A Novel MAPK–microRNA Signature Is Predictive of Hormone-Therapy Resistance and Poor Outcome in ER-Positive Breast Cancer

Philip C. Miller1,2, Jennifer Clarke1,2,3,4,5, Tulay Koru-Sengul1,4, Joeli Brinkman1, and Dorrya El-Ashry1,2,3

Abstract

Purpose: Hyperactivation of ERK1/2 MAPK (hMAPK) leads to loss of estrogen receptor (ER) expression and poor outcome in breast cancer. microRNAs (miRNA) play important regulatory roles and serve as biomarkers of disease. Here, we describe molecular, pathologic, and clinical outcome associations of an hMAPK–miRNA expression signature in breast cancer.

Experimental Design: An hMAPK–miRNA signature was identified, and associations of this signature with molecular and genetic alterations, gene expression, pathologic features, and clinical outcomes were determined in primary breast cancers from training data and validated using independent datasets. Univariate and multivariate analyses identified subsignatures associated with increased disease recurrence and poorer disease survival among ER-positive (ER+) patients, respectively.

Results: High-hMAPK–miRNA status significantly correlated with ER-negativity, enrichment for basal and HER2-subtypes, and reduced recurrence-free and disease-specific survival in publicly available datasets. A robust determination of a recurrence signature and a survival signature identified hMAPK-miRNAs commonly associated with poor clinical outcome, and specific subsets associated more closely with either disease recurrence or disease survival, especially among ER+ cancers of both luminal A and luminal B subtypes. Multivariate analyses indicated that these recurrence and survival signatures significantly associated with increased risk of disease-specific death and disease recurrence in ER+ cancer and ER+ cancers treated with hormone therapy.

Conclusions: We report an hMAPK–miRNA signature and two subsignatures derived from it that associate significantly with adverse clinical features, poor clinical outcome, and poor response to hormone therapy in breast cancer, thus identifying potential effectors of MAPK signaling, and novel predictive and prognostic biomarkers or therapeutic targets in breast cancer.

Introduction

Treatment decisions in breast cancer are informed not only by tumor grade, stage, and lymph node status, but also by genomic biomarkers, estrogen receptor (ER), and HER2. Increasingly, multigene assays such as Oncotype DX and Mammaprint are being used in prognostic capacities for patients with breast cancer (1, 2). ER+ tumors have better overall prognosis and response to hormonal therapy, while ER–tumors are more aggressive, resistant to hormonal therapies, and frequently display elevated expression of growth factor receptors such as erbB family members, EGFR, or HER2 (3–5). Overexpression or amplification of HER2 occurs in approximately 25% of breast cancers, while EGFR overexpression occurs in up to 30% of breast cancers. Both are poor prognostic indicators independent of ER status; EGFR expression has an inverse correlation with ER and is more common in triple-negative and basal breast cancers (6–8), while HER2 amplification tends to correlate with ER-negativity or lower levels of ER expression. Downstream effector pathways activated by erbB family signaling include the RAS–RAF–MEK–ERK, PI3K–AKT–MTOR, and the MEKK1–SEK1–JNK pathways. Over the past decade, prominent roles for PI3K–AKT–MTOR signaling in the biology of breast cancer have been established, in particular downstream of HER2 (ERBB2: gene product; HER2 protein product) signaling and in tamoxifen resistance mechanisms (9–11).

We have previously examined the role of hyperactivation of ERK1/2 MAPK (hMAPK) occurring downstream of EGFR or HER2 in breast cancer. We established that hMAPK induces loss of ER expression, and that inhibition of hMAPK in ER–breast cancer cell lines and primary cultures from ER–breast tumors restores ER expression, suggesting that hMAPK plays a direct role in establishing the ER–phenotype in breast cancer (12–14). mRNA expression profiling of these hMAPK cell lines defined an “hMAPK gene expression profile” that strongly correlates with ER–status in clinical samples, establishing a link between hMAPK and the transcriptomes of ER–breast tumors (15).
MicroRNAs (miRNA) play critical regulatory roles in a wide range of biologic and pathologic processes, and miRNA expression profiling reveals substantial differences between cancer and normal tissue, and among different cancer types (16–18). In the context of breast cancer, miRNA expression has been correlated with numerous biopathologic features such as tumor grade, ER, PR, and HER2 status (19–21), TP53 mutation status, or proliferation status (22), and has also been used to classify the breast cancer subtypes identified by gene expression profiling (23). miRNA expression studies have demonstrated the utility for miRNA profiles as disease classifiers and prognostic tools in breast cancer (24, 25). Functional studies indicate that miRNAs regulate important genes and processes germane to breast cancer pathology, and are effectors of disease-associated signaling pathways (26–30).

In the present study, we identify miRNAs associated with hyperactivation of ERK1/2–MAPK signaling and investigate their involvement in facilitating gene expression, adverse clinical features, poor clinical outcome, and poor response to hormone therapy. Analysis of gene and protein expression, molecular alterations, and pathway enrichment of gene targets of these hMAPK–miRNA s identifies miRNAs that may potentially act as important regulators of MAPK signaling and downstream pathways, and which may serve as novel biomarkers or therapeutic targets in the treatment of breast cancer. Multivariate analysis indicates that discrete signatures made up of subsets of these hMAPK–miRNAs are independently associated with increased risk of disease-specific death and disease recurrence, particularly among patients with ER— disease and in ER— cancers treated with hormone therapy.

Materials and Methods

Clinical datasets

The following publicly available breast cancer datasets were used in this study: GSE22220 (Buffa dataset, accessed April 2012; ref. 31), GSE19536 (Enerly dataset, accessed April 2012; ref. 22), the Cancer Genome Atlas (TCGA; accessed April 2012; ref. 32), Lyng (GSE37405, accessed May 2012; ref. 33), and the METABRIC breast cancer dataset (accessed July 2013; refs. 34, 35). Patient populations have been previously described in detail (22, 31–35). Enerly and Lyng cohorts have previously been used as validation or comparison datasets for identification of novel roles of miRNAs in breast cancer (35–40).

