Journal of Human Genetics **(JHG-20-640) revise manuscript**

Molecular pathogenesis of breast cancer: impact of *miR-99a-5p* and *miR-99a-3p* regulation on oncogenic genes

Yoshiaki Shinden[1,&], Tadahiro Hirashima[1,&], Nijiro Nohata[2], Hiroko Toda[3], Reona Okada[4], Shunichi Asai[4], Takako Tanaka[1], Yuto Hozaka[1], Takao Ohtsuka[1], Yuko Kijima[3], Naohiko Seki[4,#]

1, Department of Digestive Surgery, Breast and Thyroid Surgery, Graduate School of Medical and Dental Sciences, Kagoshima University, Kagoshima, Japan
2, MSD K.K., Tokyo, Japan
3, Department of Breast Surgery, School of Medicine, Fujita Health University, Japan
4, Department of Functional Genomics, Chiba University Graduate School of Medicine, Chuo-ku, Chiba 260-8670, Japan

& These authors contributed equally to this work.

Running title: Impact of *miR-99a* regulation on oncogenic genes in BrCa

[#]Correspondence to:
Naohiko Seki, Ph.D.
Associate Professor of Functional Genomics
Department of Functional Genomics
Chiba University Graduate School of Medicine
1-8-1 Inohana Chuo-ku, Chiba 260-8670, Japan
Tel: +81-43-226-2971
Fax: +81-43-227-3442

Abstract

Our recent research has revealed that passenger strands of certain microRNAs (miRNAs) function as tumor-suppressive miRNAs in cancer cells, e.g., *miR-101-5p*, *miR-143-5p*, *miR-144-5p*, *miR-145-3p*, and *miR-150-3p*. Thus, they are important in cancer pathogenesis. Analysis of the miRNA expression signature of breast cancer (BrCa) showed that the expression levels of 2 miRNAs derived from pre-*miR-99a* (*miR-99a-5p* and *miR-99a-3p*) were suppressed in cancerous tissues. The aim of this study was to identify oncogenic genes controlled by pre-*miR-99a* that are closely involved in the molecular pathogenesis of BrCa. A total of 113 genes were identified as targets of pre-*miR-99a* regulation (19 genes modulated by *miR-99a-5p*, and 95 genes regulated by *miR-99a-3p*) in BrCa cells. Notably, *FAM64A* was targeted by both of the miRNAs. Among these targets, high expression of 16 genes (*C5orf22*, *YOD1*, *SLBP*, *F11R*, *C12orf49*, *SRPK1*, *ZNF250*, *ZNF695*, *CDK1*, *DNMT3B*, *TRIM25*, *MCM4*, *CDKN3*, *PRPS*, *FAM64A*, and *DESI2*) significantly predicted reduced survival of BrCa patients based upon The Cancer Genome Atlas (TCGA) database. In this study, we focused on *FAM64A* and investigated the relationship between *FAM64A* expression and molecular pathogenesis of BrCa subtypes. The upregulation of *FAM64A* was confirmed in BrCa clinical specimens. Importantly, the expression of *FAM64A* significantly differed between patients with Luminal-A and Luminal-B subtypes. Our data strongly suggest that the aberrant expression of *FAM64A* is involved in the malignant transformation of BrCa. Our miRNA-based approaches (identification of tumor-suppressive miRNAs and their controlled targets) will provide novel information regarding the molecular pathogenesis of BrCa.

Keywords: Breast cancer, microRNA, *miR-99a-5p*, *miR-99a-3p*, luminal-A, luminal-B, *FAM64A*

Introduction

Breast cancer (BrCa) is the most common cancer affecting women. Approximately 2 million new cases are diagnosed each year, and over 600,000 people die of BrCa annually [1,2]. Approximately 5-10% of all BrCas are characterized by genetic factors, and among hereditary BrCa, at least 30% of cases result from germline mutations in the *BRCA1* and *BRCA2* genes [2-4]. The risk of developing BrCa increases to ~70% in women who inherit *BRCA1* or *BRCA2* mutations [2-4]. According to recent genomic analysis of BrCa, mutations in several cancer-related genes are involved in the development of BrCa, e.g., *TP53*, *PTEN*, *CHEK2*, *ATM*, and *PALB2* [2].

In general, the treatment strategies for BrCa are guided by immunohistochemical markers (hormone receptors, [ER, PR], growth receptor [HER2], a cell cycle marker [Ki67]) and pathological features [5,6]. BrCa is a heterogeneous cancer and can be classified into unique molecular subtypes (e.g., Luminal-A, Luminal-B, HER2-enriched, basal and normal breast-like) based on gene expression profile analysis [7-9]. These intrinsic molecular subtypes are associated with the biological characteristics of BrCa and are essential for therapeutic selection. However, within each subtype, there are patient populations that differ in treatment responses and require further classification.

Examination of the human genome has revealed that a large number (and types) of RNA molecules are transcribed. Among them, the noncoding RNAs are involved in various intracellular signal pathways, and their analysis is ongoing [10]. MicroRNA (miRNA) is a member of the noncoding RNAs (19–22 nucleotides in length, single-stranded RNA molecules). They function as fine-tuners of post-transcriptional RNA regulation [11,12]. A single miRNA controls many RNA transcripts. More than half of the RNA transcripts derived from the genome are controlled by miRNAs [11,12]. Numerous studies have revealed that aberrant expression of miRNAs and their controlled targets act as pivotal players of malignant transformation of human cancer cells [13,14].

We have been long working on miRNA-based approaches to identify novel molecular targets and pathways regulated in BrCa cells.

Recently, we created miRNA expression signatures of BrCa by RNA-sequencing [15,16]. Based on the signatures, we identified tumor-suppressive miRNAs and their directly regulated oncogenes. For example, *miR-204-5p* targets *AP1S3* whereas *miR-101-5p* targets *GINS1* [15,16]. Interestingly, these tumor-suppressive miRNAs controlled genes that are closely involved in BrCa molecular pathogenesis. Based on our signatures, identifying new prognostic and therapeutic targets for BrCa is accelerating.

Analysis of BrCa signatures revealed that two miRNAs derived from pre-*miR-99a* (*miR-99a-5p*: the guide strand, and *miR-99a-3p*: the passenger strand) were downregulated in cancerous tissues, suggesting that these miRNAs had tumor-suppressive roles in BrCa cells. In general, in miRNA biogenesis, the guide strands of miRNAs control genes. In contrast, passenger strands are degraded in the cytoplasm and lack any known function [11,12]. However, both strands of miRNAs have recently been recognized as functional, and recent analyses include passenger strands of miRNAs [17-21]. The aim of this study was to identify pre-*miR-99a*-controlled genes that are involved in BrCa molecular pathogenesis.

In this study, a total of 113 genes were identified as targets of pre-*miR-99a* regulation in BrCa cells. Among these targets, high expression of 16 genes (*C5orf22*, *YOD1*, *SLBP*, *F11R*, *C12orf49*, *SRPK1*, *ZNF250*, *ZNF695*, *CDK1*, *DNMT3B*, *TRIM25*, *MCM4*, *CDKN3*, *PRPS*, *FAM64A*, and *DESI2*) significantly predicted shorter survival of BrCa patients. Importantly, the expression of *FAM64A* was significantly different between patients with Luminal-A and Luminal-B types. Our miRNA-based study will provide novel insights into the molecular pathogenesis of BrCa.

Materials and Methods

Clinical BrCa specimens and BrCa cell lines

This study was approved by the Kagoshima University Bioethics Committee (approval number 160038:28-65 and 409 in Kagoshima, Japan). Prior written informed consent and approval was obtained from all patients from whom we collected clinical specimens. The study

methodologies conformed to the standards set by the Declaration of Helsinki.

In order to validate the expression levels of FAM64A, 15 clinical specimens (ER-positive and HER2-negative recurrent or metastatic breast cancer patients who had started CDK4/6 inhibitor: Palbociclib and/or Abemaciclib) were collected at Kagoshima University Hospital. The clinical features of the patients are shown in Supplemental Table 1. Progression-free survival was defined as the time from allocation to progressive disease according to the Response Evaluation Criteria in Solid Tumors (RECIST) version 1.1 or death due to any cause, whichever occurred first.

Three BrCa cell lines (MDA-MB-231, MDA-MB-157, and MCF-7) were used in this study.

Data mining of miRNA target genes and their expression in BrCa clinical specimens

The clinical significance of miRNAs and target genes was assessed by analysis of RNA-seq data of the Breast Cancer Cohort of TCGA (The Cancer Genome Atlas: https://cancergenome.nih.gov/) [22], and microarray data of METABRIC [23]. The gene expression data from TCGA and METABRIC were retrieved in March 2020 from cBioPortal (http://www.cbioportal.org/) [24].

The mRNA expression Z-scores and information on the clinical samples corresponding to BrCa patients were collected from cBioPortal. In order to categorize genes into molecular pathways based on gene set enrichment analysis (GSEA) (http://software.broadinstitute.org/gsea/index.jsp), the WebGestalt program was employed (http://www.webgestalt.org/).

Putative target genes possessing binding sequences to *miR-99a-5p* and *miR-99a-3p* were identified using the TargetScan Human database ver.7.2 (http://www.targetscan.org/vert_72/).

Comprehensive correlations between mRNA and miRNA gene expression in BrCa samples from TCGA were analyzed by LinkedOmics (http://www.linkedomics.org/).

Transfection of miRNAs, siRNAs, and plasmid vectors into BrCa cells

and functional assays

The procedures for transfecting miRNAs, siRNAs, and plasmid vectors were described in our previous studies [15,16]. Cell proliferation, migration, and invasion assays were performed with BrCa cell lines (MDA-MB-231, MDA-MB-157 and MCF-7), as previously outlined [15,16]. The reagents used are listed in Supplemental Table 2.

Incorporation of miRNAs (*miR-99a-5p* and *miR-99a-3p*) into RNA-induced silencing complex (RISC) by Agonaute-2 (Ago2) immunoprecipitation

To measurement of incorporated miRNAs into RISC in PDAC cells, we applied to Ago2 immunoprecipitation by using a microRNA Isolation Kit, Human Ago2 (Wako Pure Chemical Industries, Ltd., Osaka, Japan). The number of Ago2-conjugated miRNAs were assessed by qRT-PCR assay. The experimental procedure was described in previous studies [15,16].

Identification of putative target genes controlled by *miR-99a-5p* and *miR-99a-3p* in BrCa cells

We selected putative target genes having binding sites for *miR-99a-5p* and *miR-99a-3p* using TargetScanHuman ver.7.2 (http://www.targetscan.org/vert_72/; data were downloaded on 13 July 2018). Our microarray data (*miR-99a-5p* or *miR-99a-3p* transfected cells) were deposited in the GEO repository under accession number GSE113066. To examine upregulated genes in BrCa clinical specimens, we used expression data deposited in the GEO database (accession number; GSE118539).

Plasmid construction and dual-luciferase reporter assays

Vector construction and the dual-luciferase reporter assays were done as described in our previous studies [15,16]. Vector insertion sequences are shown in Supplemental Figure 2. The reagents used are listed in Supplemental Table 2.

Immunohistochemistry

Formalin-fixed, paraffin-embedded tissues were analyzed after immunohistochemical staining following the manufacturer's protocol. Tissues were treated with anti-PIMREG antibodies (1:100, ab251896,

Abcam, Cambridge, UK), and a single investigator scored the degree of immunostaining of the sections.

