

Analysis of TCGA data of differentially expressed EMT-related genes and miRNAs across various malignancies to identify potential biomarkers

KONSTANTINOS A. KYRITSIS¹, MELPOMENI G. AKRIVOU¹,
LEFKI-PAVLINA N. GIASSAFAKI¹, NIKOLAOS G. GRIGORIADIS^{1,2} and IOANNIS S. VIZIRIANAKIS^{1,3,4}

¹Laboratory of Pharmacology, School of Pharmacy, Aristotle University of Thessaloniki, 54124 Thessaloniki;

²Biogenea Pharmaceuticals Ltd., 54627 Thessaloniki; ³Functional Proteomics and Systems Biology Research Group at AUTH (FunPATH) Research Group, KEDEK-Aristotle University of Thessaloniki, Balkan Center, 57001 Thessaloniki, Greece; ⁴Department of Life and Health Sciences, University of Nicosia, CY-1700 Nicosia, Cyprus

Received September 8, 2020; Accepted November 30, 2020

DOI: 10.3892/wasj.2020.77

Abstract. Tumor heterogeneity presents a hindering factor that leads to therapeutic failures and limits the improvement of clinical outcomes within the concept of precision medicine. This heterogeneous characteristic provides the epithelial mesenchymal plasticity that is considered an advantage for cancer cell metabolism and genome function to be adjusted within the microenvironment, and also plays a role in the development of drug resistance and metastasis. To this respect, identifying druggable molecular targets that modulate signaling networks, which contribute to cancer cell heterogeneity, could provide innovative therapeutics with improved safety and efficacy profiles. The present study attempted to identify potentially druggable molecular targets that have been connected to the process of epithelial-to-mesenchymal transition (EMT). Towards this goal, gene and miRNA differential expression analyses were performed for cancer patients with 4 and 3 different tumor types, respectively, using data that were retrieved from The Cancer Genome Atlas (TCGA) program. Furthermore, the dbEMT 1.0 database was used to limit the results to differentially expressed molecular targets that have already been associated with EMT. The analysis resulted in the identification of multiple EMT-associated genes and miRNAs for all types of cancer, which, through pairwise comparisons, were separated into groups of common potential targets for different malignancies. Differential gene expression profiling by RT-qPCR analysis was also carried out for a number of

selected genes and miR-21 in human cancer cell lines. Notably, EMT-associated homeobox B9 (HOXB9) and miR-137 were found to have a deregulated expression in all malignancies examined, thus increasing their potential as druggable targets for cancer therapy. Overall, the present study presents an approach that, through systematic *in silico* analysis, could lead to the selection of potential druggable biomarkers of broader utility for several tumor types, irrespective of their tissue of origin.

Introduction

Cellular and genomic heterogeneity in tumors involves the microenvironment, cancer stem cells and epithelial-to-mesenchymal transition (EMT) (1,2). EMT constitutes a complex and dynamic biological process during which epithelial cells transdifferentiate towards a mesenchymal phenotype. Unlike epithelial cells, which are characterized by polarity and maintain firmly cell-to-cell adhesion contacts through cellular adhesion molecules (CAMs), mesenchymal cells display an increased mobility and loose organization within the extracellular matrix. EMT plays an important role in physiological processes, including organ formation during embryogenesis and tissue regeneration; however, its involvement has also been confirmed in tumor initiation, progression and metastasis. The transition can be triggered through various stimuli, including different factors of the tumor microenvironment (cytokines, growth factors, etc.), as well as immune responses, hypoxia and antitumor drug treatment. Notably, EMT is reversible, exhibiting plasticity, with mesenchymal cells being capable of converting back to an epithelial phenotype through a process known as mesenchymal-to-epithelial transition (MET). The combination of EMT and MET can lead to a mixed, dynamic population of cancer cells exhibiting both epithelial and mesenchymal characteristics that also promote circulating tumor cell (CTC) formation. This can result in the disruption of cellular adhesion and increased migratory and invasive capabilities, which can lead to metastasis (3,4).

Correspondence to: Professor Ioannis S. Vizirianakis, Laboratory of Pharmacology, School of Pharmacy, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
E-mail: ivizir@pharm.auth.gr

Key words: epithelial-to-mesenchymal transition, the cancer genome atlas, RNA-Seq, miRNA-Seq, meta-analysis, druggable targets, anticancer therapeutics

The EMT phenotypic plasticity of tumor cells contributes to molecular and cellular heterogeneity, that leads to acquired drug resistance to cytotoxic or molecularly targeted therapy in clinical practice. Moreover, the differential pharmacological response limits the productivity and clinical outcomes of innovative therapeutic approaches (5-7). Moreover, the interplay of transcription (including the Snail, Twist and Zeb families) and epigenetic factors (e.g., the miR-200 family, miR-205, miR-203, miR-34 and miR-29b) drive the regulatory network program of EMT plasticity in cancer (6). It would be of interest if common molecular drivers in these EMT and MET processes that are deregulated in various types of tumors could be characterized; following clinical validation, biomarkers could be developed which may be used in cancer therapy.

Previously the authors characterized the expression levels of several epithelial markers, namely desmoglein 3 (DSG3), E-cadherin and β - γ -catenins (β - γ -catenins) in monolayer (ML) and multicellular aggregates (MCAs) of the HSC-3 cell line (oral squamous carcinoma) *in vitro*, as well as in clinical samples of oral leukoplakia (OL) and oral squamous cell carcinoma (OSCC) *in vivo* (8). Of note, the downregulation of DSG3, E-cadherin and β - γ -catenins was observed to be significantly associated with the grade of OL-dysplasia and OSCC samples (8). Furthermore, the switch of expression and potent perinuclear aggregation of DSG3 and γ -catenin were observed in both HSC-3 cells and OL/OSCC samples. These observations support the involvement of DSG3 and γ -catenin in the progression of oral epithelial cell malignancy. It was also suggested that these genes may serve as potential predictive biomarkers, along with E-cadherin and β -catenin, of the malignant transformation risk of oral dysplasia and the biological behavior (aggressiveness) of oral cancer, respectively (8).

In the present study, an *in silico* analysis was performed using RNA-Seq data from The Cancer Genome Atlas (TCGA) database (9), in order to identify genes that are involved in the EMT and MET processes, and that may serve as possible biomarkers and/or therapeutic targets in clinical practice. For this purpose, differential gene and microRNA (miRNA/miR) expression analyses were carried out between solid tissue normal (STN) and primary solid tumor (PST) samples from 4 different types of cancer (head and neck, prostate and breast cancer, and glioblastoma). The accurate separation of STN and PST samples into distinct clusters was confirmed using principal component analysis (PCA) based solely on the identified differentially expressed (DE) genes/miRNAs (Fig. 1). Additionally, dbEMT 1.0, a database containing EMT-related genes collected through extensive literature search (10), was used to further filter DE genes and miRNAs, and keep those that have been found to be involved in EMT and/or MET. On the whole, the present study identified and reported several DE and EMT/MET-related genes and miRNAs in various types of malignancies, whose potential in clinical utilization will be further evaluated by characterizing their expression levels in cell line EMT models and in clinical samples in the future.