Gene and miRNA expression datasets

Gene and/or miRNA expression data for Buffa, Enerly, and Lyng datasets were retrieved from NCBI GEO. Gene expression, DNA copy number, reverse phase protein array (RPPA) protein expression, miRNA expression, and mutation status for breast cancers from the TCGA dataset were obtained using the CGDS-R package v 1.1.30. miRNA expression for the METABRIC dataset was retrieved from the European Genome-phenome Archive.

Signature generation and statistical analysis

miRNA expression and miRNA expression values were standardized by scaling expression of an individual mRNA/miRNA within each individual dataset to have mean 0 and standard deviation 1. Buffa and TCGA datasets were used as training datasets for generating the hMAPK–miRNA signature. Cancers from each dataset were classified as high-hMAPK–mRNA or low-hMAPK–mRNA by Pearson correlation to an ideal hMAPK–mRNA signature as performed by Creighton and colleagues (15). Briefly, the ideal signature genes upregulated in the signature were given an arbitrary value of “1,” and genes downregulated in the signature an arbitrary value of “−1”; Pearson correlation between the ideal signature and z-score standardized gene expression for breast cancers from the training and validation datasets were determined and cancers with positive Pearson correlation were classified as “high-hMAPK–mRNA,” whereas cancers with negative Pearson correlation were classified as “low-hMAPK–mRNA.”

miRNAs significantly differentially expressed (P ≤ 0.05) between breast cancers from training datasets classified as hMAPK–mRNA or not-hMAPK–mRNA were identified by the Student t test. P values for expression differences between groups were permutation-adjusted by randomly assigning each sample to a group (creating a random group assignment for each sample) and performing the Student t test for between-group differences for each gene; this was repeated 1,000 times, and the rank of the P value for each individual gene from the nonpermuted analysis was divided by 1,001, resulting in the permutation-adjusted P value (e.g., if a nonadjusted P = 0.022 ranked 34th lowest, the adjusted P value would be 34/1,001 = 0.0339; ref. 41). This procedure is effectively a Westfall–Young correction for multiple comparisons (42). Only miRNAs that were commonly differentially expressed between hMAPK–mRNA tumors and not-hMAPK–mRNA tumors from both the Buffa and TCGA training datasets comprised the total hMAPK–miRNA signature.

Correlation to ideal hMAPK–miRNA signature was performed for all available primary breast cancers from both training and validation datasets, as described for hMAPK–mRNA signature above. Differential mRNA expression between high-hMAPK–miRNA and low-hMAPK–miRNA cancers from TCGA training dataset, and likewise from METABRIC validation dataset, was analyzed by the Student t test.

miRNA target prediction was performed using the miRWalk database (40), a database that uses 10 different miRNA target prediction algorithms. A miRNA was considered to be significantly upregulated in a subset of breast cancers if it was identified by at least five of seven algorithms.
prediction programs to identify putative miRNA targets. Genes were considered putative targets of hMAPK–miRNAs if they were predicted to be a target by three or more of the following prediction programs: DIANA-mt, miRanda, miRDB, mirWalk, RNAhybrid, PICTAR4, PICTAR5, PITA, RNA22, or TargetScan.

Pearson correlation of hMAPK–miRNA target expression to hMAPK–miRNA expression was performed to identify regulatory relationships that may differ between cancers classified as high-hMAPK–miRNA and those classified as low-hMAPK–miRNA. Standard statistical tests were performed, as described in Supplementary Materials and Methods. All statistical analyses were performed using R statistical software v 2.15.2 (43). Microsoft Excel 2007 and R statistical software v 2.15.2 were used for figure generation. All t tests performed were two-sided.

Figure 1.
Breakdown of ER status of primary breast cancers with paired mRNA and miRNA expression data from the two training and two validation datasets classified according to the hMAPK-mRNA signature established by Creighton et al. A, association of hMAPK–mRNA status with ER status in the TCGA and Buffa training datasets. B, association of hMAPK–mRNA status with ER status in the Enerly and METABRIC validation datasets. C and D, the Kaplan–Meier analysis of clinical outcome in patients with tumors classified according to hMAPK–mRNA signature. C, recurrence-free survival in Buffa dataset. D, recurrence-free survival (left), disease-specific survival in Enerly dataset (middle), and disease-specific survival in METABRIC dataset (right). Bar graphs: white, not hMAPK-mRNA; gray, hMAPK-mRNA; the Kaplan–Meier curves: dashed, low-hMAPK–mRNA; solid, high-hMAPK–mRNA. P values from the log-rank tests are indicated.

hMAPK–miRNA Signature and Hormone Resistance
www.aacrjournals.org Clin Cancer Res; 21(2) January 15, 2015
OF3

Research.
A, expression of genes downregulated (ESR1, CDKN1B, TOB1, PDCD4, and RPS6KA5) and upregulated (ETV5, RPS6KA4, CREB1, SRF, ANGPTL4, ETV1, ELK4, AT4, and RPS6KA1) [underlined and bolded genes reported to be downregulated or upregulated in hMAPK mRNA signature (15)] by activated MAPK signaling was examined in primary tumors from the TCGA dataset that were classified as high-hMAPK–miRNA (top) or low-hMAPK–miRNA (bottom) by our hMAPK–miRNA signature. Proportions of tumors with a given alteration ≥1.5-fold are reported. Blue, gene expression downregulation; and green, gene expression upregulation.