We evaluated the degree of staining in the nuclei of cancer cells. We used the same scoring method described in the previous study [25]. Proportions of tumor cells were scored as follows: 0 (no positive tumor cells), 1 (<10% positive tumor cells), 2 (10–50% positive tumor cells), and 3 (>50% positive tumor cells). The intensity of staining was graded according to the following criteria: 0 (no staining), 1 (weak staining), 2 (moderate staining), and 3 (strong staining). The staining index (SI) was calculated as the staining intensity score proportion of positive tumor cells. The expression of FAM64A (also termed as PIMREG) was evaluated by determining the SI, with scores as 0, 1, 2, 3, 4, 6, and 9. We set samples with SI $\geqq$ 6 as high expression group and samples with SI $\leqq$ 4 were determined as low expression group.

Statistical analyses

Statistical analyses were performed with GraphPad Prism 7 software (GraphPad Software, La Jolla, CA, USA) and JMP Pro 14 software (SAS Institute Inc., Cary, NC, USA). The Mann–Whitney U test was used to determine the significance of differences between 2 groups, and one-way analysis of variance and Tukey's test for post-hoc analysis were used for multiple group comparisons.

To evaluate the correlation between 2 variables, we applied Spearman's rank test. Overall survival (OS) and disease-free survival (DFS) were assessed using the Kaplan-Meier method and log-rank or Wilcoxon test. To identify independent factors predicting OS and DFS, we utilized multivariate Cox proportional hazards models.

Results

Downregulation of *miR-99a-5p* and *miR-99a-3p* in BrCa clinical specimens

The expression levels of *miR-99a-5p* and *miR-99a-3p* were evaluated in BRCA cohort data in TCGA. The cohort data showed that *miR-99a-5p* and *miR-99a-3p* were significantly downregulated in BrCa tissues compared with normal breast tissues (Figure 1A).

The expression levels of *miR-99a-5p* and *miR-99a-3p* were evaluated according to BrCa subtypes, e.g., ER(+)/HER2(-), ER(-)/HER2(+), ER(-)/HER2(-) and others. Downregulation of *miR-99a-5p* and *miR-99a-3p* was confirmed in all subtypes compared to normal breast tissues (Figure 1B).

Cell proliferation assays and expression of *miR-99a-5p* or *miR-99a-3p* in BrCa cell lines

To investigate the tumor-suppressive functions of *miR-99a-5p* and *miR-99a-3p* in BrCa cells, we assessed changes in cell proliferation after ectopic expression of these miRNAs in MDA-MB-231, MDA-MB-157, and MCF-7 cells.

Cell proliferation was significantly inhibited by *miR-99a-5p* or *miR-99a-3p* transfection in BrCa cells (Figure 1C).

Incorporation of *miR-99a-5p* and *miR-99a-3p* into the RISC in BrCa cells

Ago2 is an essential components of the RISC that binds to miRNAs. Whether the transfected miRNAs were incorporated into RISC in PDAC cells was analyzed by immunoprecipitation using an Ago2 antibody.

In *miR-99a-5p* transfected cells (MDA-MB-231, MDA-MB-157 and MCF-7), it was confirmed that a large amount of *miR-99a-5p* was incorporated into RISC (Supplemental Figure 1). Similarly, it was confirmed that transfected *miR-99a-3p* into BrCa cells were incorporated into RISC (Supplemental Figure 1). From these facts, it was shown that the transfected miRNAs (*miR-99a-5p* and *miR-99a-3*p) were functioning in the BrCa cells.

Identification of *miR-99a-5p* and *miR-99a-3p* target genes in BrCa cells

To identify putative targets of *miR-99a-5p* and *miR-99a-3p* regulation in BrCa cells, we assessed 3 datasets: (i) the TargetScan database to identify putative targets of *miR-99a-5p* and *miR-99a-3p in silico*; (ii) gene expression data for genes that were downregulated in *miR-99a-5p*- or *miR-99a-3p*-transfected BrCa cells; and (iii) gene expression data for genes that were upregulated in BrCa clinical specimens.

A total of 95 genes that were putative targets of *miR-99a-3p* regulation in BrCa cells were identified (Table 1A). Likewise, a total of 19 genes were identified as *miR-99a-5p*-regulated in BrCa cells (Table 1B). Notably, *FAM64A* was identified as a target of both *miR-99a-5p* and *miR-99a-3p*.

Furthermore, in order to predict the functions of *miR-99a-5p* and *miR-99a-3p* target genes, Gene Ontology (GO) classification was performed using the GeneCodis tool (https://genecodis.genyo.es). According to the GO classification (Biological Process), the genes associated with the "cell cycle" (GO:0007049) were most significantly contained. The analysis data is shown in the Supplemental Table 3.

Clinical significance of *miR-99a-5p* and *miR-99a-3p* target genes in BrCa pathogenesis

To investigate the clinical significance of the target genes in BrCa pathogenesis, we evaluated the associations between their expression levels and patient survival using TCGA and GEO datasets. Among the 113 genes, expression levels of 16 genes (*C5orf22*, *YOD1*, *SLBP*, *F11R*, *C12orf49*, *SRPK1*, *ZNF250*, *ZNF695*, *CDK1*, *DNMT3B*, *TRIM25*, *MCM4*, *CDKN3*, *PRPS1*, *FAM64A*, and *DESI2*) significantly predicted poorer survival in BrCa patients (OS: 10-year survival rates, $p < 0.05$; Figure 2).

We validated the expression levels of the 16 genes. All genes were upregulated in cancer tissues compared with normal tissues (Figure 3).

Direct regulation of *FAM64A* by *miR-99a-5p* and *miR-99a-3p* in BrCa cells

We focused on *FAM64A* because its expression was found to be controlled by both strands of pre-*miR-99a* (*miR-99a-5p* and *miR-99a-3p)*. The expression levels *FAM64A* were significantly reduced after transfection with *miR-99a-5p* and *miR-99a-3p* in BrCa cells (Supplemental Figure 2A).

There is one miRNA binding site for *miR-99a-5p* and one for *miR-99a-3p* in the 3′-UTR of *FAM64A* (Supplemental Figure 2B). The luciferase activities were significantly decreased by co-transfection with *miR-99a-5p* or *miR-99a-3p*, and the vector containing the wild-type 3′-UTR of *FAM64A*, whereas transfection with

the deletion vector blocked the decrease in luminescence in MDA-MB-231 cells (Supplemental Figure 2C). These data demonstrated that *miR-99a-5p* and *miR-99a-3p* directly bound to the 3′-UTR of *FAM64A* in BrCa cells. Vector-insertion sequences are shown in Supplemental Figure 2D.

Effects of *FAM64A* knockdown on cell proliferation in BrCa cells

To investigate the oncogenic function of *FAM64A* in BrCa cells (MDA-MB-231, MDA-MB-157, and MCF-7), we performed knockdown assays using siRNAs. The expression levels of *FAM64A* were successfully reduced by two different siRNAs (si*FAM64A*-1 and si*FAM64A*-2; Supplemental Figure 3A).

The proliferation of BrCa cells was attenuated by the transfection of each si*FAM64A* (Supplemental Figure 3B).

Moreover, cell migration and invasion activities were significantly downregulated by si*FAM64A* transfection in BrCa cells (Supplemental Figures 4A and 4B).

Clinical significance of *FAM64A* in BrCa pathogenesis

We focused on *FAM64A* because its expression was directly controlled by both *miR-99a-5p* and *miR-99a-3p* in BrCa cells. Furthermore, *FAM64A* expression was analyzed in each BrCa subtypes. We found that the expression gradually increased in the three subtypes in the following order: ER(+)/HER2(-), HER2(+) and ER(-)/HER2(-) (Figure 4A). Interestingly, expression levels of *FAM64A* were almost the same as in Luminal-A tissues and normal tissue (Figure 4A). However, expression levels of *FAM64A* were elevated in Luminal-B, HER2-enriched, and basal types compared to normal tissues (Figure 4A).

Expression of *FAM64A* and prognosis of the BrCa patients were investigated by limiting the analysis to the ER(+)/HER2(-) subtype. High expression of *FAM64A* predicted a significantly shorter overall survival of ER(+)/HER2(-) subtype patients ($p < 0.0001$; Figure 4B). Moreover, expression levels of *FAM64A* genes were an independent prognostic factor in multivariate analyses for the survival of patients with ER(+)/HER2(-) subtype ($p < 0.0001$; Figure 4C).

Effects of *FAM64A* on molecular pathways in BrCa

We compared differentially expressed genes in the *FAM64A*-high group and the *FAM64A*-low group in TCGA-BRCA. The GSEA showed that the top signaling pathways enriched in the *FAM64A*-high expression group were cell cycle-associated terms, such as G2/M checkpoint and E2F targets (Figure 4D). On the other hand, EMT and interferon response pathways were enriched in the *FAM64A*-low expression group (Figure 4D).

We also found that the proportion of genomic alterations (percentage of chromosome regions with copy number alterations relative to all regions evaluated) and the mutation count (the number of mutational events per case) were significantly increased in the high-*FAM64A* expression group (Figure 4E). Those findings suggested that *FAM64A* expression may be associated with genetic mutations and genomic instability in BrCa cells.

Expression of FAM64A protein and the therapeutic effects of CDK4/6 inhibitors in patients with ER(+)/HER2(-)

We used immunochemistry to assess the relationship between FAM64 protein expression and the therapeutic efficacy of CDK4/6 inhibitors (Palbociclib and/or Abemaciclib). The characteristics of the patients are shown in Supplemental Table 1. Examples of FAM64A immunostaining are shown in Figures 5A and 5B.

The response to CDK4/6 inhibitors was better in the FAM64A-high expression group than the low group. The FAM64A high expression group had better prognosis in PFS after using the CDK4/6 inhibitors than the low expression group (Figure 5C). The data suggest that FAM64A high-expression tumors had good responses to therapy, including the CDK4/6 inhibitors. OS after using CDK4/6 inhibitors was not different significantly in the two groups.

Discussion

The expression signature of miRNAs based on RNA-sequencing reveals the presence of novel miRNAs, the expressions of which are altered in cancer cells. We recently created miRNA expression signatures by RNA-sequencing in several types of cancers, e.g., BrCa, esophageal

squamous cell carcinoma, pancreatic ductal adenocarcinoma, and head and neck squamous cell carcinoma, [15,16,20,26,27]. Importantly, our studies demonstrated that passenger strands of pre-miRNAs functioned through controlling several genes that are closely involved in cancer pathogenesis [15,16,20,26,27].

Our present study shows that both strands of *miR-99a* act as tumor-suppressive miRNAs in BrCa cells. Moreover, they both regulate 113 genes that are involved in BrCa pathogenesis. There have been multiple reports of the tumor-suppressive function of *miR-99a-5p* (the guide strand) in BrCa cells. Previous studies showed that *mTOR*, *HOXA1*, *IGF-1R*, *CDC25*, and *FGF3* were directly regulated by *miR-99a-5p*. Decreased expression of *miR-99a-5p* was a cause of abnormal expression of these oncogenes in BrCa cells [28-33].

