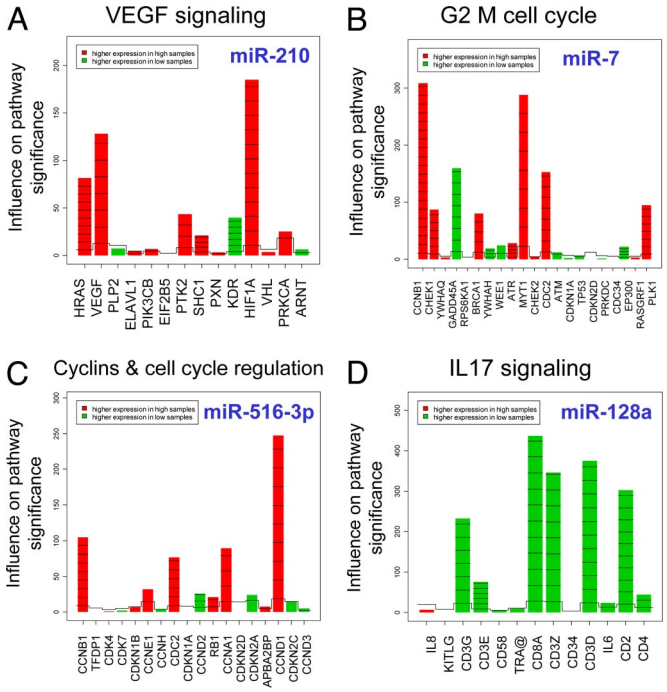

Fig. 3. Pathways associated with clinically relevant miRNAs. Pathway analysis plots of Biocarta pathways most strongly associated with *miR-210* (*A*), *miR-7* (*B*), *miR-516–3p* (*C*), and *miR-128a* (*D*). A pathway plot [generated by the Global Test program (16)] shows the genes annotated to the indicated pathway in Biocarta with their association with tumors showing a low versus high expression of the indicated miRNA. Each bar represents a gene in the pathway. The height of the bar indicates the contribution (influence) of each individual gene to the significance of the pathway. Horizontal markers in a bar indicate 1 SD away from the reference point, and two or more horizontal lines in a bar indicate a statistically significant association of the corresponding gene with the subgroups identified by a specific miRNA. Red and green colors indicate positive and negative associations with the tumor subgroup having high versus low expression of a particular miRNA.

(*miR-7*, *miR-128a*, *miR-210*, and *miR-516–3p*) associated with tumor aggressiveness in ER^+/LNN and one miRNA (*miR-210*) linked with early relapse in ER^-/LNN breast cancer. *MiRNA-210* was also significantly associated with poor outcome in the 69 triple-negative breast cancers present among the specimens studied, and this miRNA may prove to be a valuable marker, because strong prognostic markers in this clinically important breast cancer subgroup are sparsely available.

Noteworthy is that our large-scale screen in clinical specimens did not single out miRNAs *miR-10b*, *miR-126*, *miR-335*, *miR-373*, and *miR-520c*, which were recently functionally linked to breast cancer migration and invasion in animal models and subsequently linked to progression of clinical breast cancer (12–14). Even though these miRNAs were detectable by our method, some of them may be more relevant in ER^- or other specific subtypes of breast cancer. Alternatively, they could have remained unrevealed in our study because of the notorious heterogeneous nature of breast cancer. On the other hand, our study did confirm the clinical significance of *miR-210*, which was just recently connected to breast cancer outcome in a heterogeneous cohort of 219 breast cancer patients (22), which, unlike ours, also included patients with nodal involvement and those having received various types of adjuvant treatment. That previous study does not necessarily associate *miR-210* with pure prognosis of breast cancer as was established in our current study, because the observed effects may have been confounded by the treatments given.