Alterations in copy number, mutational status, gene expression, and protein expression of numerous genes in the EGFR–ERK signaling pathway found in cancers form the TCGA dataset classified by our hMAPK–miRNA signature. (Continued on the following page.)
To generate an hMAPK–miRNA recurrence signature, univariate analyses were first performed on the total hMAPK–miRNA signature to determine whether expression of individual miRNAs (median cutoff) was associated with increased rate of disease recurrence in the Buffa training dataset. Breast cancers from the Buffa dataset were then partitioned into two groups based on median expression of every hMAPK–miRNA on an individual basis. Using this grouping as a classifier, the Kaplan–Meier survival curves were generated for disease recurrence, and the log-rank test P-values were determined. 22 miRNAs for which this log-rank test P-value was ≤0.1 were included in the recurrence signature. Correlation to an ideal hMAPK-recurrence signature was determined as described above for mRNA signature analysis.

To generate an hMAPK–miRNA survival signature, primary Breast Cancers from the METABRIC dataset were first randomized into two populations. A leave-one-out analysis was then performed, sequentially removing single individual miRNAs from the hMAPK–miRNA signature, and determining whether removal of each miRNA improved the ability of the signature to predict poor survival events in ER+ population at 5 years by the Cox proportional hazards analysis, separately for each randomized population. The analysis was limited to the ER+ population because the overwhelming majority of the ER- cancers were already classified as "high hMAPK." miRNAs whose removal improved the model in both randomized populations from the METABRIC dataset were tabulated. This randomization and leave-one-out analysis was repeated in 1,000 random permutations of the dataset to ensure robustness and serve as a self-validation. miRNAs whose removal improved the model in more than 20% of randomizations were removed from the signature, establishing a 21-member refined hMAPK–miRNA signature. Correlation to an ideal hMAPK–miRNA survival signature was determined as described above for mRNA signature analysis.

miRNA expression array and qRT-PCR analysis

miRNA expression array and qRT-PCR analysis of cell lines were performed according to standard practices, and are described in Supplementary Materials and Methods.

Results

hMAPK–miRNA signature associates with adverse features of breast cancer and poor clinical outcome

We previously observed that the hMAPK–miRNA signature is present in a majority of ER+ breast cancers and a small subset of ER- breast cancers (13, 14). We confirmed that the hMAPK–miRNA signature was similarly represented in the primary breast tumor datasets with paired mRNA–miRNA expression data analyzed in this study [TCGA: “TCGA” (training cohort); GEO GSE22220: "Buffa dataset” (training cohort); GEO GSE19536: "Enerly dataset” (validation cohort); METABRIC miRNA data: “METABRIC” (validation cohort); datasets previously described (22, 31, 32, 34, 35); clinical characteristics in Supplementary Table S1]. Tumors in each of these datasets classified as "high-hMAPK–mRNA" according to our hMAPK–mRNA signature (13) exhibited a significant association with ER-negativity in training and validation datasets (Fig. 1A and B). hMAPK–mRNA classification associated with a trend toward increased incidence of disease recurrence in the training dataset (Fig. 1C), and increased incidence of disease recurrence and poorer disease-specific survival in the validation datasets (Fig. 1D).

Generation and characterization of a breast cancer hMAPK–miRNA signature

We identified 16 commonly underexpressed and 41 commonly overexpressed miRNAs in tumors classified as hMAPK–mRNA from both training cohorts (P ≤ 0.05; see Supplementary Fig. S1 for study overview), establishing a 57-member hMAPK–miRNA signature (Supplementary Table S2). Primary breast cancers from training and validation datasets were classified as "high-hMAPK–miRNA" or "low-hMAPK–miRNA" according to this miRNA signature. To confirm that this hMAPK–miRNA signature identifies tumors with activated ERK1/2–MAPK signaling, we analyzed the expression of several genes regulated by ERK1/2 in breast cancer in the TCGA training cohort. Genes known to be down-regulated by ERK1/2 (ESR1, CDKN1B, TOB1, and PDCD4) are underexpressed in a higher proportion of high-hMAPK–miRNA tumors and, similarly, genes upregulated by ERK1/2 (ETV5, RPS6KA4, CREB1, SRF, ANGPT4, ETV1, ELK4, ATF4, and RPS6KA1) are overexpressed in a higher proportion of high-hMAPK–miRNA cancers compared with low-hMAPK–miRNA cancers (Fig. 2A). We also investigated mutational status, alterations in DNA copy number, and expression of upstream regulators of ERK1/2 signaling (Fig. 2B). Alterations indicative of activation of ERK1/2–MAPK signaling pathways, such as enhanced expression of EGFR, ERBB2, SOS1, NRAS, RAF1, and MAPK1 and lower expression of ERBB3 and RKIP, are seen in a higher proportion of high-hMAPK–miRNA cancers compared with low-hMAPK–miRNA cancers. Gene set enrichment analysis of predicted targets of hMAPK–miRNAs identifies substantial enrichment for genes involved in the MAPK signaling pathway (Supplementary Fig. S2A and S2B), suggesting that a number of these hMAPK–miRNAs may not only be regulated by MAPK signaling, but also contribute to sustaining hyperactivation of ERK1/2–MAPK signaling. Analysis of protein expression data from the TCGA dataset available for breast cancers with miRNA expression data revealed significant differences in expression and phosphorylation of numerous proteins between high-hMAPK–miRNA and low-hMAPK–miRNA tumors, including proteins involved in or regulated by ERK1/2–MAPK signaling, as well as epithelial–mesenchymal transition (EMT) proteins, protein
markers of basal-like breast cancer, and proteins involved in mediating tamoxifen resistance (Supplementary Table S3).