This is the first report of *miR-99a-3p* (the passenger strand) involved in the molecular pathogenesis of BrCa. A small number of studies have confirmed the tumor-suppressive roles of *miR-99a-3p* in human cancers. In renal cell carcinoma cells, *miR-99a-3p* was significantly downregulated in sunitinib-resistant cells, and its expression induced apoptosis through its targeting of *RRM2* [34]. In prostate cancer, expression of *miR-99a-3p* was reduced in castration-resistant prostate cancer tissues compared to hormone-naïve prostate cancer tissues [35]. Ectopic expression of *miR-99a-3p* inhibited features of cancer cell aggressiveness through regulating *NCAPG* [36]. Our recent study of head and neck squamous cell carcinoma (HNSCC) showed that *miR-99a-3p* acted as a tumor-suppressive miRNA, and a total of 5 genes (*STAMBP*, *TIMP4*, *TMEM14C*, *CANX*, and *SUV420H1*) were independent prognostic markers of HNSCC by multivariate analyses [36]. These data indicate that *miR-99a-3p* is closely involved in the molecular pathogenesis of human cancers.

In this study, we identified 113 genes that were targeted by pre-*miR-99a* in BrCa cells. Importantly, expression of 16 genes (*C5orf22*, *YOD1*, *SLBP*, *F11R*, *C12orf49*, *SRPK1*, *ZNF250*, *ZNF695*, *CDK1*, *DNMT3B*, *TRIM25*, *MCM4*, *CDKN3*, *PRPS*, *DESI2*, and *FAM64A*) predicted reduced survival of the patients. Functional analysis of these genes will enhance our understanding of the novel molecular mechanisms

underlying the progression of BrCa. Among these target genes, recent studies showed that several genes were closely associated with BrCa molecular pathogenesis. For example, overexpression of *F11R* (alias; Junctional adhesion molecule-A) was observed on BrCa with aggressive phenotypes [37]. Knockdown of *F11R* induced downregulation of HER2 expression and enhanced anti-proliferation effects in trastuzumab- and lapatinib-resistant BrCa cells [37]. *DNMT3B* is a member of DNA methyltransferases and its overexpression was observed in a wide range of human cancers, including BrCa. More recently, a study revealed that *DNMT3B* contributed to malignant transformation and distant metastasis in BrCa cells through various molecular pathways, e.g., STAT3, NFκB, PI3K/Akt, β-catenin, and Notch signaling pathways [38]. Recent study showed that *TRIM25* (Tripartite motif-containing protein 25) was a master regulator by controlling the pro-metastasis network of BrCa cells [39].

We focused on *FAM64A* in this study and demonstrated that its expression was directly controlled by both *miR-99a-5p* and *miR-99a-3p* in BrCa cells. Initially, *FAM64A* (also termed *PIMREG*, *CAKM*, *CATS*, and *RCS1*) was identified as a CALM/PICALM-interacting protein [40]. The chimeric gene, CALM/AF10: t(10;11)(p13;q14), plays a crucial role in the development of acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL) and malignant lymphoma [41,42]. Recent TCGA-based analysis demonstrated that *FAM64A* was upregulated in multiple cancer types compared to adjacent normal tissues [43]. In BrCa cells, overexpression of *FAM64A* enhanced cancer cell aggressiveness, e.g., proliferation, stemness, and epithelial-to-mesenchymal transition [44,45]. Moreover, overexpression of FAM64A contributed to the constitutive activation of NF-κB signaling in BrCa cells through disruption of the NF-κB/IκBα negative feedback loop [25]. These findings suggest that *FAM64A* is a potential target for diagnosis and therapy of BrCa.

In this study, high expression of *FAM64A* significantly predicted the shortened survival of patients with ER(+)/HER(-) disease. Moreover, expression levels of *FAM64A* differed among patients with Luminal-A and Luminal-B types. The expression status of *FAM64A* may help identify patients with the highly aggressive ER(+)/HER(-)

subtype. Luminal-A subtype BrCa is a unique feature that may have favorable cancer biology [46,47]. In the diagnosis of BrCa, the Luminal-A subtype makes up at least half of cases. It is important to determine the best way to identify the Luminal-A subtype and the best approach to treating this type of disease [46,47].

The frequency of gene mutations affecting cell cycle regulators in BrCa depends on the subtype [48]. Gene amplification of cyclin D1, the ER target, frequently occurred in ER-positive BrCa, 29% in the Luminal-A type, and 58% in the Luminal-B type [49]. Recently, CDK4/6 inhibitors have been used in the treatment of ER(+)/HER2(-) metastatic and recurrent BrCa patients [50-52]. We investigated whether *FAM64A* expression is a useful marker for predicting the therapeutic efficacy of CDK4/6 inhibitors. Our preliminary study showed that the *FAM64A*-high expression group had better PFS than the low expression group. Our results suggest that BrCa cells with high expression of *FAM64A* may have a better response to CDK4/6 inhibitors. This survey was too small to establish firm conclusions, and further clinical research is needed.

In conclusion, we demonstrated that both strands of pre-*99a* (*miR-99a-5p* and *miR-99a-3p*) act as tumor-suppressive miRNAs in BrCa cells. These miRNAs regulate a total of 16 genes (*C5orf22*, *YOD1*, *SLBP*, *F11R*, *C12orf49*, *SRPK1*, *ZNF250*, *ZNF695*, *CDK1*, *DNMT3B*, *TRIM25*, *MCM4*, *CDKN3*, *PRPS*, *DESI2,* and *FAM64A*), and high expression levels of these genes were significantly predictive of shorter survival times in BrCa patients. Overexpression of *FAM64A* was confirmed in BrCa clinical specimens. Moreover, the level of *FAM64A* expression may distinguish between Luminal-A and Luminal-B subtypes. Our miRNA-based strategy thus provides novel insights contributing to our overall understanding of the molecular pathogenesis of BrCa.

Acknowledgements: The present study was supported by KAKENHI grants 18K09338, 19K09049, 19K18060.

Conflicts of Interest: The authors declare no conflicts of interest.

NN is an employee of MSD K.K., a subsidiary of Merck & Co., Inc. and reports personal fees from MSD K.K. outside this study.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018; 68: 394–424.

2. Economopoulou P, Dimitriadis G, Psyrri A. Beyond BRCA: new hereditary breast cancer susceptibility genes. Cancer Treat Rev. 2015; 41: 1-8.

3. Lord CJ, Ashworth A. BRCAness revisited. Nat Rev Cancer. 2016; 16: 599-612.

4. Kuchenbaecker KB, Hopper JL, Barnes DR, Phillips KA, Mooij TM, Roos-Blom MJ, et al. Risks of breast, ovarian, and contralateral breast cancer for BRCA1 and BRCA2 mutation carriers. JAMA. 2017; 317: 2402–2416.

5. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000; 406: 747–752.

6. Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thurlimann B, Senn HJ and Panel Members. Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer. Ann Oncol. 2011; 22: 1736–1747.

7. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, et al. Breast cancer. Nat Rev Dis Primers. 2019; 5: 66.

8. Rakha EA, Fresia GP. New Advances in Molecular Breast Cancer Pathology. Semin Cancer Biol. 2020.

9. Jennifer JG, Sandra MS. Luminal A Breast Cancer and Molecular

Assays: A Review. 2018; 23: 556-565.

10. Anfossi S, Babayan A, Pantel K, Calin GA. Clinical utility of circulating non-coding RNAs—an update. Nat Rev Clin Oncol. 2018;15:541–63.

11. Ha M, Kim VN. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol. 2014;15:509–24

12. Gebert LFR, MacRae IJ. Regulation of microRNA function in animals. Nat Rev Mol Cell Biol. 2019;20:21–37.

13. Lin S, Gregory RI. MicroRNA biogenesis pathways in cancer. Nat Rev Cancer. 2015;15:321–33.

14. Rupaimoole R, Slack FJ. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. Nat Rev Drug Disco. 2017;16:203–22.

15. Toda H, Kurozumi S, Kijima Y, Idichi T, Shinden Y, Yamada Y, et al. Molecular pathogenesis of triple-negative breast cancer based on microRNA expression signatures: antitumor miR-204-5p targets AP1S3. J Hum Genet. 2018;63:1197-1210.

16. Toda H, Seki N, Kurozumi S, Shinden Y, Yamada Y, Nohata N, et al. RNA-sequence-based microRNA expression signature in breast cancer: tumor-suppressive miR-101-5p regulates molecular pathogenesis. Mol Oncol. 2020;14:426–46.

17. Osako Y, Seki N, Koshizuka K, Okato A, Idichi T, Arai T,et al. Regulation of SPOCK1 by dual strands of pre-miR-150 inhibit cancer cell migration and invasion in esophageal squamous cell carcinoma. J Hum Genet. 2017;62:935-944.

18. Misono S, Seki N, Mizuno K, Yamada Y, Uchida A, Arai T, et al. Dual strands of the miR-145 duplex (miR-145-5p and miR-145-3p)

regulate oncogenes in lung adenocarcinoma pathogenesis. J Hum Genet. 2018;63:1015-1028.

19. Uchida A, Seki N, Mizuno K, Misono S, Yamada Y, Kikkawa N, et al. Involvement of dual-strand of the miR-144 duplex and their targets in the pathogenesis of lung squamous cell carcinoma. Cancer Sci. 2019;110:420-432.

20. Wada M, Goto Y, Tanaka T, Okada R, Moriya S, Idichi T, et al. RNA sequencing-based microRNA expression signature in esophageal squamous cell carcinoma: oncogenic targets by antitumor miR-143-5p and miR-143-3p regulation. J Hum Genet. 2020

21. Mitra R, Adams CM, Jiang W, Greenawalt E, Eischen CM. Pan-cancer analysis reveals cooperativity of bothstrands of microRNA that regulate tumorigenesis and patient survival. Nat Commun. 2020;11:968.

22. Cancer Genome Atlas Network. Comprehensive Molecular Portraits of Human Breast Tumours. Nature. 2012;490:61–70.

23. Pereira B, Chin S, Rueda OM, Moen VH, Provenzano E, Bardwell HA, et al. The Somatic Mutation Profiles of 2,433 Breast Cancers Refines Their Genomic and Transcriptomic Landscapes. Nat Commun. 2016;7:11479.

24. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6:pl1.

25. Jiang L, Ren L, Zhang X, Chen H, Chen X, Lin C, et al. Overexpression of PIMREG promotes breast cancer aggressiveness via constitutive activation of NF-□B signaling. EBioMedicine. 2019 May;43:188-200.

26. Yonemori K, Seki N, Idichi T, Kurahara H, Osako Y, Koshizuka K,

et al. The microRNA expression signature of pancreatic ductal adenocarcinoma by RNA sequencing: anti-tumour functions of the *microRNA-216* cluster. Oncotarget. 2017;8:70097-70115.

27. Koshizuka K, Nohata N, Hanazawa T, Kikkawa N, Arai T, Okato A, et al. Deep sequencing-based microRNA expression signatures in head and neck squamous cell carcinoma: dual strands of pre-miR-150 as antitumor miRNAs. Oncotarget. 2017;8:30288-30304.

28. Hu Y, Zhu Q, Tang L. MiR-99a antitumor activity in human breast cancer cells through targeting of mTOR expression. PLoS One. 2014;9:e92099.

29. Yang Z, Han Y, Cheng K, Zhang G, Wang X. miR-99a directly targets the mTOR signalling pathway in breast cancer side population cells. Cell Prolif. 2014;47:587-95.

30. Wang X, Li Y, Qi W, Zhang N, Sun M, Huo Q, et al. MicroRNA-99a inhibits tumor aggressive phenotypes through regulating HOXA1 in breast cancer cells. Oncotarget. 2015;6:32737-47.