The potential biological significance of our identified miRNAs in breast cancer was revealed by studying co-expressed messenger RNAs. This way, *miR-210* was connected to hypoxia in agreement with the fact that *miR-210* expression is induced under hypoxic conditions in breast and ovarian cancer cell lines and is under direct transcriptional control of the transcription factor HIF-1α (22–24). The independent conformation by several groups that *miR-210* has a link with hypoxia confirms that global testing of co-expressed messengers grouped in biological pathways is a feasible approach. In a similar way, *miR-7* and *miR-516–3p,* which both were associated with pathological grade, were linked to cell cycle deregulation, the pathway currently most frequently linked to prognosis, particularly in ER^+ breast cancer (1, 3, 25, 26). Finally, *miR-128a* was inversely associated with genes usually expressed in lymphocytes, suggestive of a role of this miRNA in lymphocytic infiltration or immune response in breast cancer metastasis (26).

In conclusion, this study provides a leap forward because it has connected expression levels of four miRNAs and associated pathways with breast tumor aggressiveness. Only *miR-210* was also associated with clinical outcome in ER^-/LNN breast cancers, including the clinically important triple-negative subgroup. Finally, if the miRNAs identified here, which are all overexpressed in tumors that relapse early, are causally linked to breast cancer progression, silencing them has great therapeutic potential and may guide the biotechnical development of a new type of targeted drugs (27, 28).

Materials and Methods

Patients' Samples. The institutional medical ethics committee approved the current study (medical ethics committee number 02·953), which was performed in accordance with the Code of Conduct of the Federation of Medical Scientific Societies in The Netherlands (http://www.fmwv.nl/) and which is reported according to the REMARK guidelines (29). The 344-LNN patient cohort from which our 299 samples were selected because of the availability of sufficient good-quality miRNA and Affymetrix U133A mRNA expression data has been described in detail previously (2, 26). None of the patients had received any systemic (neo)adjuvant therapy. From 185 ER^+ and 114 ER^- tumors, sufficient total RNA was present to perform the miRNA expression profiling analysis. In the prescreen, 38 ER^+ samples were analyzed, leaving 147 ER^+ cases in which the miRNA candidates from the prescreen were subsequently studied. For one prescreen sample, we ran out of RNA during the study, leaving 184 ER^+ cases for the final candidate analysis. For these 184 ER^+ and 114 ER^- cases (298 in total), the patient and tumor characteristics are presented in detail in Table S1. The median age of the patients in the study at surgery was 52 years (range: 26–83 years). Two hundred seventeen patients (72%) had undergone breast-conserving surgery, and the remaining patients were treated with radical mastectomy. Time to distant metastasis was defined as the time between surgery and the development of distant metastasis. Median follow-up time of patients alive was 99 months (range: 20–186 months). One hundred one (34%) patients developed a distant metastasis, which counted as an event in the analysis for TDM. Patients who died without evidence of disease were censored at last follow-up in the analysis of TDM.

Tissue Processing, RNA Isolation, Multiplex and Uniplex cDNA Synthesis, and Quantification of miRNA Species. Tissue processing, total RNA isolation, and total RNA quality control checks have been described elsewhere (30, 31). The TaqMan Human MicroRNA Assay Set (release date March 2006; see Table S4; Applied Biosystems [ABI]), consisting of 259 unique assays to quantify 249 miRNAs and 10 controls (Z30 and nine different SNORs/RNUs), was used to screen a set of 38 ER^+ breast cancer specimens for differentially expressed miRNAs. The specimens were selected from the ER^+ cases in the cohort at random with the restriction that 19 were from patients who developed a distant metastasis within 3 years and 19 were from patients who did not develop a distant metastasis within 7 years. The use of a specific primer with a hairpin structure during cDNA synthesis and mature miRNA specific detection probes precluded the detection of precursor miRNAs (ABI). For analysis of the remaining cohort with the 12 selected miRNAs, the individual TaqMan human MicroRNA Assay kits from ABI were used. In brief, up to 48 different RT primers (250 nM each) were pooled according to the recommendations of ABI, concentrated for 30 min in a speed vacuum centrifuge at 50°C and resuspended in nuclease-free ddH_2O to a final concentration of 50 nM each.