Materials and methods

Cancer data. RNA-Seq and miRNA-Seq data from TCGA (9) were retrieved using TCGAbiolinks (11). Specifically, from the available pre-processed data types, gene and miRNA count

data, derived from HTSeq software (12), were selected for 4 different types of cancer: i) Head-and-neck squamous cell carcinoma (TCGA-HNSC); ii) breast cancer (TCGA-BRCA); iii) prostate adenocarcinoma (TCGA-PRAD); and iv) glioblastoma multiforme (TCGA-GBM).

Differential expression analysis. To ensure the power of statistical testing, only STN and PST samples were selected to perform differential expression analysis for genes and miRNAs using DESeq2 (13).

Due to the lack of sufficient miRNA-Seq samples in TCGA-GBM, miRNA differential expression analysis was performed only for TCGA-BRCA, TCGA-HNSC and TCGA-PRAD.

A minimum threshold of 100 and 10 total number of counts was set to filter out genes and miRNAs with very low counts, respectively. Genes and miRNAs with an absolute log2 fold change (LFC) >1 and a false discovery rate (FDR) (14) adjusted P-value <0.001 were reported as statistically significant, DE targets. Variance-stabilizing transformation (VST) was applied to DE gene or miRNA expression values of all samples, followed by PCA. Both VST and PCA were performed using DESeq2 (13).

EMT-associated genes and miRNAs. Genes and miRNAs that have been associated with EMT were retrieved from dbEMT 1.0, a database containing EMT-related genes and miRNAs that were collected through extensive literature search (10). Comparisons of EMT targets with DE genes and miRNAs was performed with R statistical programming language. Venn diagrams were created using limma (15).

Survival analysis of DE genes and miRNAs. Survival analysis was performed on DE genes and miRNAs using clinical metadata from TCGA. Specifically, for each DE gene or miRNA, PST samples were assigned into 2 separate groups, depending on whether the target expression of each sample was higher (high expression) or lower (low expression) than the median. Kaplan-Meier analysis on the 2 groups was performed for each DE gene or miRNA and selected, statistically significant targets are reported in Kaplan-Meier survival curves. For the statistical analysis of DE genes (P-value <0.001) and miRNAs (P-value <0.1) the non-parametric log-rank test was used. An exception to this was miR-16-1 in HNSC cancer samples, with a P-value of ~0.12. Survival analysis was performed using a survival analysis package (16,17) and survminer (18).

Cell cultures, RNA isolation and RT-qPCR analysis. The established cell lines of human breast epithelial carcinoma MCF-7 (Cellosaurus, CVCL-0031) and MDA-MB-231 (Cellosaurus, CVCL-0062), as well as the human tongue squamous carcinoma HSC-3 (Cellosaurus, CVCL-1288) and keratinocyte HaCaT (Cellosaurus, CVCL-0038) cancer cells that are routinely used in the authors' laboratory were cultured as previously described (19,20). Moreover, the RNA isolation, RT-cDNA synthesis, as well as the RT-qPCR analysis were carried out as previously described (19,20). In brief, total RNA was extracted from cells using TRIzol G (Panreac, Applichem), quantified using a Nanodrop ND-100

Table I. Number of genes and miRNAs identified as differentially expressed and EMT-associated in different malignancies.

TCGA project ID	Type of malignancy	Number of DE genes (LFC >1 and FDR adjusted P-value <0.001)	Number of DE Genes reported by dbEMT 1.0	Number of DE miRNAs (LFC >1 and FDR adjusted P-value <0.001)	Number of DE miRNAs reported by dbEMT 1.0
TCGA-HNSC	Head and neck squamous carcinoma	2,509	37	255	6
TCGA-BRCA	Breast cancer	4,227	73	271	10
TCGA-PRAD	Prostate cancer	1,049	11	110	4
TCGA-GBM	Glioblastoma multiforme	2,350	44		

DE, differentially expressed; EMT, epithelial-to-mesenchymal transition.

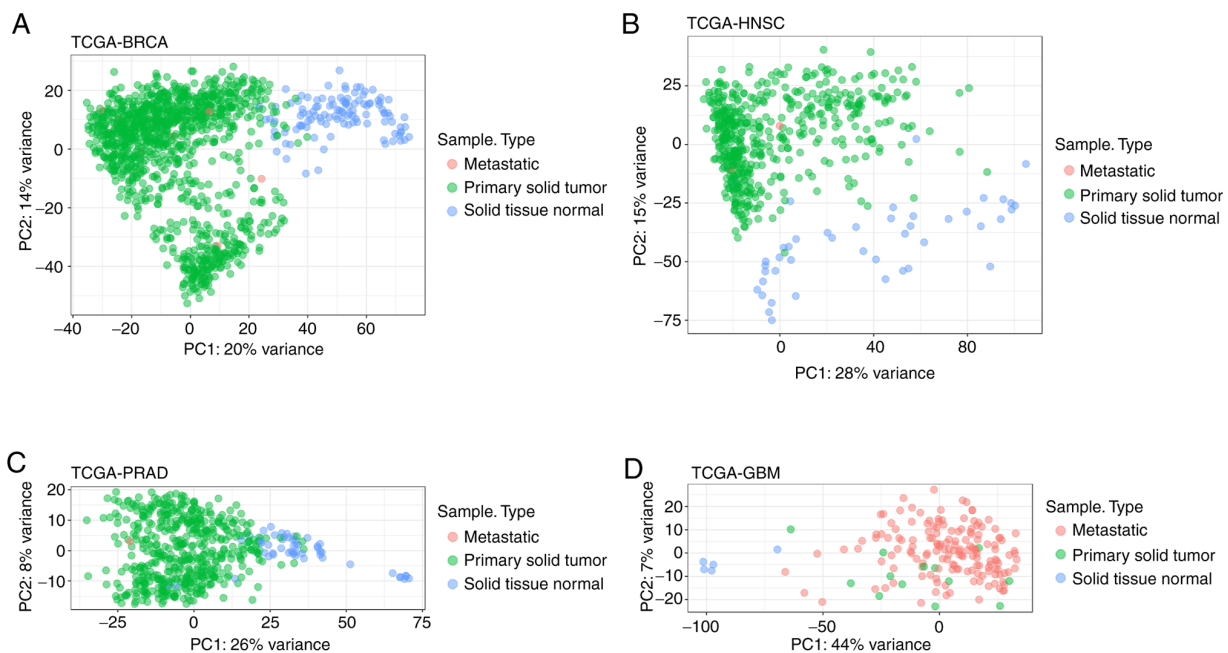


Figure 1. PCA of all samples from (A) breast cancer, (B) head and neck cancer, (C) prostate cancer and (D) glioblastoma using the expression values of differentially expressed genes. PCA was performed and scatterplots were created using DESeq2 (13). PCA, principal components analysis; TCGA, The Cancer Genome Atlas; BRCA, breast cancer; HNSC, head and neck squamous cell carcinoma; PRAD, prostate adenocarcinoma; GBM, glioblastoma multiforme.