31. Xia M, Li H, Wang JJ, Zeng HJ, Wang SH. MiR-99a suppress proliferation, migration and invasion through regulating insulin-like growth factor 1 receptor in breast cancer. Eur Rev Med Pharmacol Sci. 2016;20:1755-63.

32. Qin H, Liu W. MicroRNA-99a-5p suppresses breast cancer progression and cell-cycle pathway through downregulating CDC25A. J Cell Physiol. 2019;234:3526-3537.

33. Long X, Shi Y, Ye P, Guo J, Zhou Q, Tang Y, et al. MicroRNA-99a Suppresses Breast Cancer Progression by Targeting FGFR3. Front Oncol. 2020;9:1473.

34. Osako Y, Yoshino H, Sakaguchi T, Sugita S, Yonemori M, Nakagawa M, et al. Potential tumor-suppressive role of microRNA-99a-3p in

sunitinib-resistant renal cell carcinoma cells through the regulation of RRM2. Int J Oncol. 2019;54:1759-1770.

35. Arai T, Okato A, Yamada Y, Sugawara S, Kurozumi A, Kojima S, et al. Regulation of NCAPG by miR-99a-3p (passenger strand) inhibits cancer cell aggressiveness and is involved in CRPC. Cancer Med. 2018;7:1988-2002.

36. Okada R, Koshizuka K, Yamada Y, Moriya S, Kikkawa N, Kinoshita T, et al. Regulation of Oncogenic Targets by *miR-99a-3p* (Passenger Strand of *miR-99a*- Duplex) in Head and Neck Squamous Cell Carcinoma. Cells. 2019;8:1535.

37. Leech AO, Vellanki SH, Rutherford EJ, Keogh A, Jahns H, Hudson L, et al. Cleavage of the extracellular domain of junctional adhesion molecule-A is associated with resistance to anti-HER2 therapies in breast cancer settings. Breast Cancer Res. 2018;20:140.

38. So JY, Skrypek N, Yang HH, Merchant AS, Nelson GW, Chen WD, et al. Induction of DNMT3B by PGE2 and IL6 at Distant Metastatic Sites Promotes Epigenetic Modification and Breast Cancer Colonization. Cancer Res. 2020;80:2612-2627.

39. Walsh LA, Alvarez MJ, Sabio EY, Reyngold M, Makarov V, Mukherjee S, et al. An Integrated Systems Biology Approach Identifies TRIM25 as a Key Determinant of Breast Cancer Metastasis. Cell Rep. 2017;20:1623-1640.

40. Archangelo LF, Gläsner J, Krause A, Bohlander SK. The novel CALM interactor CATS influences the subcellular localization of the leukemogenic fusion protein CALM/AF10. Oncogene. 2006;*25*:4099-4109.

41. Archangelo LF, Greif PA, Hölzel M, Harasim T, Kremmer E, Przemeck GK, et al. The CALM and CALM/AF10 interactor CATS is a marker for

proliferation. Mol Oncol. 2008;2:356-367.

42. Caudell D, Aplan PD. The role of CALM-AF10 gene fusion in acute leukemia. Leukemia. 2008;22:678-685.

43. Hu S, Yuan H, Li Z, Zhang J, Wu J, Chen Y, et al. Transcriptional response profiles of paired tumor-normal samples offer novel perspectives in pan-cancer analysis. Oncotarget. 2017;8:41334-41347.

44. Yao Z, Zheng X, Lu S, He Z, Miao Y, Huang H, et al. Knockdown of FAM64A suppresses proliferation and migration of breast cancer cells. Breast Cancer. 2019;26:835-845.

45. Zhang J, Qian L, Wu J, Lu D, Yuan H, Li W, et al. Up-regulation of FAM64A promotes epithelial-to-mesenchymal transition and enhances stemness features in breast cancer cells. Biochem Biophys Res Commun. 2019;513:472-478.

46. Ignatiadis M, Sotiriou C. Luminal breast cancer: from biology to treatment. Nat Rev Clin Oncol. 2013;10:494-506.

47. Gao JJ, Swain SM. Luminal A Breast Cancer and Molecular Assays: A Review. Oncologist. 2018;23:556-565.

48. Kyrochristos ID, Ziogas DE, Roukos DH. Dynamic genome and transcriptional network-based biomarkers and drugs: precision in breast cancer therapy. Med Res Rev. 2019;39:1205-1227.

49. Araki k, Miyoshi Y. Mechanism of resistance to endocrine therapy in breast cancer: the important role of PI3K/Akt/mTOR in estrogen receptor-positive, HER2-negative breast cancer. Breast Cancer. 2018;25:392-401.

50. O'Leary B, Finn RS, Turner NC. Treating cancer with selective CDK4/6 inhibitors. Nat Rev Clin Oncol. 2016;13:417-30.

51. Sobhani N, D'Angelo A, Pittacolo M, Roviello G, Miccoli A, Corona SP, et al. Updates on the CDK4/6 Inhibitory Strategy and Combinations in Breast Cancer. Cells. 2019;8:321.

52. Pandey K, An HJ, Kim SK, Lee SA, Kim S, Lim SM, et al. Molecular mechanisms of resistance to CDK4/6 inhibitors in breast cancer: A review. Int J Cancer. 2019;145:1179-1188.

Figure Legends

Figure 1: Tumor-suppressive roles of *miR-99a-5p* and *miR-99a-3p* in BrCa cells.

(A) Downregulation of *miR-99a-5p* and *miR-99a-3p* in BrCa clinical specimens by TCGA database analysis. (B) The expression levels of *miR-99a-5p* and *miR-99a-3p* were analyzed separately for each subtype of BrCa patient. (C) Cell proliferation was assessed using XTT assays in BrCa cell lines, MDA-MB-231, MBA-MB-157, and MCF-7. Data were collected 72 h after miRNA transfection (*$p < 0.0001$).

Figure 2: Clinical significance of *miR-99a-5p* and *miR-99a-3p* target genes according to TCGA database.

Among the putative target genes of *miR-99a-5p* and *miR-99a-3p* regulated in BrCa cells, high expression levels of 16 genes (*C5orf22*, *YOD1*, *SLBP*, *F11R*, *C12orf49*, *SRPK1*, *ZNF250*, *ZNF695*, *CDK1*, *DNMT3B*, *TRIM25*, *MCM4*, *CDKN3*, *PRPS*, *DESI2*, and *FAM64A*) were found to significantly predict poorer prognoses in patients with BrCa ($p < 0.05$). Kaplan-Meier curves of the 10-year overall survival rate for each gene are presented.

Figure 3: Expression levels of 16 target genes (that predicted 10-year survival) modulated by *miR-99a-5p* and *miR-99a-3p* in BrCa clinical specimens from TCGA analyses.

Expression levels of 16 target genes of *miR-99a* (Figure 2) were evaluated by TCGA database analyses. All genes were found to be upregulated in BrCa tissues (n = 1093) compared to normal tissues (n = 112).

Figure 4: Clinical significance of *FAM64A* expression in BrCa.

(A) Expression levels of *FAM64A* were investigated based on each subtype of BrCa patient. (B) Kaplan-Meier curves of the 10-year overall survival frequencies for patients with ER(+)/HER2(-) are presented. (C) Forest plot of multivariate analysis of *FAM64A* expression. The expression level of *FAM64A* is an independent prognostic factor for 10-year overall survival for patients with ER(+)/HER2(-). (D) Gene Set Enrichment Analysis (GSEA) by mRNA

expression levels of *FAM64A* in BrCa patients. Four representative GSEA plots are shown. G2/M checkpoint and E2F target pathways were significantly enriched in the *FAM64A*-high group. EMT and interferon γ response pathways were significantly enriched in the *FAM64A*-low group. (E) Fraction of genome alteration (% of copy number altered chromosome regions out of measured regions; left) and the mutation count (number of mutational events per cases; right) were significantly increased in the *FAM64A* high group.

Figure 5: FAM64A expression in ER-positive HER2-negative recurrent/ metastatic breast cancer.
Expression of FAM64A in BrCa clinical specimens (ER+ and HER2-) by immunohistochemistry. Examples of low expression cases (A) and high expression cases (B) were shown. FAM64A expression was evaluated with staining of nuclei of breast cancer cells. (C) Kaplan-Meier curves of the progression-free survival frequencies for ER(+)/HER2(-) patients with recurrence or metastasis after CDK4/6 inhibitor (Palbociclib and/or Abemaciclib) treatment is presented.

Supplemental Figure 1: Incorporation of *miR-99a-5p* and *miR-99a-3p* into the RISC in BrCa cells.
Isolated Ago2-bound miRNAs were analyzed by RT-qPCR to confirm whether *miR-99a-5p* and *miR-99a-3p* bound to Ago2. Data were normalized by the expression of *miR-21.* In BrCa cells (MDA-MB-231, MDA-MB-157, and MCF-7), *miR-99a-5p* transfectants demonstrated higher uptake levels of *miR-99a-5p* than mock transfectants, miR-control or *miR-99a-3p* transfectants. Similarly, following *miR-99a-3p* transfection, *miR-99a-3p* was detected by Ago2 immunoprecipitation. (*$p < 0.0001$).

Supplemental Figure 2: Direct regulation of *FAM64A* by *miR-99a-5p and miR-99a-3p* in BrCa cells.
(A) Expression levels of *FAM64A* were significantly reduced by *miR-99a-5p* and *miR-99a-3p* transfection into MDA-MB-231 cells. (B) TargetScan database analyses predicted that *miR-99a-5p* and *miR-99a-3p* each had one binding site in the 3′-UTR of *FAM64A.* (c)

Dual-luciferase reporter assays showed that luminescence activities were reduced by co-transfection of wild-type vector (with *miR-99a-5*p binding site) and *miR-99a-5p*, and wild-type vector (with *miR-99a-3*p binding site) and *miR-99a-3p* in MDA-MB-231 cells. Normalized data were calculated as *Renilla*/firefly luciferase activity ratios (**p* < 0.0001; N.S., Not significant). (D) The nucleotide sequences inserted into the vectors are shown.

Supplemental Figure 3: Effects of knockdown of *FAM64A* on cell proliferation of BrCa cells.
(A) Expression levels of *FAM64A* were successfully suppressed by si*FAM64A*-1 and si*FAM64A*-2 transfection into MDA-MB-231, MDA-MB-157, and MCF-7 cells. (B) Cell proliferation was assessed using XTT assays. Data were collected 72 h after miRNA transfection (**p* < 0.0001).

Supplemental Figure 4: Effects of knockdown of *FAM64A* on cell migration and invasion of BrCa cells.
(A) Cell migration was measured using a membrane culture system. Data were collected 48 h after seeding the cells into chambers (**p* < 0.0001). (B) Cell invasion was determined 48 h after seeding miRNA-transfected cells into Matrigel invasion chambers (**p* < 0.0001).