To study the impact that altering MAPK signaling has on miRNA expression in established models of hyperactive MAPK signaling in breast cancer, we altered MAPK signaling in our stable hMAPK cell lines (15) and investigated effects on select miRNAs from the hMAPK–miRNA signature that had been previously implicated in breast tumor biology, or that have predicted targets in genes validated by Creighton and colleagues (15) to be regulated by MAPK signaling. In particular, we treated ca-RAF/MCF-7 (which exhibits constitutive activation of MAPK signaling) with a MEK inhibitor to abrogate MAPK signaling, and in a complementary approach, we stimulated MAPK signaling in EGFR/MCF7 cell line (which overexpresses EGFR and is a ligand-inducible model for hyperactivation of MAPK signaling). Several hMAPK–miRNAs were altered in their expression by abrogating or inducing hMAPK (Supplementary Fig. S3A and S3B): overexpressed hMAPK–miRNAs significantly decreased following abrogation of MAPK signaling with U0126 treatment (hsa-miR221 and hsa-miR222) or increased following EGF treatment (hsa-miR378). Likewise, underexpressed hMAPK–miRNAs were increased following abrogation of MAPK signaling with U0126 treatment (hsa-let-7a, hsa-miR30a, hsa-miR30a* , hsa-miR125a-5p, and hsa-miR375) or decreased following EGF treatment (hsa-let-7a, hsa-let-7e, hsa-miR29c, and hsa-miR30c), indicating that these miRNAs are regulated by MAPK signaling. We observed that target genes of these validated hMAPK–miRNAs are differentially expressed between tumors classified as high-hMAPK–miRNA cancers with those classified as low-hMAPK–miRNA in the TCGA training dataset (Fig. 2C and D), and verified these relationships in the METABRIC validation dataset (Supplementary Fig. S4A–S4F). In addition, we observed differential protein expression of targets of hMAPK–miRNAs in breast cancers from the TCGA dataset classified as “high-hMAPK” versus “low-hMAPK” by our miRNA signature, and inverse relationships between expression of miRNAs and protein expression of selected gene targets in the TCGA dataset, suggesting that differential expression of these miRNAs may have a regulatory impact on the protein expression of their predicted gene targets (Supplementary Fig. S5A–S5H).

hMAPK–miRNA signature is associated with adverse clinical features of breast cancer

Cancers classified as high-hMAPK–miRNA in the TCGA and Buffa training data were enriched for ER− status (TCGA, P = 4.308e−16; Buffa, P = 8.508e−14; Supplementary Fig. S6A), high tumor grade (Buffa, P = 7.711e−11; Supplementary Fig. S6B), and basal-like or HER2+ PAM50 molecular subtypes (TCGA, P < 2.2e−16; Fig. 3A, top). High-hMAPK–miRNA cancers also displayed significantly earlier disease recurrence among all patients (Buffa, P = 0.00242; Fig. 3B); segregation of patients by cancer ER status revealed significant trends for earlier disease recurrence among patients with ER− disease (P = 0.0835) and ER+ disease (P = 0.167; Fig. 3B).

In the Enerly and METABRIC validation datasets, the hMAPK–miRNA signature was significantly associated with cancers that are ER− (P values: Enerly, 2.722e−08; METABRIC, 2.2e−16; Supplementary Fig. S7A), higher grade (P values: Enerly, 1.198e−09; METABRIC, 2.2e−16; Supplementary Fig. S7B), classified by a high proliferation metric (described by Enerly and colleagues (22); P = 2.218e−05; Supplementary Fig. S7C), HER2+ (P values: Enerly, 0.02115; METABRIC, <2.2e−16; Supplementary Fig. S7D), and basal and HER2+ PAM50 subtypes (P values: Enerly, 9.621e−16; METABRIC, 2.2e−16; Fig. 3A, middle and bottom). Survival analyses indicate that, among all patients, cancers classified as high-hMAPK–miRNA demonstrated significantly increased disease recurrence (P value: Enerly, 0.0119; Fig. 3C, top) and decreased disease-specific survival (P values: Enerly, 0.00194; METABRIC, 6.2e−06; Fig. 3C, bottom) at 5 years. Stratifying these analyses by ER status indicates that classification as high-hMAPK–miRNA identifies a population of ER− cancers with significantly increased disease recurrence (Enerly, P = 0.0059; Fig. 3C, bottom) and significantly poorer disease-specific survival (Enerly, P = 0.0191; METABRIC, P = 0.0818; Fig. 3C, bottom).

hMAPK–miRNAs associated with recurrence identify an hMAPK–miRNA recurrence signature

By limiting the signature to miRNAs whose individual expression was significantly associated with reduced recurrence-free survival on univariate analysis among patients from the Buffa training dataset, we identified a 22-member hMAPK–miRNA recurrence signature (Supplementary Table S4). Cancers from the Buffa dataset classified as high-hMAPK–miRNA by this recurrence signature demonstrated significantly reduced recurrence-free survival at 5 years in all patients (P = 5.52e−06; Supplementary Fig. S8) and in ER− and ER+–specific cohorts (ER− P = 0.00117; ER+ P = 0.0177; Supplementary Fig. S8). In the Enerly validation dataset, classification by this recurrence signature significantly associated with reduced recurrence-free survival among all patients and patients with ER− disease (P = 0.015 and 0.0523, respectively; Fig. 4A). In the Enerly and METABRIC datasets, high-hMAPK–miRNA status also significantly associated with shorter disease-specific survival in all patients (P values: Enerly, 0.00121; Fig. 4B; METABRIC, 4.49e−10; Fig. 4C) and patients with ER+ disease (P values: Enerly, 0.0672; Fig. 4B; METABRIC, 0.00144; Fig. 4C).