Spectrometer and reverse transcribed into cDNA by applying the QuantiTect Reverse Transcription kit (Qiagen, Inc.). qPCR was performed on a 7500 Real-Time PCR System (Applied Biosystems; Thermo Fisher Scientific, Inc.) using KAPA SYBR® FAST qPCR Master Mix (KAPA Biosystems) under optimized conditions: 95°C for 3 min followed by 40 cycles at 95°C for 3 sec and 60°C for 20 sec. Primers designed and used during the present study were as follows: (PPARG forward, 5'-TCG-AGG-ACA-CCG-GAG-AGG-3' and reverse, 5'-CAC-GGA-GCT-GAT-CCC-AAA-GT-3'; HMG2 forward, 5'-GAA-AAA-CGG-CCA-AGA-GGC-AG-3' and reverse, 5'-AGA-GCT-ATC-CTG-GAC-TCC-TCC-3'; FOXM1 forward, 5'-ACC-GCT-ACT-TGA-CAT-TGG-AC-3' and reverse, 5'-GGG-AGT-TCG-GTT-TTG-ATG-GTC-3'; CAV-1 forward, 5'-CCC-AGG-GAA-ACC-TCC-TCA-CAG-3' and reverse, 5'-GGC-AGA-TAG-CAG-AAG-CGG-AC-3'; TGFB1-F

forward, 5'-ACT-GCG-GAT-CTC-TGT-GTC-ATT-G-3' and reverse, 5'-ACA-GTA-GTG-TTC-CCC-ACT-GGT-C-3'; Vimentin forward, 5'-GGC-TCG-TCA-CCT-TCG-TGA-AT-3' and reverse, 5'-GAG-AAA-TCC-TGC-TCT-CCT-CGC-3'; β -actin forward, 5'-TTG-CTG-ACA-GGA-TGC-AGA-AG-3' and reverse, 5'-TGA-TCC-ACA-TCT-GCT-GGA-AG-3'). β -actin was used as an endogenous control to normalize the gene expression levels.

The expression of miRNAs was also carried out by RT-qPCR using the miScript SYBR®-Green PCR kit (Qiagen, Inc.). Total cellular RNA extraction and quantification was performed as indicated above, whereas cDNA synthesis was executed with the miScript II RT kit (Qiagen, Inc.). Hsa-miR-21-5p (miR-21, 5'-UAGCUUAUCAGACUGAUG UUGA-3') was designed and used during this experiment and SNORD6 small nucleolar RNA, C/D box 6 (also known

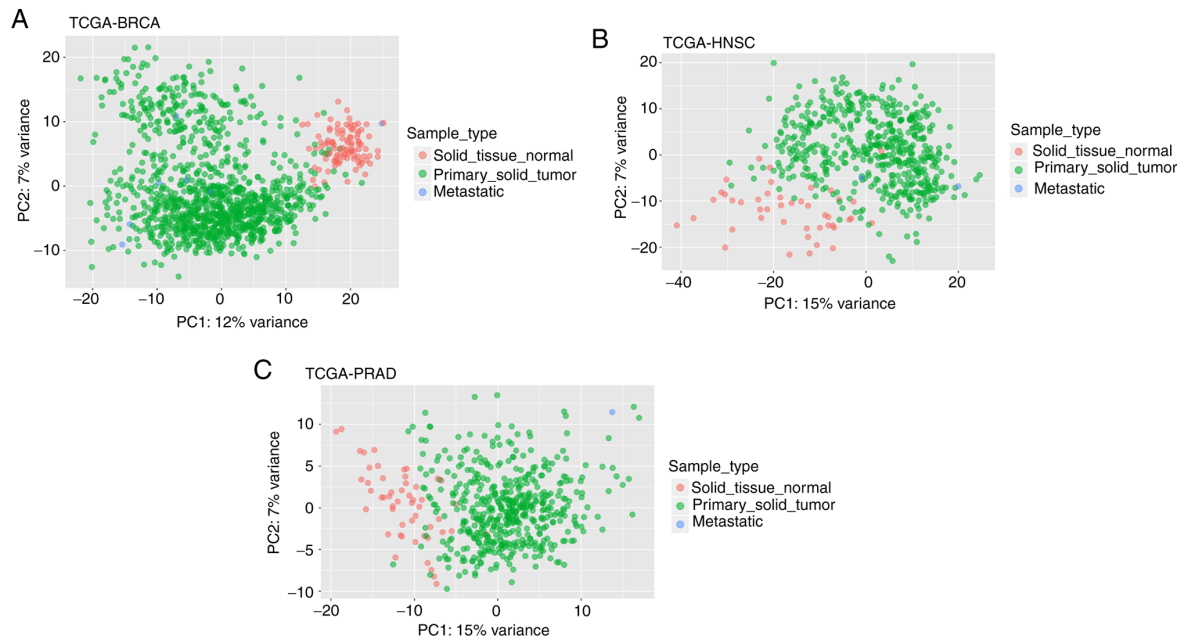


Figure 2. PCA of patient samples from (A) breast cancer, (B) head and neck cancer and (C) prostate cancer using the expression values of differentially expressed miRNAs (please see the ‘Materials and methods’ section). PCA was performed and scatterplots were created using DESeq2 (13). differentially expressed genes. PCA was performed and scatterplots were created using DESeq2 (13). PCA, principal components analysis; TCGA, The Cancer Genome Atlas; BRCA, breast cancer; HNSC, head and neck squamous cell carcinoma; PRAD, prostate adenocarcinoma.