Table 1A Putative targets by *miR-99a-3*p regulation in BrCa cells

Entrez Gene	Gene Symbol	Gene name	Total sites	Upregulated genes in BrCa clinical specimens Fold change (log2 > 1.0)	Downregulated genes by miR-99a-5p transfection in MDA-MB-231 cells Fold change (log2 < -0.5)	10-year overall survival rate p-value
55322	*C5orf22*	chromosome 5 open reading frame 22	1	1.01738602	-0.8554292	0.0017
55432	*YOD1*	YOD1 deubiquitinase	1	1.16307434	-0.54761785	0.0131
7884	*SLBP*	stem-loop binding protein	1	1.49767138	-1.5822946	0.0139
50848	*F11R*	F11 receptor	1	1.10770003	-0.8633478	0.0164
79794	*C12orf49*	chromosome 12 open reading frame 49	1	1.2251136	-0.93862456	0.0187
6732	*SRPK1*	SRSF protein kinase 1	1	2.39980645	-0.78706837	0.0201
58500	*ZNF250*	zinc finger protein 250	2	1.21513278	-0.5442901	0.0216
57116	*ZNF695*	zinc finger protein 695	1	4.44321566	-0.56834084	0.0291
983	*CDK1*	cyclin-dependent kinase 1	1	5.9451053	-0.8056192	0.0306
1789	*DNMT3B*	DNA (cytosine-5-)-methyltransferase 3 beta	1	2.096505296	-0.83369714	0.0325
7706	*TRIM25*	tripartite motif containing 25	2	1.18288208	-0.52675897	0.0334
4173	*MCM4*	minichromosome maintenance complex component 4	1	3.28578355	-0.7807608	0.0371
1033	*CDKN3*	cyclin-dependent kinase inhibitor 3	1	5.94608437	-1.4748689	0.0393
5631	*PRPS1*	phosphoribosyl pyrophosphate synthetase 1	1	1.0804877	-0.91920686	0.0457
54478	*FAM64A*	family with sequence similarity 64, member A	1	4.34909188	-0.93843186	0.0494
55215	*FANCI*	Fanconi anemia, complementation group I	1	3.61349032	-0.7373419	0.0512
64151	*NCAPG*	non-SMC condensin I complex, subunit G	1	4.276076288	-0.86872	0.0523
10863	*ADAM28*	ADAM metallopeptidase domain 28	1	1.1427985	-0.7693379	0.0537
4751	*NEK2*	NIMA-related kinase 2	1	5.15460919	-1.5273995	0.0606
4288	*MKI67*	antigen identified by monoclonal antibody Ki-67	1	5.328904474	-0.5673255	0.0606
3832	*KIF11*	kinesin family member 11	1	3.80302127	-1.0077753	0.0689
7371	*UCK2*	uridine-cytidine kinase 2	2	3.206800186	-0.70736885	0.0702
126731	*CCSAP*	centriole, cilia and spindle-associated protein	1	1.77402136	-0.7316693	0.0747
114971	*PTPMT1*	protein tyrosine phosphatase, mitochondrial 1	1	1.258667825	-1.3865844	0.0767
11130	*ZWINT*	ZW10 interacting kinetochore protein	1	4.47834363	-0.59883296	0.0801
10825	*NEU3*	sialidase 3 (membrane sialidase)	1	1.449446693	-0.7443413	0.0913
5596	*MAPK4*	mitogen-activated protein kinase 4	1	2.15477954	-0.8519361	0.0934
113115	*MTFR2*	mitochondrial fission regulator 2	1	4.03328261	-0.5805406	0.0991
729230	*CCR2*	chemokine (C-C motif) receptor 2	1	2.16469888	-1.1817989	0.1168
26191	*PTPN22*	protein tyrosine phosphatase, non-receptor type 22 (lymphoid)	1	2.023809545	-0.8555168	0.1282
57619	*SHROOM3*	shroom family member 3	1	1.239897	-0.50767225	0.1307
22856	*CHSY1*	chondroitin sulfate synthase 1	1	1.02702899	-0.98781383	0.1317
119	*ADD2*	adducin 2 (beta)	1	3.04006748	-0.8456229	0.1362
54875	*CNTLN*	centlein, centrosomal protein	1	1.29824186	-0.81310636	0.1454
9282	*MED14*	mediator complex subunit 14	1	1.016861558	-0.9899208	0.1507
3766	*KCNJ10*	potassium inwardly-rectifying channel, subfamily J, member 10	1	1.93407872	-0.82971394	0.1642
5557	*PRIM1*	primase, DNA, polypeptide 1 (49kDa)	1	1.32324429	-0.6942426	0.1739
56952	*PRTFDC1*	phosphoribosyl transferase domain containing 1	1	2.405149203	-1.4599757	0.1953
140883	*ZNF280B*	zinc finger protein 280B	1	1.50408313	-0.8540934	0.2012
3782	*KCNN3*	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 3	3	1.236191655	-0.85277057	0.2064
4507	*MTAP*	methylthioadenosine phosphorylase	1	1.565450716	-0.8150412	0.21
11321	*GPN1*	GPN-loop GTPase 1	1	1.167405798	-0.80830187	0.2136
6241	*RRM2*	ribonucleotide reductase M2	1	3.743160542	-1.8221378	0.2151
341640	*FREM2*	FRAS1 related extracellular matrix protein 2	1	1.3222878	-1.2619936	0.2256
204	*AK2*	adenylate kinase 2	1	2.05410731	-0.5060893	0.232
4905	*NSF*	N-ethylmaleimide-sensitive factor	1	1.1664766	-1.3393241	0.2482
57684	*ZBTB26*	zinc finger and BTB domain containing 26	1	1.78874605	-1.1690493	0.2573
81796	*SLCO5A1*	solute carrier organic anion transporter family, member 5A1	1	2.76534195	-0.86495584	0.2641
348235	*SKA2*	spindle and kinetochore associated complex subunit 2	1	2.47732328	-0.69469357	0.265
8477	*GPR65*	G protein-coupled receptor 65	1	2.7438532	-1.5752689	0.2947
92292	*GLYATL1*	glycine-N-acyltransferase-like 1	1	2.24989688	-0.83099866	0.3061
2971	*GTF3A*	general transcription factor IIIA	1	1.22873624	-1.1663675	0.3126
139886	*SPIN4*	spindlin family, member 4	1	1.94368063	-0.576191	0.3188
729830	*FAM160A1*	family with sequence similarity 160, member A1	1	1.23539917	-0.80362284	0.3574
5027	*P2RX7*	purinergic receptor P2X, ligand-gated ion channel, 7	1	1.603523625	-0.84574753	0.3594
5150	*PDE7A*	phosphodiesterase 7A	2	1.61001512	-0.55403656	0.361
4436	*MSH2*	mutS homolog 2	1	3.218406643	-0.58050627	0.3899
10964	*IFI44L*	interferon-induced protein 44-like	1	2.4641337	-0.68931293	0.4103
51642	*MRPL48*	mitochondrial ribosomal protein L48	1	1.43577964	-1.0735437	0.436
55916	*NXT2*	nuclear transport factor 2-like export factor 2	1	1.239889236	-1.5145341	0.4499

9709	*HERPUD1*	homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1	1	1.034342165	-0.8028811	0.4501
5163	*PDK1*	pyruvate dehydrogenase kinase, isozyme 1	1	1.99820907	-0.9712486	0.4619
3676	*ITGA4*	integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)	1	2.24218054	-0.8488977	0.4741
219790	*RTKN2*	rhotekin 2	1	1.94469413	-0.61484104	0.5059
9507	*ADAMTS4*	ADAM metallopeptidase with thrombospondin type 1 motif, 4	2	4.55619766	-0.8408131	0.5151
5307	*PITX1*	paired-like homeodomain 1	1	5.90232915	-0.5850916	0.5334
55632	*G2E3*	G2/M-phase specific E3 ubiquitin protein ligase	1	1.08945445	-0.5671835	0.537
1E+08	*SIGLEC14*	sialic acid binding Ig-like lectin 14	1	2.883414066	-0.86273104	0.5422
9188	*DDX21*	DEAD (Asp-Glu-Ala-Asp) box helicase 21	1	1.1816007	-0.69152737	0.5554
285382	*C3orf70*	chromosome 3 open reading frame 70	1	2.0665948	-1.2021079	0.5632
84908	*FAM136A*	family with sequence similarity 136, member A	1	1.58857094	-0.6684657	0.5657
79152	*FA2H*	fatty acid 2-hydroxylase	1	1.72207075	-1.9765699	0.5682
4481	*MSR1*	macrophage scavenger receptor 1	1	3.37965255	-0.8699817	0.6068
23225	*NUP210*	nucleoporin 210kDa	1	2.679524564	-0.7378459	0.627
1729	*DIAPH1*	diaphanous-related formin 1	1	1.75873225	-0.73850775	0.6338
139285	*AMER1*	APC membrane recruitment protein 1	1	1.69424082	-0.8348991	0.6651
64770	*CCDC14*	coiled-coil domain containing 14	1	1.16539666	-0.65920436	0.6975
23114	*NFASC*	neurofascin	1	1.658357719	-0.85161525	0.7281
3037	*HAS2*	hyaluronan synthase 2	1	1.2195903	-1.0471387	0.7503
353355	*ZNF233*	zinc finger protein 233	1	1.34180749	-0.8719424	0.782
353500	*BMP8A*	bone morphogenetic protein 8a	1	2.37134127	-0.8491625	0.823
5965	*RECQL*	RecQ protein-like (DNA helicase Q1-like)	3	1.01599909	-1.2409371	0.8293
114088	*TRIM9*	tripartite motif containing 9	1	1.3785621	-1.3875808	0.8305
8819	*SAP30*	Sin3A-associated protein, 30kDa	1	2.400758735	-0.76948166	0.8442
387119	*CEP85L*	centrosomal protein 85kDa-like	2	1.84140808	-0.8404673	0.8488
9352	*TXNL1*	thioredoxin-like 1	1	1.352058307	-1.2844013	0.8494
127703	*C1orf216*	chromosome 1 open reading frame 216	1	1.28944233	-1.1623883	0.8545
9659	*PDE4DIP*	phosphodiesterase 4D interacting protein	1	1.41972497	-1.4432238	0.9033
55339	*WDR33*	WD repeat domain 33	1	1.50067893	-0.6801382	0.9243
8832	*CD84*	CD84 molecule	2	1.7966833	-0.85565436	0.9334
149420	*PDIK1L*	PDLIM1 interacting kinase 1 like	1	1.17156335	-0.5682502	0.9396
728	*C5AR1*	complement component 5a receptor 1	1	1.9558317	-0.8240501	0.9668
23279	*NUP160*	nucleoporin 160kDa	1	1.685647435	-0.6186448	0.9861
3609	*ILF3*	interleukin enhancer binding factor 3, 90kDa	1	1.27175378	-0.7088605	0.9939
91607	*SLFN11*	schlafen family member 11	1	1.29068583	-1.8289578	0.9952