Cancers from the Lyng ER+ dataset (GSE37405; ref. 33) classified as high-hMAPK–miRNA by this recurrence signature demonstrated significantly earlier disease recurrence among breast cancers treated with adjuvant tamoxifen monotherapy (P = 0.00348; Fig. 4D). ER+ cancers from the METABRIC dataset receiving hormone therapy, either alone or in conjunction with chemotherapy or radiotherapy, which were classified as high-hMAPK by the hMAPK–miRNA recurrence signature demonstrated significantly poorer disease-specific survival (P = 0.00301; Fig. 4D).

Multivariate analysis of hMAPK–miRNA recurrence signature

Multivariate analysis of the hMAPK–miRNA recurrence signature in the METABRIC dataset using the Cox proportional hazards analyses revealed that among all patients, high-hMAPK–miRNA status is associated with a significant increase in risk of breast cancer-specific death within 5 years from diagnosis [multiplicative hazard factor, 1.48; 95% confidence interval (CI), 1.002–2.175; P = 0.04886; Supplementary Table S5A]. Because we observed that classification as high-hMAPK–miRNA by the hMAPK–miRNA recurrence signature was significantly associated with poor outcome in ER− breast cancers, and ER+ breast cancers displayed a large variation in correlation values with this recurrence signature ranging from very positive to very negative (in contrast to ER− breast cancers, whose correlations were almost entirely positive), we sought to determine whether an incremental
change in correlation with the hMAPK–miRNA recurrence signature (represented as a 1% increase in Pearson correlation coefficient) associated with change in risk of disease-specific death. In the ER+ METABRIC cohort, multivariate analysis indicated that incremental increase in Pearson correlation with the ideal hMAPK–miRNA recurrence signature carries a significant correspondingly incremental increase in risk of breast cancer-specific death within 5 years from diagnosis (multiplicative hazard factor, 1.01; 95% CI, 1.001–1.011; \( P = 0.0185 \); Supplementary Table S5B).

Similarly, incremental increased Pearson correlation with the ideal hMAPK–miRNA recurrence signature is also associated with a significant incremental increase in risk of breast cancer-specific death within 5 years from diagnosis among patients with ER+ cancer receiving hormone therapy, either alone or in conjunction with chemotherapy or radiotherapy (multiplicative hazard factor, 1.008; 95% CI, 1.002–1.014; \( P = 0.006285 \); Supplementary Table S5C). Multivariate analysis of the Lyng cohort indicated a signficant association with high-hMAPK–miRNA status according to the hMAPK–miRNA

---

**Figure 3.** Molecular subtyping and the Kaplan-Meier survival analysis of disease recurrence and survival in patients with tumors classified according to hMAPK-miRNA signature from training (Buffa and TCGA) and validation datasets (Enerly and METABRIC). A, association of hMAPK-miRNA status with PAM50 molecular subtypes in TCGA (top), Enerly (middle), and METABRIC (bottom) datasets. B, the Kaplan-Meier survival analysis of disease recurrence in patients with tumors classified according to the hMAPK-miRNA signature from training (Buffa dataset); disease recurrence among all patients, patients with ER+ disease, patients with ER+ disease (Buffa). C, disease recurrence and disease survival among all patients (Enerly) and disease survival among all patients from METABRIC dataset (top), disease recurrence and disease survival among patients with ER+ disease (Enerly) and disease survival among patients from METABRIC dataset with ER+ disease (bottom). Bar graphs: white, low-hMAPK-miRNA; gray, high-hMAPK-miRNA; P values given are for the \( \chi^2 \) test or the Fisher exact test, as indicated. The Kaplan-Meier curves: dashed, low-hMAPK-miRNA; solid, high-hMAPK-miRNA; the log-rank test \( P \) values are indicated.
recurrence signature and increased risk of disease recurrence within 5 years (multiplicative hazard factor, 2.109; 95% CI, 1.199–3.708; \( P = 0.009583 \); Supplementary Table S5D). In addition, multivariate analysis indicated that increased Pearson correlation with the ideal hMAPK–miRNA recurrence signature associates with a significant increase in risk of breast cancer-specific death within 5 years from diagnosis among ER\(^+\) patients from METABRIC dataset stratified by adjuvant therapy received (multiplicative hazard factor, 1.01; 95% CI, 1.0006–1.011; \( P = 0.029455 \); Supplementary Table S6A).

Leave-one-out analysis of hMAPK–miRNA signature identifies a subset associated with poor disease-specific survival outcomes in ER\(^+\) population

We hypothesized that miRNAs from the hMAPK–miRNA signature may contribute to de novo or acquired resistance to
hormone therapy, and may be associated with poor survival outcomes in patients with ER− disease. To test this hypothesis, we performed a leave-one-out analysis in the large METABRIC dataset. Briefly, miRNAs were individually excluded from analysis of the 57-member hMAPK–miRNA signature, and the Cox proportional hazards analysis for hazard of disease-specific death event within 5 years among patients with ER− disease from the METABRIC dataset was performed (see Materials and Methods for additional details). Individual miRNAs whose removal improved the ability of the hMAPK–miRNA signature to predict poor survival events in the ER− population at 5 years by the Cox proportional hazards analysis were tabulated (see Materials and Methods), and the 21 retained miRNAs make up a “hMAPK–miRNA survival signature” (see Supplementary Table S7). As expected, this survival signature has substantially improved the ability to predict poor disease survival in the Kaplan–Meier survival analysis and the Cox univariate hazard analysis of patients in the METABRIC dataset, in both total patient population ($P = 4.12e−17$; univariate HR, 3.7 [2.67–5.13]) and in the ER− patient population ($P = 1.44e−08$; univariate HR, 2.87 [1.96–4.21]; Fig. 5A) compared with either the total hMAPK–miRNA signature or the hMAPK–miRNA recurrence signature.