Twenty-five to 50 ng of total RNA sample aliquots were reverse-transcribed in a final volume of 20 μl with a final concentration of 12.5 nM for each RT primer using the TaqMan MicroRNA Reverse Transcription Kit (ABI) according to the manufacturer's instructions. To verify that the multiplex RT approach did not affect the quantification of specific miRNAs, some RT primers were included in different RT pools and expression levels of a number of the more lowly expressed miRNAs were validated in a uniplex RT reaction. These experiments showed that especially for the lower expressed miRNAs, some gain in sensitivity could be achieved in a uniplex RT reaction compared with a multiplex RT reaction. We therefore decided to assess expression levels of all the low expressed miRNAs (requiring >30 cycles before passing the threshold level) in a uniplex RT format.

Before performing a duplicate real-time PCR for each of the miRNAs separately, RT samples were diluted in nuclease-free ddH_2O and amplified in a final volume of 20 μl containing cDNA synthesized from 0.25 ng of total RNA, 6 μl of TaqMan universal PCR master mix without UNG, and 6 μl of the individual TaqMan MicroRNA primer and probe assays. Real-time PCR was performed in an Mx3000P Real-Time PCR System (Stratagene) with cycling conditions according to the manufacturer's instructions. A pool consisting of RNA from all 38 human breast tissues included in the prescreening set was included in each cDNA synthesis run, and the resulting data were used to normalize for in-between experimental variations. In addition, all cDNA synthesis runs incorporated a minus RT reaction, which proved to be negative for all assays in this study. PCR efficiency, linearity, and the upper and lower detection limits of each of the individual miRNA assays were validated with a standard curve constructed from a simultaneously run serially diluted pool of cDNA from all samples included in the particular run. To check for RNA input, integrity, and yield of amplifiable cDNA, the reverse primer (*HPRT1-R*) of the medium abundant reference gene hypoxanthine-guanine phosphoribosyltransferase (*HPRT1*) was used in a uniplex cDNA synthesis reaction, followed by amplification of the *HPRT1* gene: *HPRT1-F*, 5′-TTC CTT GGT CAG GCA GTA TAA TCC-3; *HPRT1-R*, 5′-GGT CCT TTT CAC CAG CAA GCT-3; *HPRT1* probe, FAM-[CTG GCT TCA CCA TCG]-BHQ1, TAMRA (Sigma Genosys). Because appropriate reference molecules for small RNAs are still unknown for clinical breast cancer, the most stable miRNA in our prescreening set of 38 ER^+ breast cancer specimens was calculated with geNorm (32) and NormFinder (33) with the Datan Framework GenEx Pro package version 4.3.2 and was shown to be *miR-132*. In addition, geNorm and NormFinder identified *miR-374* as second best. Reassuring is that we found no correlation between these two miRNAs and ER-protein levels (Spearman $r_s = -0.03$ for *miR-374* and 0.11 for *miR-132*, respectively) which is the most important confounding factor in clinical breast cancer. Thus, we selected *miR-132* (median Ct: 26.71) and *miR-374* (median Ct: 25.24) to normalize the raw expression data of the 213 different miRNAs that could be reliably assessed in our prescreening breast tumor cohort (median Ct: 27.94, range: 20.22–39.55). Finally, *miR-374* and *miR-132* turned out to be the least variable miRNAs in the extended series that were analyzed for the 12 candidate miRNAs. Concentrations of the miRNAs, expressed relative to our miRNA reference set (*miR-132* and *miR-374*), were quantified as follows: miRNA target $= 2^{(\text{average Ct miR-132 + miR-374 } - \text{ mean Ct target miR})}$.