Table 1B Putative targets by *miR-99a-5p* regulation in BrCa cells

Entrez Gene	Gene Symbol	Gene name	Total binding site	Upregulated genes in BrCa clinical specimens Fold change (log2 > 1.0)	Downregulated genes by miR-99a-5p transfection in MDA-MB-231 cells Fold change (log2 < -0.5)	10-year overrall survival rate p-value
51029	*DESI2*	desumoylating isopeptidase 2	1	5.18661384	-0.5243997	0.0392
54478	*FAM64A*	family with sequence similarity 64, member A	0	4.34909188	-1.1865349	0.0494
23066	*CAND2*	cullin-associated and neddylation-dissociated 2 (putative)	0	3.38351938	-0.9171996	0.0983
55754	*TMEM30A*	transmembrane protein 30A	1	2.7490854	-1.2071294	0.1066
8541	*PPFIA3*	protein tyrosine phosphatase, receptor type, f polypeptide (PT	0	2.71854478	-1.4380178	0.1199
2124	*EVI2B*	ecotropic viral integration site 2B	0	2.4210338	-1.6084512	0.1205
400746	*NCMAP*	noncompact myelin associated protein	0	2.10444327	-1.3652966	0.1482
9480	*ONECUT2*	one cut homeobox 2	0	2.04110646	-0.5466981	0.1596
54491	*FAM105A*	family with sequence similarity 105, member A	0	1.89790948	-0.8176197	0.1658
22996	*TTC39A*	tetratricopeptide repeat domain 39A	1	1.78416152	-0.7515176	0.206
2261	*FGFR3*	fibroblast growth factor receptor 3	1	1.70475088	-1.1740819	0.3477
1870	*E2F2*	E2F transcription factor 2	0	1.5753339	-1.4057112	0.4205
84620	*ST6GAL2*	ST6 beta-galactosamide alpha-2,6-sialyltranferase 2	0	1.30184504	-0.72825116	0.5752
121268	*RHEBL1*	Ras homolog enriched in brain like 1	0	1.22502847	-0.60373783	0.675
3437	*IFIT3*	interferon-induced protein with tetratricopeptide repeats 3	0	1.222916529	-1.1946698	0.7772
25780	*RASGRP3*	RAS guanyl releasing protein 3 (calcium and DAG-regulated)	1	1.2114101	-0.8922431	0.7785
3614	*IMPDH1*	IMP (inosine 5'-monophosphate) dehydrogenase 1	0	1.20418401	-0.8492775	0.8526
3433	*IFIT2*	interferon-induced protein with tetratricopeptide repeats 2	0	1.064862982	-1.1962117	0.8835
23446	*SLC44A1*	solute carrier family 44 (choline transporter), member 1	1	1.00411866	-1.1147652	0.9307

Figure 1

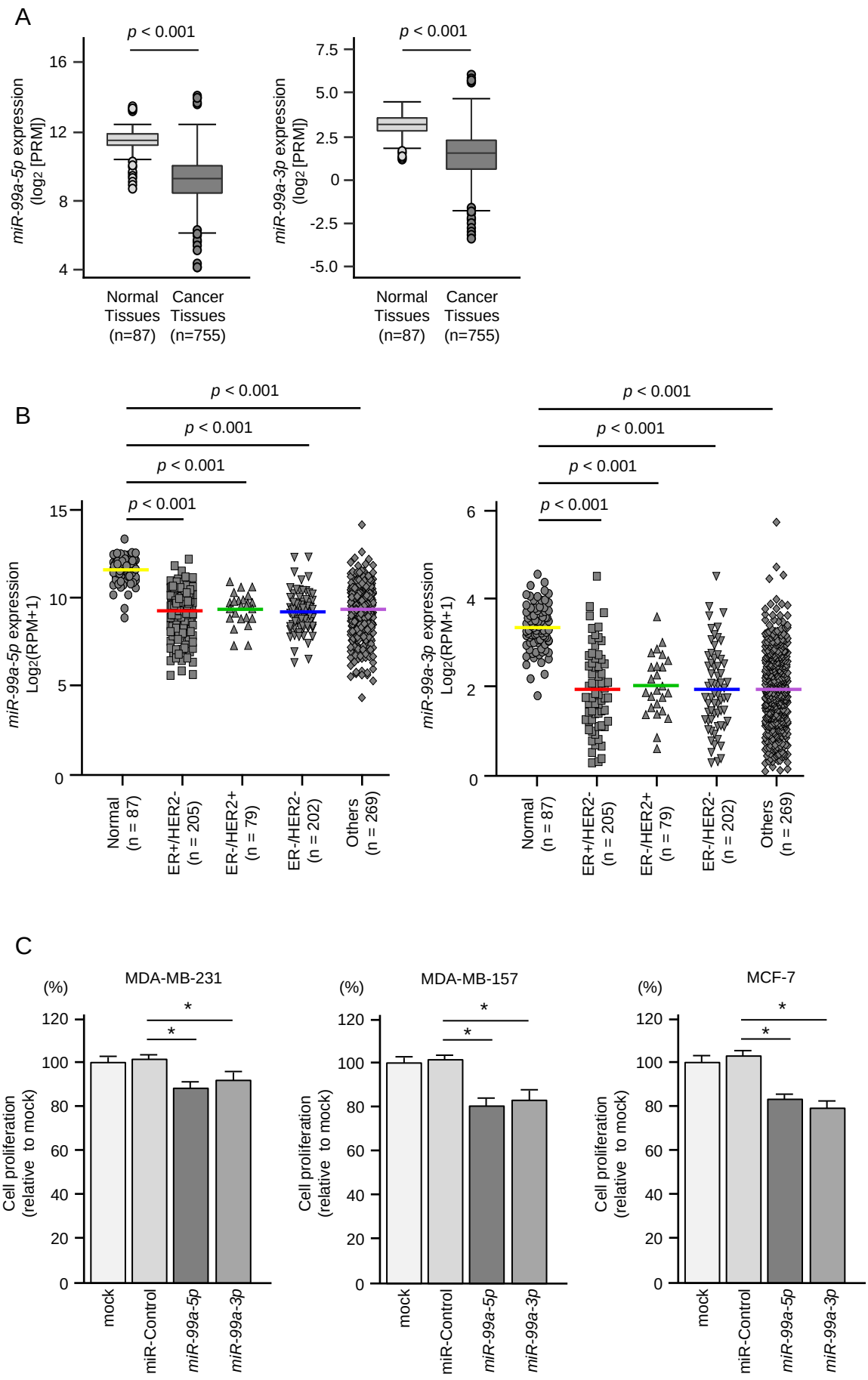

A
p < 0.001
miR-99a-5p expression (log2 [PRM])
16
12
8
4
Normal Tissues (n=87)
Cancer Tissues (n=755)
p < 0.001
miR-99a-3p expression (log2 [PRM])
7.5
5.0
2.5
0
-2.5
-5.0
Normal Tissues (n=87)
Cancer Tissues (n=755)
B
p < 0.001
p < 0.001
p < 0.001
p < 0.001
miR-99a-5p expression Log2(RPM+1)
15
10
5
0
Normal (n = 87)
ER+/HER2- (n = 205)
ER-/HER2+ (n = 79)
ER-/HER2- (n = 202)
Others (n = 269)
p < 0.001
p < 0.001
p < 0.001
p < 0.001
miR-99a-3p expression Log2(RPM+1)
6
4
2
0
Normal (n = 87)
ER+/HER2- (n = 205)
ER-/HER2+ (n = 79)
ER-/HER2- (n = 202)
Others (n = 269)
C
MDA-MB-231
(%)
Cell proliferation (relative to mock)
120
100
80
60
40
20
0
*
*
mock
miR-Control
miR-99a-5p
miR-99a-3p
MDA-MB-157
(%)
Cell proliferation (relative to mock)
120
100
80
60
40
20
0
*
*
mock
miR-Control
miR-99a-5p
miR-99a-3p
MCF-7
(%)
Cell proliferation (relative to mock)
120
100
80
60
40
20
0
*
*
mock
miR-Control
miR-99a-5p
miR-99a-3p

Figure 2

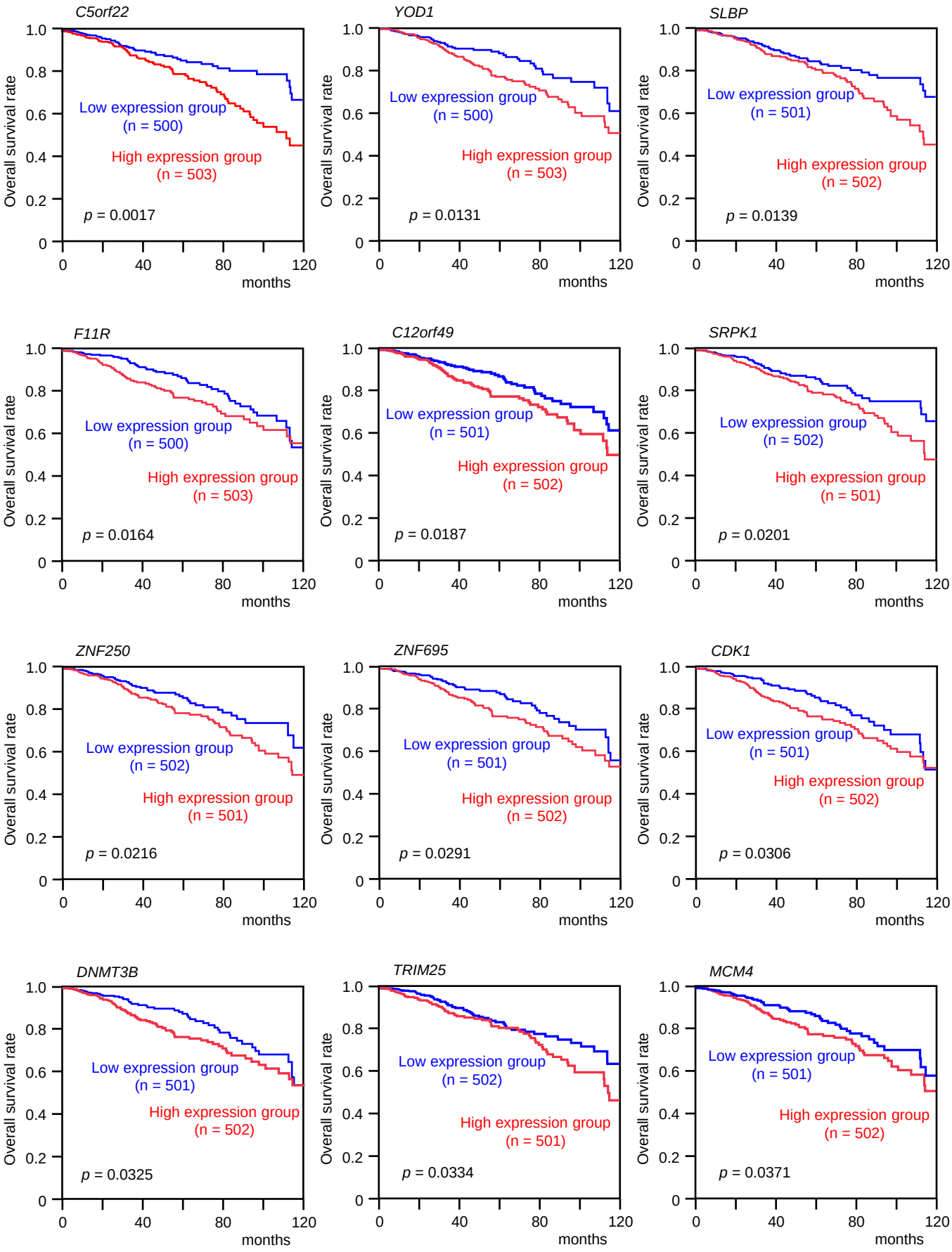

C5orf22
Low expression group (n = 500)
High expression group (n = 503)
p = 0.0017
YOD1
Low expression group (n = 500)
High expression group (n = 503)
p = 0.0131
SLBP
Low expression group (n = 501)
High expression group (n = 502)
p = 0.0139
F11R
Low expression group (n = 500)
High expression group (n = 503)
p = 0.0164
C12orf49
Low expression group (n = 501)
High expression group (n = 502)
p = 0.0187
SRPK1
Low expression group (n = 502)
High expression group (n = 501)
p = 0.0201
ZNF250
Low expression group (n = 502)
High expression group (n = 501)
p = 0.0216
ZNF695
Low expression group (n = 501)
High expression group (n = 502)
p = 0.0291
CDK1
Low expression group (n = 501)
High expression group (n = 502)
p = 0.0306
DNMT3B
Low expression group (n = 501)
High expression group (n = 502)
p = 0.0325
TRIM25
Low expression group (n = 502)
High expression group (n = 501)
p = 0.0334
MCM4
Low expression group (n = 501)
High expression group (n = 502)
p = 0.0371
Overall survival rate
months
0
0.2
0.4
0.6
0.8
1.0
0
40
80
120