Breast cancers of the luminal A and luminal B subtypes that were classified as “high-hMAPK” by this hMAPK–miRNA survival signature exhibited significantly poorer 5-year survival compared with those classified as “low-hMAPK” (luminal A: $P = 0.0151$; univariate HR, 2.28 [1.15–4.52]; luminal B: $P = 0.000614$; univariate HR, 2.48 [1.45–4.25]; Fig. 5A). Breast cancers treated with any hormone therapy that were classified as “high-hMAPK” by this survival signature were at significantly higher risk for poorer disease survival than those classified as “low-hMAPK” ($P = 1.7e−06$; univariate HR, 2.74 [1.78–4.23]; Fig. 5B). Multivariate analysis of all breast cancers and ER− breast cancers in the METABRIC dataset indicated that high-hMAPK status, as determined by our hMAPK–miRNA survival signature, is an independent risk factor for reduced disease survival at 5 years among all patients ($P = 0.0031$; multiplicative hazard factor, 1.89 [1.24–2.89]) and especially among patients with ER− disease ($P = 0.0027$; multiplicative hazard factor, 1.95 [1.26–3.02; Tables 1 and 2]).

To validate this hMAPK–miRNA ER− survival signature, we assessed its association to disease survival in all patients and patients with ER− disease in the Enerly dataset ($P = 0.00689$ and 0.0538, respectively; Fig. 5C). We additionally assessed association of this survival signature with disease recurrence in all patients and patients with ER− disease in the Buffa ($P = 1e−07$ and 0.000245, respectively; Fig. 5D) and Enerly datasets ($P = 0.00393$ and 0.0167, respectively; Fig. 5D), and in ER− patients treated with tamoxifen monotherapy from the Lyng dataset ($P = 0.0371$; Fig. 5B). This survival signature was vastly superior to the total hMAPK–miRNA signature, and performed comparably with or better than the hMAPK-recurrence signature in the ability to predict poor outcome in patients classified as “high-hMAPK” versus “low-hMAPK,” even in the smaller publicly available datasets. This strongly suggests that the miRNAs retained in the survival signature may drive the poor survival outcome associated with the hMAPK–miRNA signature.

**Discussion**

Using paired mRNA and miRNA expression data from primary breast tumor datasets (TCGA, Buffa), we report a miRNA signature that associates significantly with the ERK1/2 hMAPK–miRNA expression signature. Classification of primary breast cancers using this hMAPK–miRNA signature indicates significant association with ER− status, high tumor grade, increased proliferation, basal and HER2 molecular subtypes, and poor clinical outcomes. This signature identifies a population of ER− primary breast cancers classified as high-hMAPK–miRNA that not only exhibit gene expression patterns similar to ER− primary breast cancers, but also exhibit significantly poorer clinical outcomes than ER− cancers classified as low-hMAPK–miRNA. These observations were validated in independent, publicly available cohorts (Enerly and METABRIC). We report a subsignature composed of 22 miRNAs, an hMAPK–miRNA recurrence signature, which retains significant associations with poor clinical outcome among patients with breast cancer and significantly associates with poor response to hormone therapy in two independent validation datasets by multivariate analysis (Lyng and METABRIC datasets).

Finally, we report a subsignature of this hMAPK–miRNA signature, consisting of 21 miRNAs, which is significantly associated with poorer disease-specific survival outcome in patients with ER− disease: an hMAPK–miRNA survival signature.

The hMAPK–miRNA signature (Supplementary Table S2) contains numerous miRNAs with previously reported roles in breast cancer etiology. Several hMAPK–miRNAs are significantly differentially expressed between basal/HER2− subtype and luminal A type cancers, between ER− and ER+, and between high-grade versus low-grade cancers (overexpressed: hsa-miR150, -142-3p, -142-5p, -148a, -155, and -135b; underexpressed: hsa-miR30a-3-, -30a-5p, let-7a, and -342; ref. 35). miR29c, underexpressed in the hMAPK–miRNA survival signature, correlates with expression of GATA3, a luminal identity promoting transcription factor, and the miR29 miRNA family facilitates GATA3 mediated maintenance of luminal identity (44). The let-7 family of miRNAs, underexpressed in the hMAPK–miRNA signature, negatively regulates RAS (45), and represses self-renewal and tumorigenicity of breast cancer stem cells (30). The EMT-promoting transcription factor SNAIL1 has been shown to be a direct target of miR30a (46). Overexpression of the hMAPK-downregulated miRNAs miR29c, miR30c, and miR342 has been reported to significantly associate with good prognosis (35), and we observe significantly lower expression of GATA3 as well as miR29c in ER− tumors classified as hMAPK−miRNA, and higher expression of basal markers and genes associated with poor clinical outcome, suggesting poor outcomes associated with these tumors may be due to loss of luminal identity and acquisition of ER+ characteristics. A number of miRNAs upregulated in the hMAPK–miRNA signature (miR142-5p/3p, miR146a, miR150, and miR155; ref. 35) are associated with markers of lymphocytic infiltrate and may reflect altered host immune or inflammatory responses associated with this hMAPK–miRNA signature. Thus, the downregulation and upregulation of these family members, respectively, in the hMAPK−miRNA signature may contribute to the breast cancer aggressiveness predicted by this profile.