Statistics and Bioinformatics. Differences in expression levels between groups were assessed with the Mann–Whitney *U* test or Kruskal–Wallis test. In these tests, patient and tumor characteristics were used as grouping variables. The Fisher's exact test was used to test for associations between categorized variables. The associations between continuous variables were tested using Spearman rank correlation (r_s). In the subset of 38 ER^+/LNN primary breast cancers, the Student's *t* test was used to identify differentially expressed miRNAs between the patients with a short or long TDM. Because this prescreening aimed at identifying candidate miRNAs for confirmation in independent samples, a *P* of 0.1 without correcting for multiple comparisons was used. In the remaining 147 ER^+ and 114 ER^- cases, the prognostic value of the clinical and biological variables, with TDM as the endpoint in the univariate and multivariable analyses, was investigated with the Cox proportional hazards model, and the hazard ratio and its 95% confidence interval were estimated (34). To adjust for multiple comparisons, univariate *P*-values were considered significant using a false discovery rate of 10% (35). For this regression analysis, miRNA expression levels were either log- or Box-Cox–transformed to reduce skewness. To combine the information of multiple miRNAs, we calculated an index using the estimates from the Cox model. Kaplan-Meier survival plots, log-rank tests, and log-rank tests for trend were used to assess the differences in TDM of the different predicted risk groups of patients. Triple-negative breast cancer (i.e., ER^-, PgR^-, non-HER2 amplified), was defined with cutpoints for ER and PgR protein as described, respectively (2, 36). Her-2 mRNA levels, as determined by real-time PCR (37), displayed a clear bimodal distribution; therefore, the cutpoint for being Her-2 amplified or not was determined as described (38) and defined at 18.01355. All *P*-values are two-sided and, except for the prescreen, a $P \leq 0.05$ was considered statistically significant. Regression analysis and association studies were performed using the statistical package STATA, release 10 (Stata).

Pathway Analysis. Affymetrix microarray gene expression data (HG-U133A chips) previously deposited in the NCBI/GEO database (entries GSE2034 and GSE5327) were available from all 184 ER^+ patients and 114 ER^- patients. As done previously (2, 26), gene expression signals were calculated using Affymetrix GeneChip analysis software MAS 5.0. Global scaling was performed to bring the average signal intensity of a chip to a target of 600 before data analysis. Samples were split according to their *ESR1* mRNA expression level (39). Within each sample set, the intensities for each probe set were thresholded at 30 and were then expressed relative to the geometric mean of that probe set and 2-log transformed. A sample was labeled ''high'' or ''low'' for an miRNA if it belonged to the top 25% or bottom 25% of all samples according to the expression levels of that particular miRNA. Thus, the 25% ($n = 46$) ER^+ samples with the highest *miR-7* expression levels were compared with the 46 ER^+ samples with the lowest *miR-7* expression levels, and likewise for each miRNA of interest. The Global Test program (16) was used (version 4.2.0) to relate Biocarta pathways (http://www.biocarta.com/) to samples expressing high or low levels of a particular miRNA, irrespective of their association with TDM. All *P*-values were corrected for multiple testing and checked by re-sampling at 1000 times to ensure that an equally sized and randomly chosen group of genes would not also be significant. Pathways were considered of interest if the P value of the Global Test after correcting for multiple testing and the re-sampling P value were both below 0.05. Pathway *P*-values were two-sided and corrected for multiple testing. The contribution of individual genes in a pathway was evaluated using *z*-scores calculated by the Global Test program. Genes with *z*-scores >1.96 were considered significant contributors to the pathway. R version 2.4.1 (http://www.cran.org) was used to run the Global Test package.

ACKNOWLEDGMENTS. We thank Marion E. Meijer-van Gelder for the collection of follow-up information and Mieke Timmermans for analysis of tissue sections. This study was financially supported by the Association for International Cancer Research (Grant 07–0609) and The Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO).