Figure 2

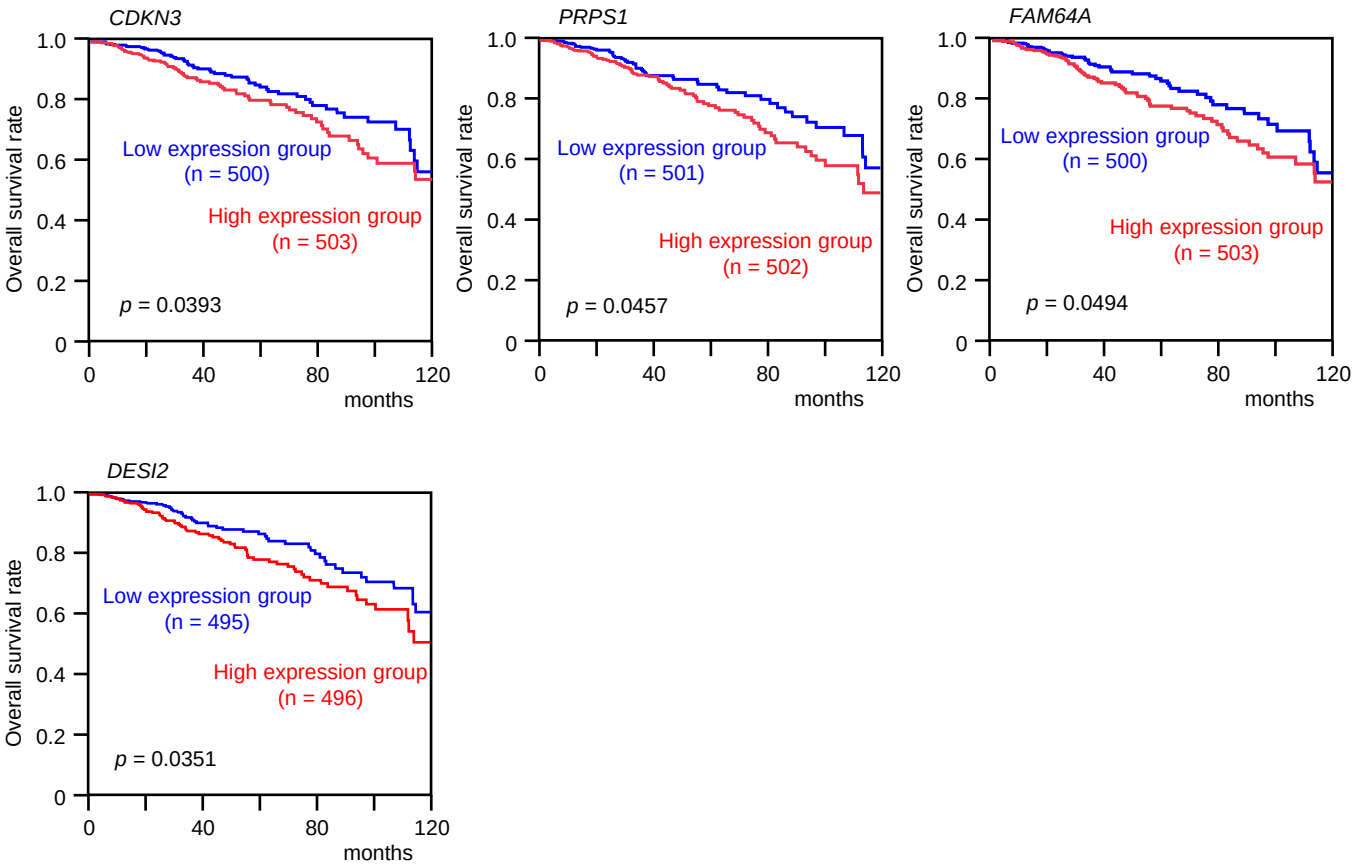

CDKN3
Overall survival rate
1.0
0.8
0.6
0.4
0.2
0
Low expression group
(n = 500)
High expression group
(n = 503)
p = 0.0393
0
40
80
120
months
PRPS1
Low expression group
(n = 501)
High expression group
(n = 502)
p = 0.0457
FAM64A
Low expression group
(n = 500)
High expression group
(n = 503)
p = 0.0494
DESI2
Low expression group
(n = 495)
High expression group
(n = 496)
p = 0.0351

Figure 3

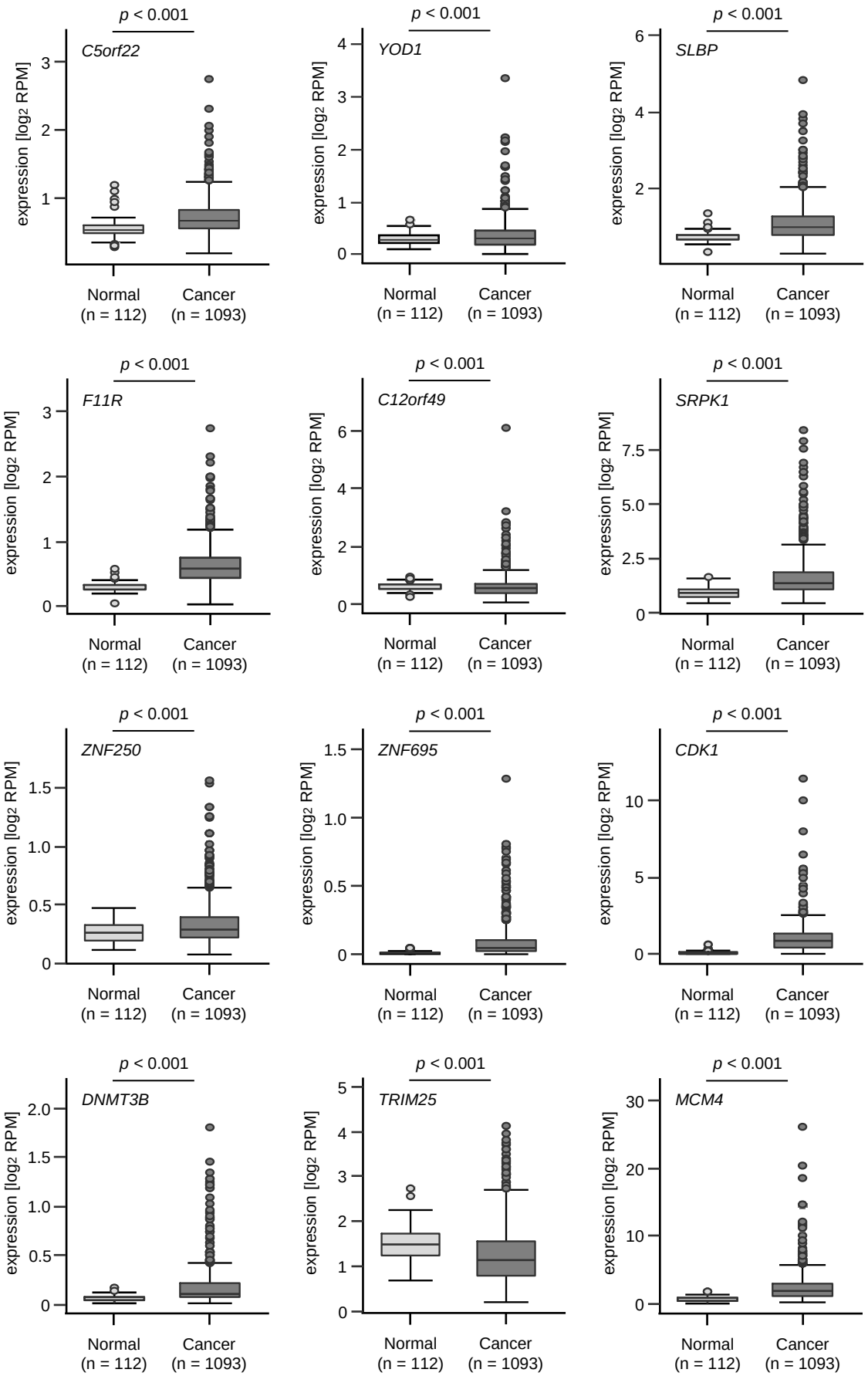

p < 0.001
C5orf22
YOD1
SLBP
F11R
C12orf49
SRPK1
ZNF250
ZNF695
CDK1
DNMT3B
TRIM25
MCM4
expression [log2 RPM]
Normal (n = 112)
Cancer (n = 1093)

Figure 3

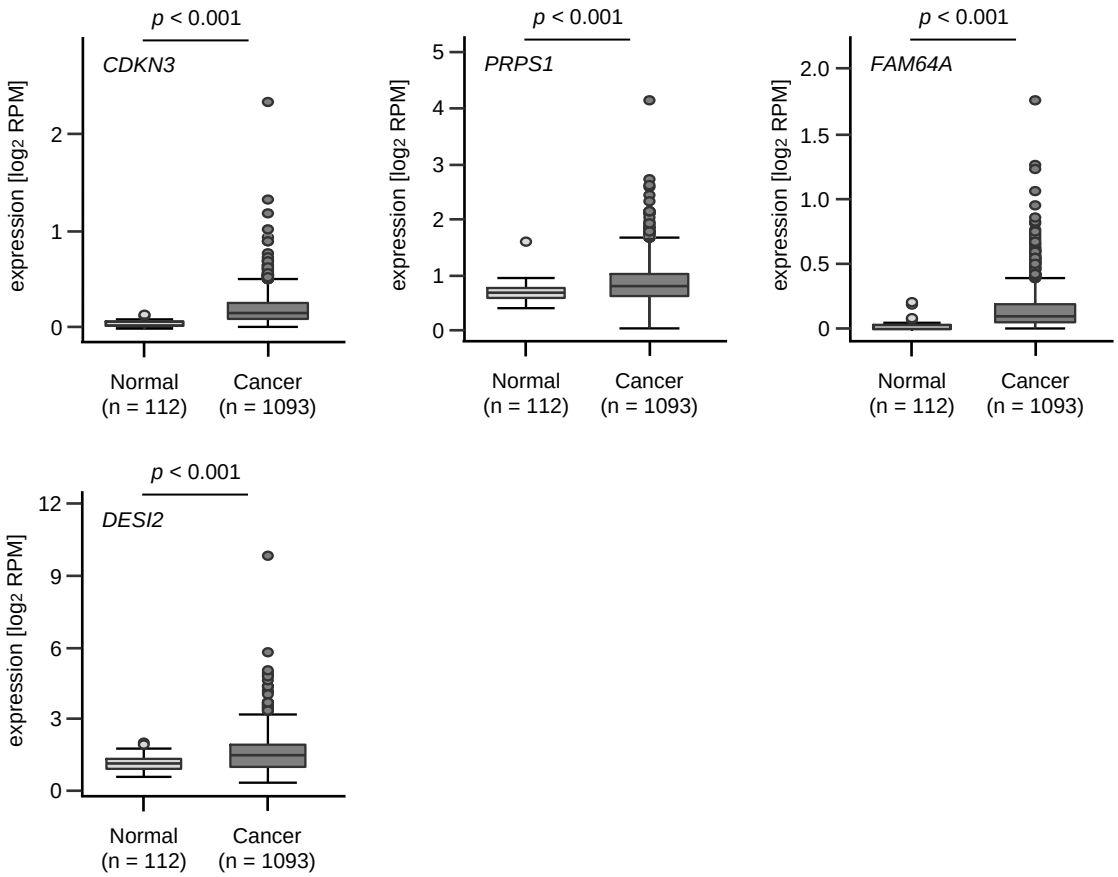

p < 0.001
CDKN3
expression [log2 RPM]
Normal (n = 112)
Cancer (n = 1093)
p < 0.001
PRPS1
expression [log2 RPM]
Normal (n = 112)
Cancer (n = 1093)
p < 0.001
FAM64A
expression [log2 RPM]
Normal (n = 112)
Cancer (n = 1093)
p < 0.001
DESI2
expression [log2 RPM]
Normal (n = 112)
Cancer (n = 1093)