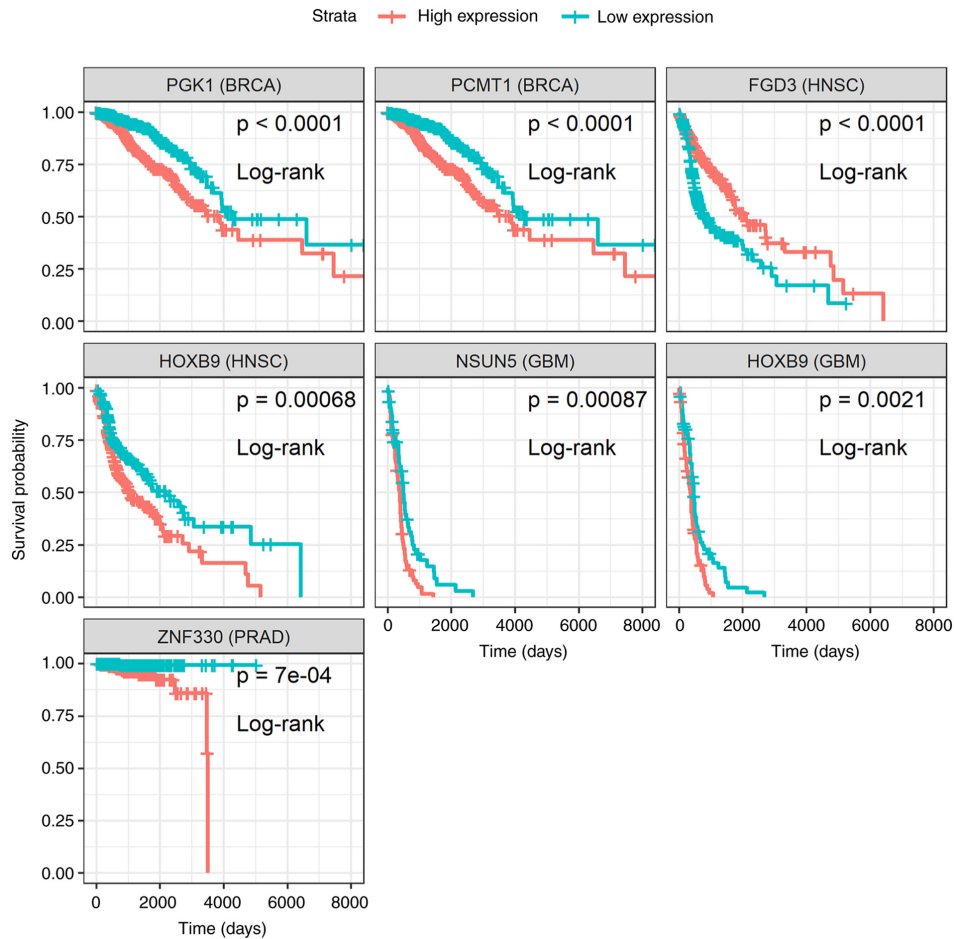


Figure 3. Kaplan-Meier survival curves of the differentially expressed genes, *PGK1*, *PCMT1*, *FGD3*, *HOXB9*, *NSUN5* and *ZNF330*, whose expression levels were significantly associated (P -value < 0.001 ; non-parametric log-rank test) with a favorable or poor survival probability of patients with breast (BRCA), head and neck (HNSC), glioblastoma (GBM) or prostate (PRAD) cancer. PST samples were assigned into two separate groups depending on whether target expression of each sample is higher (high expression) or lower (low expression) than the median. Survival analysis was performed and plots were created using survival methods (16,17) and survminer (18). PST, primary solid tumor; (Time: Is shown in days).

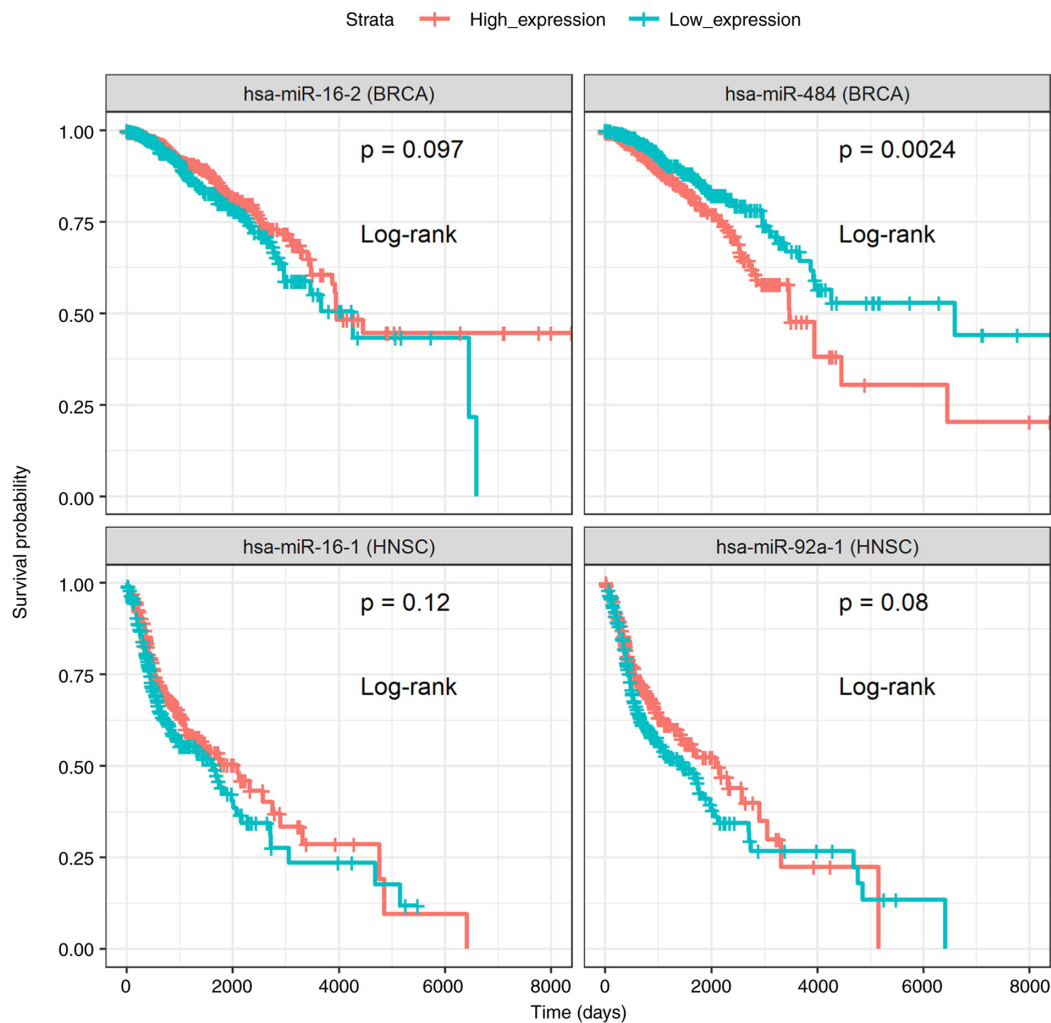


Figure 4. Kaplan-Meier survival curves of differentially expressed miRNAs miR-16-1/-2, miR-92-1 and miR-484, whose expression levels are significantly associated (P-value ~ 0.1 ; non-parametric log-rank test) with favorable or poor survival probability of patients with breast (BRCA) or head and neck (HNSC) cancer. PST samples were assigned into two separate groups depending on whether target expression of each sample is higher (high expression) or lower (low expression) than the median. Survival analysis was performed and plots were created using survival methods (16,17) and survminer (18). PST, primary solid tumor; (Time: Is shown in days).

as mgh28S-2412) (Qiagen, Inc.) was used as reference RNA gene. The reaction conditions consisted of polymerase activation/denaturation at 95°C for 15 min, followed by 40 cycles at 94°C for 15 sec and 55°C for 30 sec.

In both cases, the relative mRNA/miRNA concentrations were calculated using the $2^{-\Delta\Delta C_q}$ method (21) and the results obtained were represented as fold changes in the diagrams.

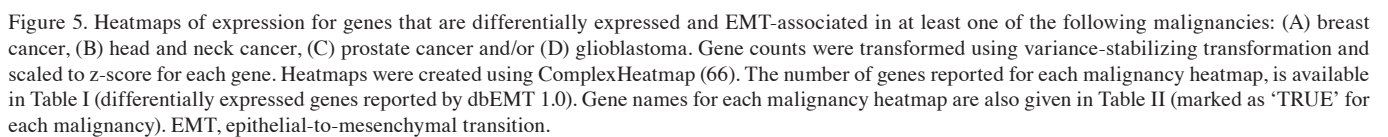
Statistical analysis. The results from 2 independent biological experiments (triplicate measurements) are shown and the data are expressed as the means \pm standard deviation (SE). Comparisons were carried out using a Student's t test, whereas the statistical analysis was performed using GraphPad Prism 6.0 (GraphPad Software, Inc.). $P < 0.05$ was considered to indicate a statistically significant difference.