A number of hMAPK–miRNAs regulate the ER (28–30, 47, 48). The hMAPK-upregulated miR221/222 family is overexpressed in ER− cancers (49), directly targets ER (49), is upregulated in HER2−amplified versus nonamplified breast cancers (29), and mediates tamoxifen resistance in MCF-7 cells by repression of CDKN1B, a cell-cycle regulator that is underexpressed in breast cancers classified as high-hMAPK by our miRNA signature (see Fig. 2; refs. 28, 29). Six underexpressed miRNAs in the hMAPK−miRNA signature
Figure 5.
Five-year outcome analysis of tumors from training (Buffa) and validation datasets (Enerly, METABRIC, and Lyng) classified according to the hMAPK–miRNA survival signature. A, the Kaplan-Meier survival analysis of disease recurrence in patients with tumors classified according to hMAPK–miRNA survival signature from METABRIC dataset; disease survival among all patients, patients with ER$^+$ disease (top), disease survival among patients with luminal A and luminal B type breast cancer (bottom). (Continued on the following page.)
Table 1. Multivariate analysis of hMAPK–miRNA survival signature in METABRIC dataset—Cox proportional hazards analysis of all patients

<table>
<thead>
<tr>
<th>Covariate</th>
<th>HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive LN (3+)</td>
<td>2.97 (1.89–4.69)α</td>
</tr>
<tr>
<td>Positive LN(2)</td>
<td>2.43 (1.37–4.31)β</td>
</tr>
<tr>
<td>Basal subtype</td>
<td>2.03 (0.95–4.31)γ</td>
</tr>
<tr>
<td>High-hMAPK</td>
<td>1.89 (1.24–2.69)β</td>
</tr>
<tr>
<td>ERBB2-positive</td>
<td>1.79 (1.18–2.73)β</td>
</tr>
<tr>
<td>Luminal B</td>
<td>1.62 (1.04–2.52)γ</td>
</tr>
<tr>
<td>PR-positive</td>
<td>0.67 (0.46–0.99)γ</td>
</tr>
</tbody>
</table>

NOTE: Cancers from METABRIC dataset were classified as high-hMAPK–miRNA or low-hMAPK–miRNA by hMAPK–miRNA survival signature, and this hMAPK–miRNA classification was included in the Cox proportional hazards analysis along with standard clinical covariates including tumor grade, stage, HER2 status, progesterone receptor status, lymph node status, and tumor PAM50 molecular subtype.

Significance codes: 0 “α”; 0.001 “β”; 0.01 “γ”; 0.05 “δ”.

Baseline: low hMAPK, grade 1, stage 0, 0 positive LN, ERBB2-negative, PR-negative, ER−, luminal A.

(hsa-miR125a, -let7c, -30d, -30a-5p, -30a-3p, and -149) are underexpressed in treatment-naïve ER+ primary cancers that expressed lower ESR1 mRNA levels, higher ERBB2 mRNA levels, and which had higher proportions of basal and HER2 subtype tumors than ER+ tumors with higher expression of these miRNAs (50). Elevated expression of miR30 family members (hsa-miR30a, -3p, and -30c) is significantly associated with tamoxifen sensitivity (51), whereas repression of miR375, underexpressed in the hMAPK–miRNA signature, is associated with resistance to tamoxifen in an ER+ breast cancer model (52). Analysis of the Lyng dataset containing only ER+ tumors treated with adjuvant tamoxifen monotherapy and the hormone therapy–treated cohort of the METABRIC dataset demonstrates that this hMAPK–miRNA signature identifies a subset of ER+ tumors that exhibit a higher incidence of recurrence and decreased disease-specific death following hormone therapy, suggesting an association of this hMAPK–miRNA signature with tamoxifen resistance and agreeing with previous observations linking activated MAPK signaling and resistance to endocrine therapy (53). The Lyng dataset consists of only ER+ cancers treated with adjuvant tamoxifen monotherapy, selected such that half of the patients exhibited disease recurrence over the course of 10-year follow-up (33). Lyng and colleagues (33) were not able to identify a miRNA signature which consistently predicted disease recurrence among ER+ breast tumors, and suggested that as numerous mechanisms contribute to hormone therapy resistance there may not be a single miRNA profile able to predict response to tamoxifen therapy. In our approach, we identified a miRNA signature indicative of a common biology, that of hMAPK signaling, and queried whether there was an association with clinical outcome, rather than taking patients with discrete clinical outcomes and searching for a miRNA signature that would predict those outcomes. That our hMAPK-recurrence and survival signatures show predictive value in the Lyng dataset reinforces the idea that activation of MAPK signaling represents a significant biologic event in the establishment of estrogen independence and tamoxifen insensitivity and that these hMAPK–miRNAs may mediate these effects. The hMAPK–miRNA recurrence signature and the hMAPK–miRNA survival signature identified here significantly associate with increased early risk of recurrence and poor disease survival in ER+ patient populations, supporting the notion that these miRNAs contribute to an ERK1/2–MAPK mechanism of tamoxifen resistance.

Multivariate analysis of the breast cancers from METABRIC and Lyng datasets indicates that classification as high-hMAPK–miRNA by the hMAPK–miRNA recurrence signature significantly contributes to increased risk of breast cancer-specific death and disease recurrence within 5 years after diagnosis. Although we report that the majority of ER+ breast cancers fall under the classification of high-hMAPK–miRNA, we did not identify ER status as a significant covariate in our multivariate analysis, indicating that increased risk associated with hMAPK–miRNA status is independent of ER status. We observed a significant increase in hazard associated with incremental increase in correlation to the idealized hMAPK–miRNA recurrence signature in ER+ breast cancers. Importantly, the hMAPK–miRNA survival signature demonstrated prognostic capability in both luminal A and luminal B subtypes of breast cancer from the METABRIC dataset. This indicates that hMAPK–miRNA survival signature is not just a surrogate for luminal B breast cancers from the ER+ group, which often exhibit higher proliferation rates, more activated growth factor signaling, and less reliance upon estrogen signaling than luminal A breast cancers. In addition, multivariate analysis of ER+ breast cancers from the METABRIC and Lyng datasets receiving hormone therapy, either alone or in combination with chemotherapy and/or radiotherapy, revealed that classification of ER+ breast cancers as high-hMAPK–miRNA by the recurrence or survival signature was significantly associated with increased risk of disease recurrence (Lyng) and disease-specific mortality (METABRIC). These results reinforce the observations we made from survival analysis of the hMAPK–miRNA signatures in the other training and validation datasets, and suggest that these hMAPK–miRNA signatures may have predictive value regarding response to tamoxifen therapy for patients with ER+ disease.