Figure 4

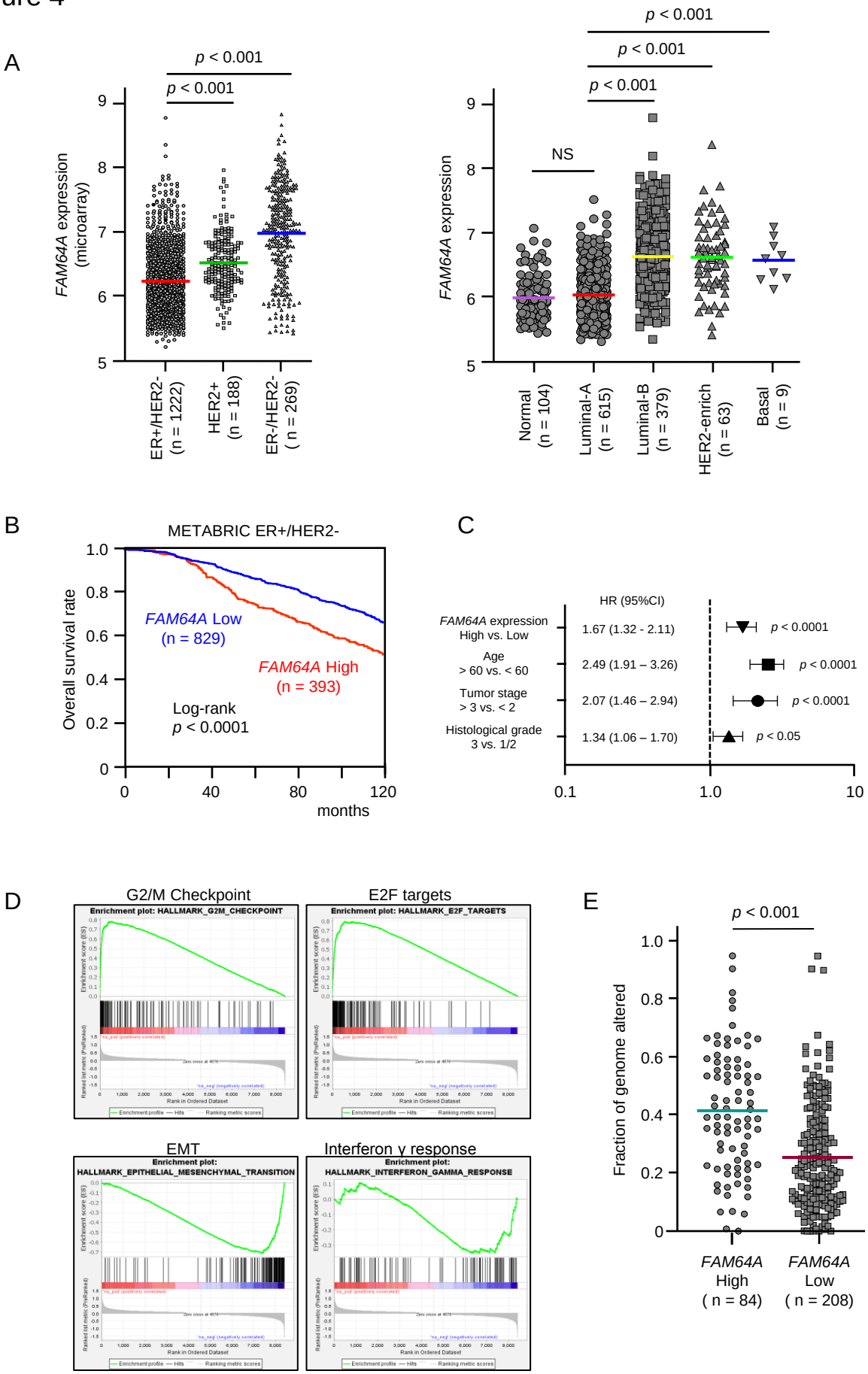
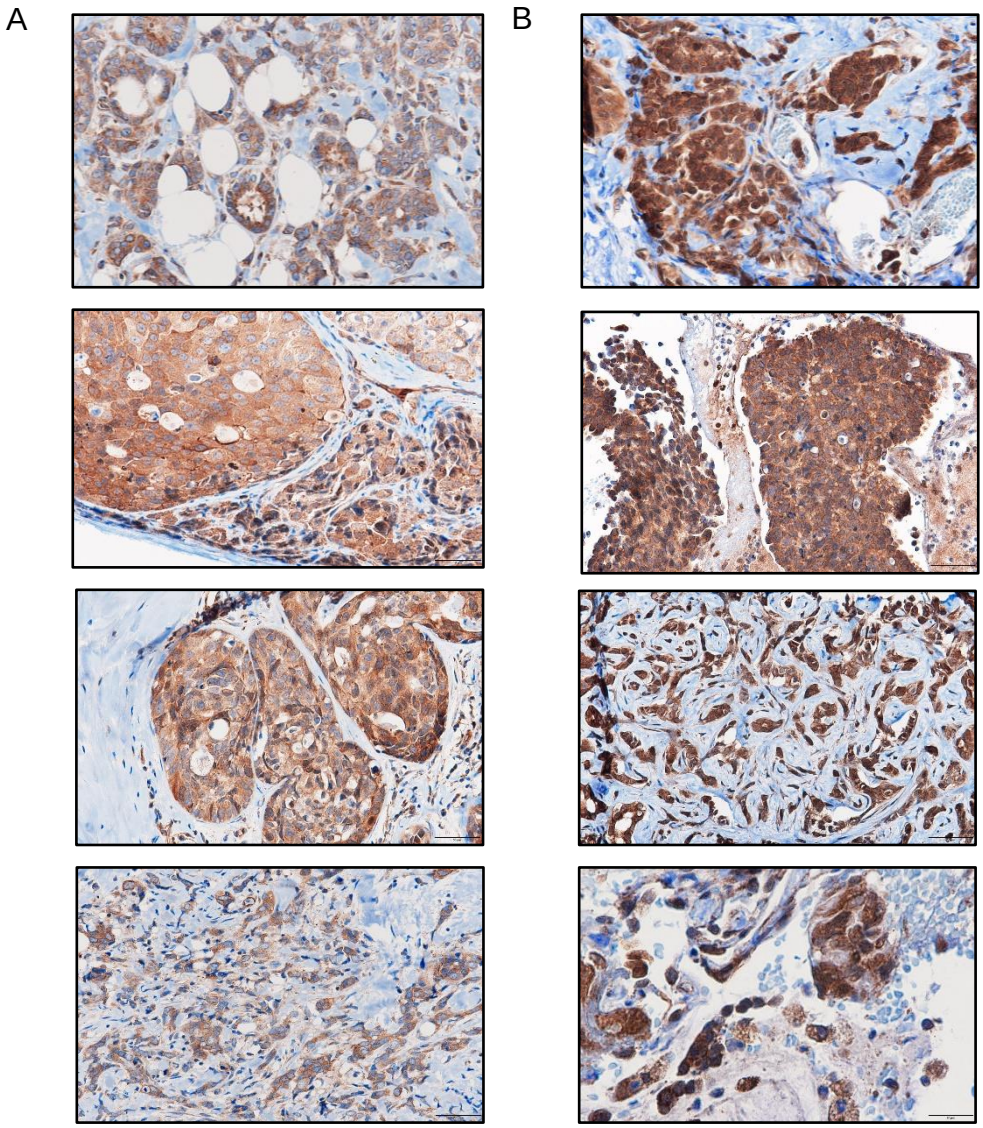
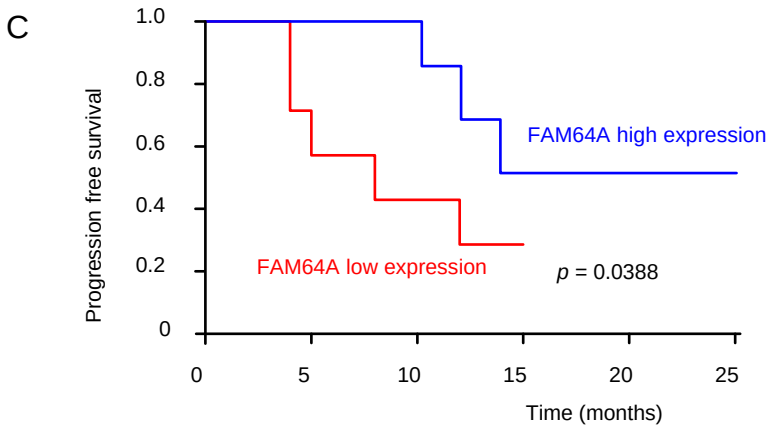

A
p < 0.001
p < 0.001
FAM64A expression (microarray)
ER+/HER2- (n = 1222)
HER2+ (n = 188)
ER-/HER2- (n = 269)
p < 0.001
p < 0.001
p < 0.001
NS
FAM64A expression
Normal (n = 104)
Luminal-A (n = 615)
Luminal-B (n = 379)
HER2-enrich (n = 63)
Basal (n = 9)
B
METABRIC ER+/HER2-
Overall survival rate
FAM64A Low (n = 829)
FAM64A High (n = 393)
Log-rank p < 0.0001
months
C
HR (95%CI)
FAM64A expression High vs. Low
1.67 (1.32 - 2.11)
p < 0.0001
Age > 60 vs. < 60
2.49 (1.91 – 3.26)
p < 0.0001
Tumor stage > 3 vs. < 2
2.07 (1.46 – 2.94)
p < 0.0001
Histological grade 3 vs. 1/2
1.34 (1.06 – 1.70)
p < 0.05
D
G2/M Checkpoint
Enrichment plot: HALLMARK_G2M_CHECKPOINT
E2F targets
Enrichment plot: HALLMARK_E2F_TARGETS
EMT
Enrichment plot: HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION
Interferon γ response
Enrichment plot: HALLMARK_INTERFERON_GAMMA_RESPONSE
E
p < 0.001
Fraction of genome altered
FAM64A High (n = 84)
FAM64A Low (n = 208)

Figure 5


A
B


C
1.0
0.8
0.6
0.4
0.2
0
Progression free survival
FAM64A high expression
FAM64A low expression
p = 0.0388
0
5
10
15
20
25
Time (months)