Results

DE genes and miRNAs. Accessing the GDC data portal through TCGAbiolinks enabled the retrieval of RNA-Seq for 4 different types of malignancies. For each project, STN and

PST were identified as the predominant categories containing the majority of samples. Specifically, for TCGA-BRCA, TCGA-HNSC, TCGA-PRAD and TCGA-GBM, STN + PST samples numbering 113 + 1,102, 44 + 523, 52 + 498 and 5 + 156, respectively, were obtained, filtered for low count genes and compared to identify DE genes. Based on strict criteria (please see the 'Materials and methods' section) a subset of all genes analyzed was determined to be DE genes in each tumor (percentage of DE genes in each TCGA-project: TCGA-BRCA, 8.95%; TCGA-HNSC, 6.1%; TCGA-PRAD, 2.56%; TCGA-GBM, 6.52 %) (Table I). Similarly, miRNA-Seq data were retrieved for 3 TCGA projects and major sample categories STN and PST were used for DE miRNA identification (TCGA-BRCA, TCGA-HNSC and TCGA-PRAD with STN + PST samples numbering 104 + 1,096, 44 + 523, 52 + 498, respectively). Following analysis, the percentage of DE miRNAs were determined to be as follows: TCGA-BRCA, 19.1%; TCGA-HNSC, 18.33%; TCGA-PRAD, 9.04% (Table I).

To further confirm these findings, PCA analysis was performed using DE gene and miRNA expression values for all samples of each TCGA-project. Following dimension



Moreover, survival analysis was performed on DE genes and miRNAs (please see the ‘Materials and methods’ section). In total, 6 genes (Fig. 3) and 3 miRNAs (Fig. 4) were reported, for which patients with high or low expression levels presented considerable differences in survival. The results concerning the association of the genes *PGK1*, *PCMT1*, *FDG3*, *HOXB9*, *NSUN5* and *ZNF330* to patient survival probability, are in agreement with those previously reported by the Human Protein Atlas project (22), with the exception of the unfavorable prognosis of *HOXB9* overexpression in glioblastoma, which was identified during the current analysis (Fig. 3). As regards the miRNAs,

EMT-associated genes and miRNAs. In total, 344 genes and 20 miRNAs that, following an exhaustive literature search, constitute a collection of well-characterized EMT-associated targets, were retrieved from dbEMT 1.0 (10). Direct comparisons between the 2 gene collections revealed that only a small fraction of DE genes has been identified as directly related to the EMT process (percentage of DE genes that were EMT-related: TCGA-BRCA, 1.7%; TCGA-HNSC, 1.5%; TCGA-PRAD, 1%; TCGA-GBM, 1.9%) (Table I). A

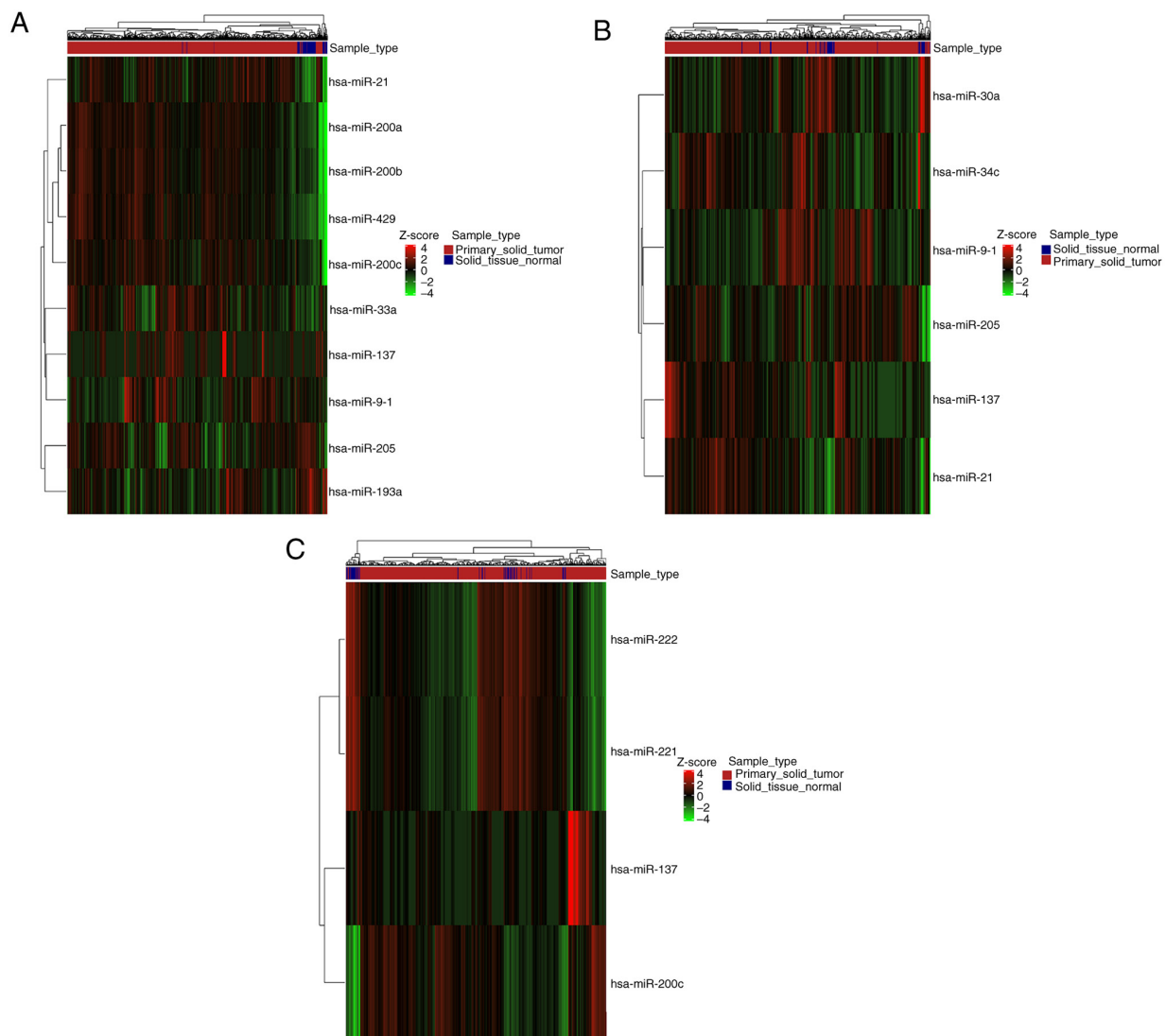


Figure 6. Heatmaps of expression for miRNAs that are differentially expressed and EMT-associated in at least one of the following malignancies: (A) breast cancer, (B) head and neck cancer and/or (C) prostate cancer. miRNA counts were transformed using variance-stabilizing transformation and scaled to z-score for each miRNA. Heatmaps were created using ComplexHeatmap (66). The number of miRNAs reported for each malignancy heatmap, is available in Table I (differentially expressed miRNAs reported by dbEMT 1.0). miRNA names for each malignancy heatmap are also given in Table III (marked as 'TRUE' for each malignancy). EMT, epithelial-to-mesenchymal transition.

small number of miRNAs reported in dbEMT 1.0 was also found in the present collection of DE miRNAs (percentage of DE miRNAs that were EMT-related: TCGA-BRCA, 3.7%; TCGA-HNSC, 2.4%; TCGA-PRAD, 3.6%) (Table I).

Furthermore, DE genes and miRNAs related to EMT were found both up- and downregulated in each type of cancer (Figs. 5 and 6). Careful inspection of these results is required to decipher the role of each gene and miRNA in EMT, at the context of each malignancy.