1. Van 't Veer LJ, *et al.* (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415:530–536.
2. Wang Y, *et al.* (2005) Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 365:671–679.
3. Sotiriou C, *et al.* (2006) Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 98:262–272.
4. Calin GA, Croce CM (2007) Chromosomal rearrangements and microRNAs: a new cancer link with clinical implications. *J Clin Invest* 117:2059–2066.
5. Calin GA, Croce CM (2006) MicroRNA signatures in human cancers. *Nat Rev Cancer* 6:857–866.
6. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116:281–297.
7. Caldas C, Brenton JD (2005) Sizing up miRNAs as cancer genes. *Nat Med* 11:712–714.
8. Calin GA, Liu CG, Ferracin M, *et al.* (2007) Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. *Cancer Cell* 12:215–229.
9. Zhang L, *et al.* (2006) MicroRNAs exhibit high frequency genomic alterations in human cancer. *Proc Natl Acad Sci USA* 103:9136–9141.
10. Iorio MV, *et al.* (2005) MicroRNA gene expression deregulation in human breast cancer. *Cancer Res* 65:7065–7070.
11. Blenkiron C, *et al.* (2007) MicroRNA expression profiling of human breast cancer identifies new markers of tumour subtype. *Genome Biol* 8:R214.
12. Ma L, Teruya-Feldstein J, Weinberg RA (2007) Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* 449:682–688.
13. Tavazoie SF, *et al.* (2008) Endogenous human microRNAs that suppress breast cancer metastasis. *Nature* 451:147–152.
14. Huang Q, *et al.* (2008) The microRNAs miR-373 and miR-520c promote tumour invasion and metastasis. *Nat Cell Biol* 10:202–210.
15. Oh DS, *et al.* (2006) Estrogen-regulated genes predict survival in hormone receptor-positive breast cancers. *J Clin Oncol* 24:1656–1664.
16. Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC (2004) A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics* 20:93–99.
17. Winter SC, *et al.* (2007) Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. *Cancer Res* 67:3441–3449.
18. Chi JT, *et al.* (2006) Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. *PLoS Med* 3:e47.
19. Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z (2006) A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat Genet* 38:1043–1048.

20. Perou CM, *et al.* (2000) Molecular portraits of human breast tumours. *Nature* 406:747–752.
21. Smid M, *et al.* (2008) Subtypes of breast cancer show preferential site of relapse. *Cancer Res* 68:3108–3114.
22. Camps C, *et al.* (2008) hsa-miR-210 is induced by hypoxia and is an independent prognostic factor in breast cancer. *Clin Cancer Res* 14:1340–1348.
23. Kulshreshtha R, *et al.* (2007) A microRNA signature of hypoxia. *Mol Cell Biol* 27:1859–1867.
24. Giannakakis A, *et al.* (2007) miR-210 links hypoxia with cell cycle regulation and is deleted in human epithelial ovarian cancer. *Cancer Biol Ther* 7:255–264.
25. Yu K, *et al.* (2004) A molecular signature of the Nottingham prognostic index in breast cancer. *Cancer Res* 64:2962–2968.
26. Yu JX, *et al.* (2007) Pathway analysis of gene signatures predicting metastasis of node-negative primary breast cancer. *BMC Cancer* 7:182–195.
27. Volinia S, *et al.* (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci USA* 103:2257–2261.
28. Negrini M, Ferracin M, Sabbioni S, Croce CM (2007) MicroRNAs in human cancer: from research to therapy. *J Cell Sci* 120:1833–1840.
29. McShane LM, *et al.* (2005) Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst* 97:1180–1184.
30. Sieuwerts AM, *et al.* (2005) How ADAM-9 and ADAM-11 differentially from estrogen receptor predict response to tamoxifen treatment in patients with recurrent breast cancer: a retrospective study. *Clin Cancer Res* 11:7311–7321.
31. Sieuwerts AM, *et al.* (2006) Which cyclin E prevails as prognostic marker for breast cancer? Results from a retrospective study involving 635 lymph node-negative breast cancer patients. *Clin Cancer Res* 12:3319–3328.
32. Vandesompele J, *et al.* (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:R34.1–R34.11.
33. Andersen CL, Jensen JL, Orntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* 64:5245–5250.
34. Cox DR, Oakes D (1984) *Analysis of Survival Data* (Chapman and Hall Ltd, London).
35. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Ser B* 57:289–300.
36. Foekens JA, *et al.* (1989) Prognostic value of estrogen and progesterone receptors measured by enzyme immunoassays in human breast tumor cytosols. *Cancer Res* 49:5823–5828.
37. Van Agthoven A, *et al.* (2008) Relevance of BCAR genes in human breast cancer progression and tamoxifen. *J Oncol*, in press.
38. Urban P, *et al.* (2006) Increased expression of urokinase-type plasminogen activator mRNA determines adverse prognosis in ErbB2-positive primary breast cancer. *J Clin Oncol* 24:4245–4253.
39. Foekens JA, *et al.* (2006) Multicenter validation of a gene expression-based prognostic signature in lymph node-negative primary breast cancer. *J Clin Oncol* 24:1665–16671.