Table 2. Multivariate analysis of hMAPK–miRNA survival signature in METABRIC dataset—Cox proportional hazards analysis of METABRIC ER+ breast cancers only, including otherwise identical covariates

<table>
<thead>
<tr>
<th>Covariate</th>
<th>HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive LN (3+)</td>
<td>2.81 (1.62–4.86)α</td>
</tr>
<tr>
<td>Positive LN(2)</td>
<td>2.16 (1.05–4.45)β</td>
</tr>
<tr>
<td>High-hMAPK</td>
<td>1.95 (1.26–3.02)β</td>
</tr>
<tr>
<td>Luminal B</td>
<td>1.64 (1.04–2.56)β</td>
</tr>
<tr>
<td>PR-positive</td>
<td>0.59 (0.39–0.87)β</td>
</tr>
</tbody>
</table>

NOTE: Only covariates with significant hazard ratios are indicated. Significance codes: 0 “α”; 0.001 “β”; 0.01 “γ”; 0.05 “δ”.

Baseline: low hMAPK, grade 1, stage 0, 0 positive LN, ERBB2-negative, PR-negative, luminal A.
Focusing the analysis on the 12 miRNAs in common between the hMAPK-recurrence and hMAPK-survival, miRNA signatures reduced the ability to prognosticate increased disease recurrence and decreased disease survival in all patient cohorts compared with the recurrence and survival signatures, but was still more significant than the total 57-member hMAPK miRNA signature (data not shown). This suggests that, while there are hMAPK–miRNAs common to both the recurrence and the survival signature that may be useful biomarkers or potential drivers of general poor outcome (i.e., low expression of miR29c, miR30 family, high expression of 221/222, etc.), alterations in particular hMAPK–miRNAs associate specifically with either increased disease recurrence or poor disease-specific survival (i.e., in the survival signature miR22 and miR224 are overexpressed, whereas let-7a is underexpressed), especially among ER+ breast cancers.

The data presented here suggest that molecular and behavioral characteristics of tumors with activated growth factor signaling through the ERK1/2–MAPK axis are in part coordinated by aberrant miRNA expression associated with such activated MAPK signaling. The reported functions of members of the hMAPK–miRNA signature, together with the survival and recurrence data from the Enerly, METABRIC, and Lyng validation datasets (Figs. 4 and 5; Tables 1 and 2), suggest that the miRNAs identified in our hMAPK–miRNA signature contribute to endocrine resistance associated with activated ERK1/2–MAPK signaling, and raise the provocative possibility that the miRNAs contained within this signature may ultimately prove to have value as prognostic indicators of clinical outcome and as predictors of ER+ breast cancers with de novo endocrine resistance. These findings support further evaluation of the predictive potential of members of this hMAPK–miRNA signature for hormone therapy resistance and suggest that a subset of these miRNAs may also prove to be potential therapeutic targets. The identification of subsignatures of hMAPK–miRNAs that significantly associate with increased disease recurrence and reduced disease-specific survival outcomes, particularly in patients with ER+ breast cancer, indicates that stratification of patients according to expression of these miRNAs may ultimately provide important information related to disease prognosis and response to therapy; these data indicate that the clinical application of these signatures warrants further investigation prospectively in large patient cohorts where patients tumors would be arrayed for miRNA expression of these 21 hMAPK–miRNAs, and patients then stratified by high hMAPK–miRNA classification or low hMAPK–miRNA. Such classification by this signature could then inform decisions on tamoxifen treatment alone or more aggressive therapy, such as in combination with MEK or other signaling inhibitors, as informed by the targets of these miRNAs. Such prospective studies could be performed in tandem with current predictive gene expression assays (such as Oncotype DX or Mammaprint) to determine the added value of incorporating miRNA analysis with existing mRNA-based assays in predicting response of ER+ breast cancer to targeted therapy.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Conception and design: P.C. Miller, D. El-Ashy
Development of methodology: P.C. Miller, J. Clarke, D. El-Ashy
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): P.C. Miller, J. Brinkman
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): P.C. Miller, J. Clarke, T. Koru-Sengul, D. El-Ashy
Writing, review, and/or revision of the manuscript: P.C. Miller, J. Clarke, T. Koru-Sengul, D. El-Ashy
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): P.C. Miller, J. Brinkman
Study supervision: D. El-Ashy

Acknowledgments

The authors thank Drs. Joyce Slingerland (University of Miami, Miami, FL), Dan Hayes (University of Michigan, Ann Arbor, MI), James Rae (University of Michigan), Jennifer Richer (University of Colorado, Boulder, CO), and Marc Lippman (University of Miami) for critical review of the article; Dr. Marc Lippman and members of the El-Ashy and Lippman laboratory group for thoughtful discussion; and The University of Miami Sylvester Comprehensive Cancer Center Oncometrics Core Facility for technical assistance and generation of microarray data.

Grant Support

Funding for this research was provided by NIH grant NIH 1R01 CA113674 and by Bankhead Coley Foundation 09BW-04. 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.

Received August 7, 2014; revised October 16, 2014; accepted October 30, 2014; published OnlineFirst November 4, 2014.

References


A Novel MAPK–microRNA Signature Is Predictive of Hormone-Therapy Resistance and Poor Outcome in ER-Positive Breast Cancer

Philip C. Miller, Jennifer Clarke, Tulay Koru-Sengul, et al.

Clin Cancer Res  Published OnlineFirst November 4, 2014.

Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-14-2053

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2014/11/05/1078-0432.CCR-14-2053.DC1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.