EMT targets amongst different malignancies. With an aim of identifying targets that are commonly deregulated and EMT-associated between different types of cancer, pairwise comparisons of DE genes and miRNAs associated with EMT were performed. It was found that different types of malignancies shared several deregulated molecules (Tables II and III), with *HOXB9* and miR-137 being common for all types of malignancies that were examined (Fig. 7).

Assessment of DE genes and miRNAs related to EMT in breast, and head and neck cell carcinoma lines. Based on the data obtained, the present study wished to assess the expression level in a number of DE genes in well-characterized human breast, and head and neck cancer cell lines. To this end, the selection was made for caveolin-1, FOXM1 and Vimentin for the human breast epithelial carcinoma cell lines, MCF-7 and MDA-MB-231. As for the head and neck cancer cell lines, the human oral HSC-3 and keratinocyte HaCaT cancer cells were used to assess the gene expression levels of HMGA2, TGFB1, FOXM1 and PPARG. Moreover, from the DE miRNAs, the expression of miR-21 was selected and was assessed in all these 4 cell lines.

As shown in Fig. 8, for the breast epithelial carcinoma cell lines, the expression of caveolin-1 (Fig. 8A) was markedly higher in the MDA-MB-231 cells than in the MCF-7 cells ($P < 0.001$). As regards the expression of miR-21 (Fig. 8B), it was higher again in the MDA-MB-231 cells compared to the MCF-7 cells,

Table II. Differentially expressed and EMT-associated genes in different malignancies.

Gene symbol	Ensemble gene ID	TCGA_BRCA	TCGA_HNSC	TCGA_PRAD	TCGA_GBM
ZEB2	ENSG00000169554	TRUE	FALSE	FALSE	FALSE
EGFR	ENSG00000146648	TRUE	FALSE	FALSE	TRUE
EPAS1	ENSG00000116016	TRUE	FALSE	FALSE	FALSE
ERBB2	ENSG00000141736	TRUE	FALSE	FALSE	FALSE
MET	ENSG00000105976	TRUE	FALSE	FALSE	FALSE
CDH2	ENSG00000170558	TRUE	FALSE	FALSE	FALSE
KLF4	ENSG00000136826	TRUE	FALSE	FALSE	FALSE
KLF6	ENSG00000067082	TRUE	FALSE	FALSE	FALSE
WT1	ENSG00000184937	TRUE	TRUE	FALSE	FALSE
LEF1	ENSG00000138795	TRUE	FALSE	FALSE	FALSE
SIM2	ENSG00000159263	TRUE	FALSE	TRUE	FALSE
FN1	ENSG00000115414	TRUE	TRUE	FALSE	TRUE
EGR1	ENSG00000120738	TRUE	FALSE	FALSE	FALSE
KIT	ENSG00000157404	TRUE	TRUE	TRUE	FALSE
BMP2	ENSG00000125845	TRUE	FALSE	FALSE	FALSE
CAV1	ENSG00000105974	TRUE	FALSE	TRUE	FALSE
PPARG	ENSG00000132170	TRUE	TRUE	FALSE	FALSE
TGFBR3	ENSG00000069702	TRUE	TRUE	TRUE	FALSE
FOXA1	ENSG00000129514	TRUE	FALSE	FALSE	FALSE
STAT5A	ENSG00000126561	TRUE	FALSE	FALSE	FALSE
GATA3	ENSG00000107485	TRUE	FALSE	TRUE	FALSE
ANXA1	ENSG00000135046	TRUE	TRUE	FALSE	TRUE
DDR2	ENSG00000162733	TRUE	FALSE	FALSE	FALSE
FOXM1	ENSG00000111206	TRUE	TRUE	FALSE	TRUE
HSPB1	ENSG00000106211	TRUE	FALSE	FALSE	FALSE
VIM	ENSG00000026025	TRUE	FALSE	FALSE	TRUE
SMAD9	ENSG00000120693	TRUE	FALSE	FALSE	FALSE
GSN	ENSG00000148180	TRUE	FALSE	FALSE	FALSE
CYR61	ENSG00000142871	TRUE	FALSE	FALSE	FALSE
MST1R	ENSG00000164078	TRUE	FALSE	FALSE	FALSE
VCAN	ENSG00000038427	TRUE	FALSE	FALSE	FALSE
MYCN	ENSG00000134323	TRUE	FALSE	FALSE	FALSE
TCF21	ENSG00000118526	TRUE	FALSE	FALSE	FALSE
HMGA2	ENSG00000149948	TRUE	TRUE	FALSE	FALSE
CDKN2A	ENSG00000147889	TRUE	TRUE	FALSE	FALSE
MUC1	ENSG00000185499	TRUE	FALSE	FALSE	FALSE
AURKA	ENSG00000087586	TRUE	TRUE	FALSE	TRUE
FGF2	ENSG00000138685	TRUE	FALSE	FALSE	FALSE
MCAM	ENSG00000076706	TRUE	FALSE	FALSE	FALSE
PAX2	ENSG00000075891	TRUE	FALSE	TRUE	FALSE
PTPN14	ENSG00000152104	TRUE	FALSE	FALSE	FALSE
SHH	ENSG00000164690	TRUE	FALSE	FALSE	FALSE
SPP1	ENSG00000118785	TRUE	TRUE	FALSE	FALSE
CDH13	ENSG00000140945	TRUE	FALSE	FALSE	FALSE
VTN	ENSG00000109072	TRUE	FALSE	FALSE	FALSE
FBLN5	ENSG00000140092	TRUE	FALSE	FALSE	FALSE
KRT19	ENSG00000171345	TRUE	FALSE	FALSE	FALSE
EZH2	ENSG00000106462	TRUE	FALSE	FALSE	TRUE
ECT2	ENSG00000114346	TRUE	TRUE	FALSE	FALSE
PLAUR	ENSG00000011422	TRUE	FALSE	FALSE	FALSE
MMP9	ENSG00000100985	TRUE	TRUE	FALSE	TRUE
PTN	ENSG00000105894	TRUE	FALSE	FALSE	FALSE

Table II. Continued.

Gene symbol	Ensemble gene ID	TCGA_BRCA	TCGA_HNSC	TCGA_PRAD	TCGA_GBM
POSTN	ENSG00000133110	TRUE	TRUE	FALSE	TRUE
CXCL12	ENSG00000107562	TRUE	FALSE	FALSE	FALSE
EPO	ENSG00000130427	TRUE	FALSE	FALSE	FALSE
FGF1	ENSG00000113578	TRUE	FALSE	FALSE	FALSE
DLX4	ENSG00000108813	TRUE	FALSE	FALSE	FALSE
MMP3	ENSG00000149968	TRUE	TRUE	FALSE	FALSE
LEP	ENSG00000174697	TRUE	FALSE	FALSE	FALSE
PRSS8	ENSG00000052344	TRUE	FALSE	FALSE	FALSE
MMP13	ENSG00000137745	TRUE	TRUE	FALSE	FALSE
KL	ENSG00000133116	TRUE	FALSE	FALSE	FALSE
HOXB9	ENSG00000170689	TRUE	TRUE	TRUE	TRUE
SLC39A6	ENSG00000141424	TRUE	FALSE	FALSE	FALSE
SDC1	ENSG00000115884	TRUE	FALSE	FALSE	FALSE
GIPC2	ENSG00000137960	TRUE	FALSE	FALSE	FALSE
HMGB3	ENSG00000029993	TRUE	FALSE	FALSE	FALSE
TMPRSS4	ENSG00000137648	TRUE	FALSE	FALSE	FALSE
RGCC	ENSG00000102760	TRUE	FALSE	FALSE	FALSE
VWCE	ENSG00000167992	TRUE	FALSE	FALSE	FALSE
MUC2	ENSG00000198788	TRUE	FALSE	FALSE	FALSE
KCNH1	ENSG00000143473	TRUE	FALSE	FALSE	TRUE
HSPB2	ENSG00000170276	TRUE	TRUE	FALSE	FALSE
TGFB1	ENSG00000105329	FALSE	TRUE	FALSE	FALSE
SNAI2	ENSG00000019549	FALSE	TRUE	FALSE	FALSE
EGF	ENSG00000138798	FALSE	TRUE	FALSE	FALSE
TNC	ENSG00000041982	FALSE	TRUE	FALSE	TRUE
ITGA6	ENSG00000091409	FALSE	TRUE	FALSE	FALSE
PTHLH	ENSG00000087494	FALSE	TRUE	FALSE	FALSE
ITGA5	ENSG00000161638	FALSE	TRUE	FALSE	TRUE
ROR2	ENSG00000169071	FALSE	TRUE	FALSE	FALSE
CLDN4	ENSG00000189143	FALSE	TRUE	FALSE	FALSE
HPGD	ENSG00000164120	FALSE	TRUE	FALSE	FALSE
GREM1	ENSG00000166923	FALSE	TRUE	FALSE	FALSE
CLU	ENSG00000120885	FALSE	TRUE	TRUE	FALSE
LAMA1	ENSG00000101680	FALSE	TRUE	FALSE	FALSE
PROM1	ENSG00000007062	FALSE	TRUE	FALSE	FALSE
LOXL2	ENSG00000134013	FALSE	TRUE	FALSE	TRUE
FSCN1	ENSG00000075618	FALSE	TRUE	FALSE	FALSE
COL8A1	ENSG00000144810	FALSE	TRUE	FALSE	TRUE
EIF5A2	ENSG00000163577	FALSE	TRUE	FALSE	FALSE
PDPN	ENSG00000162493	FALSE	TRUE	FALSE	TRUE
TWIST1	ENSG00000122691	FALSE	FALSE	TRUE	FALSE
PTGS2	ENSG00000073756	FALSE	FALSE	TRUE	FALSE
FOXQ1	ENSG00000164379	FALSE	FALSE	TRUE	FALSE
TP53	ENSG00000141510	FALSE	FALSE	FALSE	TRUE
MAP2K1	ENSG00000169032	FALSE	FALSE	FALSE	TRUE
PRKCE	ENSG00000171132	FALSE	FALSE	FALSE	TRUE
CD44	ENSG00000026508	FALSE	FALSE	FALSE	TRUE
MMP2	ENSG00000087245	FALSE	FALSE	FALSE	TRUE
YBX1	ENSG00000065978	FALSE	FALSE	FALSE	TRUE
TGFB1I1	ENSG00000140682	FALSE	FALSE	FALSE	TRUE
CXCR4	ENSG00000121966	FALSE	FALSE	FALSE	TRUE
MDK	ENSG00000110492	FALSE	FALSE	FALSE	TRUE

Table II. Continued.

Gene symbol	Ensemble gene ID	TCGA_BRCA	TCGA_HNSC	TCGA_PRAD	TCGA_GBM
MSN	ENSG00000147065	FALSE	FALSE	FALSE	TRUE
SIX1	ENSG00000126778	FALSE	FALSE	FALSE	TRUE
S100A4	ENSG00000196154	FALSE	FALSE	FALSE	TRUE
PAK1	ENSG00000149269	FALSE	FALSE	FALSE	TRUE
IGFBP3	ENSG00000146674	FALSE	FALSE	FALSE	TRUE
MMP14	ENSG00000157227	FALSE	FALSE	FALSE	TRUE
ST14	ENSG00000149418	FALSE	FALSE	FALSE	TRUE
MKL2	ENSG00000186260	FALSE	FALSE	FALSE	TRUE
ETV4	ENSG00000175832	FALSE	FALSE	FALSE	TRUE
WNT1	ENSG00000125084	FALSE	FALSE	FALSE	TRUE
LOX	ENSG00000113083	FALSE	FALSE	FALSE	TRUE
LIMA1	ENSG00000050405	FALSE	FALSE	FALSE	TRUE
GRIN1	ENSG00000176884	FALSE	FALSE	FALSE	TRUE
LOXL3	ENSG00000115318	FALSE	FALSE	FALSE	TRUE
VSNL1	ENSG00000163032	FALSE	FALSE	FALSE	TRUE
IDH1	ENSG00000138413	FALSE	FALSE	FALSE	TRUE
CAMK1D	ENSG00000183049	FALSE	FALSE	FALSE	TRUE
MGAT3	ENSG00000128268	FALSE	FALSE	FALSE	TRUE
HAS2	ENSG00000170961	FALSE	FALSE	FALSE	TRUE

EMT, epithelial-to-mesenchymal transition; The Cancer Genome Atlas; BRCA, breast cancer; HNSC, head and neck squamous cell carcinoma; PRAD, prostate adenocarcinoma; GBM, glioblastoma multiforme.

Table III. Differentially expressed and EMT-associated miRNAs in different malignancies.

miRNA	TCGA_BRCA	TCGA_HNSC	TCGA_PRAD
hsa-mir-137	TRUE	TRUE	TRUE
hsa-mir-193a	TRUE	FALSE	FALSE
hsa-mir-200a	TRUE	FALSE	FALSE
hsa-mir-200b	TRUE	FALSE	FALSE
hsa-mir-200c	TRUE	FALSE	TRUE
hsa-mir-205	TRUE	TRUE	FALSE
hsa-mir-21	TRUE	TRUE	FALSE
hsa-mir-33a	TRUE	FALSE	FALSE
hsa-mir-9-1	TRUE	TRUE	FALSE
hsa-mir-429	TRUE	FALSE	FALSE
hsa-mir-30a	FALSE	TRUE	FALSE
hsa-mir-34c	FALSE	TRUE	FALSE
hsa-mir-221	FALSE	FALSE	TRUE
hsa-mir-222	FALSE	FALSE	TRUE

EMT, epithelial-to-mesenchymal transition; The Cancer Genome Atlas; BRCA, breast cancer; HNSC, head and neck squamous cell carcinoma; PRAD, prostate adenocarcinoma; GBM, glioblastoma multiforme.

although the difference was not statistically significant. Such gene and miR-21 differential expression profiles may contribute to the observed metabolic and phenotypic behavior of these 2

cell lines. Indeed, although they are both invasive ductal breast carcinoma cells, they have a number of phenotypic and genotypic differences. MCF-7 are estrogen receptor-positive cells, whereas MDA-MB-231 cells are estrogen and progesterone receptor-negative; in addition, MCF-7 cells express the epithelial phenotype in contrast to the MDA-MB-231 cells that are more mesenchymal (30-32).

Similarly, as shown in Fig. 9, the expression profiles of the FOXM1 and HMGA2 genes exhibited higher levels in the HSC-3 as compared to the HaCaT cells (Fig. 9A), although this difference was not statistically significant. On the contrary, miR-21 was found to have an increased expression in HaCaT compared to HSC-3 cells, a result that was again, not significant. It is interesting that the pattern of expression of these EMT-related genes and that of miR-21 varies in these 2 cell lines; however, it is still not known to what extent such behavior contributes to any metabolic and phenotypic property seen in these 2 cell lines.

Discussion

Through preliminary *in silico* analysis, the present study identified a list of genes and miRNAs that were DE and associated with EMT and/or MET in different types of cancer, such as head and neck, breast and prostate cancer, and glioblastoma. Moreover, RT-qPCR analysis revealed differential expression profiles of selected EMT-related genes and miR-21 in a number of human breast, and head and neck carcinoma cell lines. TCGA-BRCA and -HNSC cancer samples shared the most DE and dbEMT 1.0 reported genes and miRNAs, in accordance with their common epithelial tissue origin. The present study,

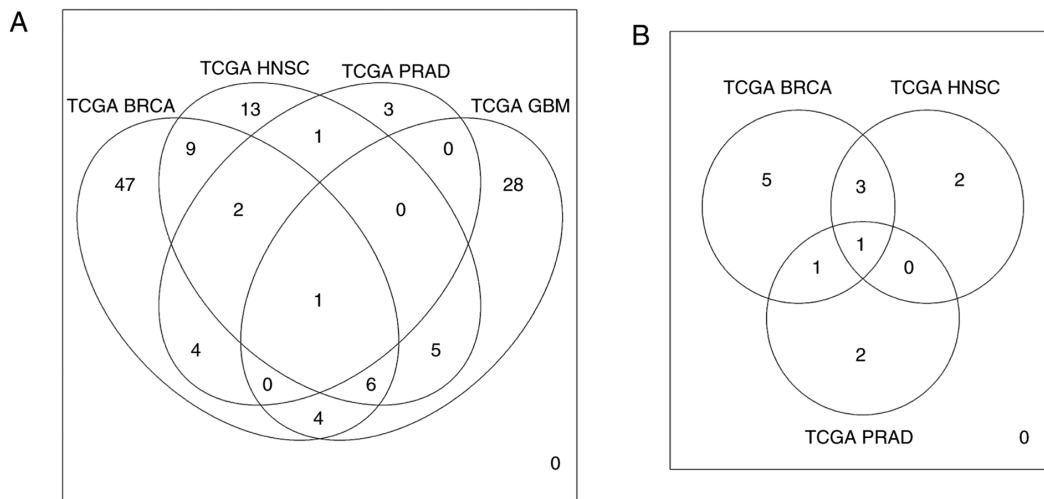


Figure 7. Venn diagrams showing common and unique differentially expressed and EMT reported (A) genes and (B) miRNAs for all types of malignancies that were analyzed. Plots were created using limma (15). EMT, epithelial-to-mesenchymal transition; The Cancer Genome Atlas; BRCA, breast cancer; HNSC, head and neck squamous cell carcinoma; PRAD, prostate adenocarcinoma; GBM, glioblastoma multiforme.

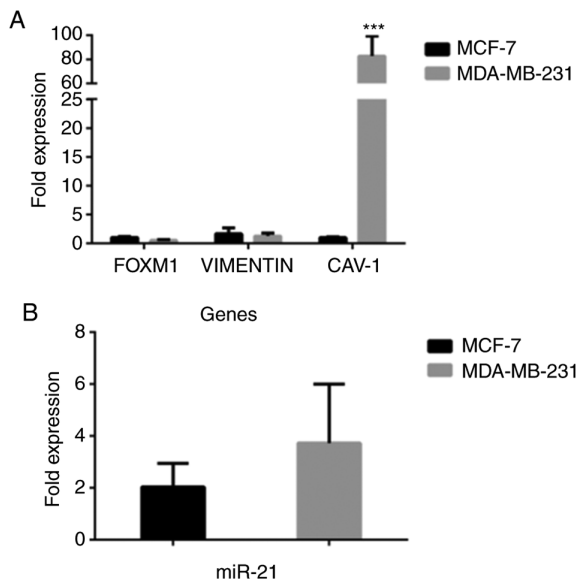


Figure 8. Assessment of expression profiles of (A) EMT-related genes and (B) miR-21 by RT-qPCR analysis of the MCF-7 and MDA-MB-231 cell lines. EMT, epithelial-to-mesenchymal transition. The results represent the means \pm SD (A) *** $P < 0.001$; (B) no statistically significant difference was observed; Student's t-test).

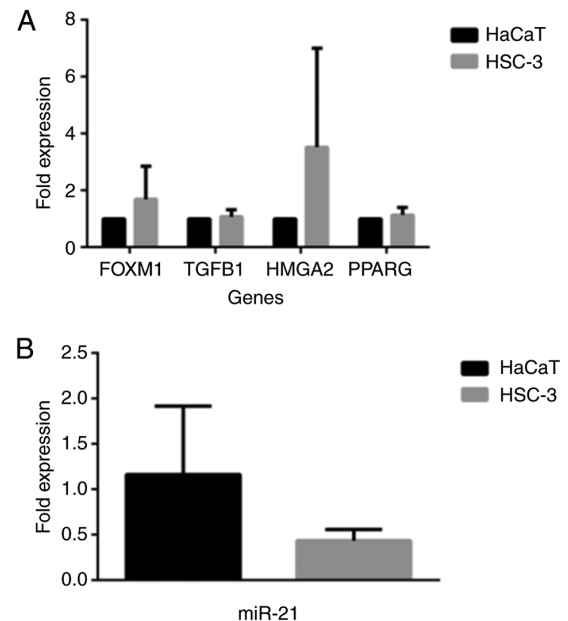


Figure 9. Assessment of expression profiles of (A) EMT-related genes and (B) miR-21 by RT-qPCR analysis of the HSC-3 and HaCaT cell lines. EMT, epithelial-to-mesenchymal transition. The results shown represent the mean \pm SD (no statistically significant difference is observed; Student's t-test).

by analyzing the expression profiles of EMT-related genes and miRNAs in patient cancer samples, suggests that HOXB9 and miR-137 present the same deregulated patterns, independent of tumor type. HOXB9 belongs to HOX gene family that in human plays crucial role in physiology and pathophysiology, by modulating cell development, differentiation and growth. It is noteworthy that the aberrant expression of HOX genes has been shown to contribute to cancer progression and development (33,34). Indeed, it has been observed that HOX genes exhibit a dysregulated expression in leukemia, ovarian and lung cancer (35-38). The function of HOXB9 novel tumor suppressor in the regulation of colon adenocarcinoma progression has also been identified (39). Moreover, it has been shown

that HOXB9 is associated with the emergence of radioresistance, as well as the development of resistance in anti-VEGF therapy in colorectal cancer (40,41). Importantly, a recent study identified HOXB9 as one key gene in a 5-gene molecular prognostic signature in patients with laryngeal cancer (42). In addition, the implication of HOX9 in prostate cancer cell progression has been recently proposed (43). Furthermore, the present study reports HOXB9, along with PGK1 (44,45), PCMT1 (46,47), NSUN5 (48,49) and ZNF330 (50), as unfavorable markers for the survival of patients with the types of cancer examined herein. Of note, the present study demonstrated showed that HOXB9 overexpression was associated with a poor prognosis for patients with both head and neck

cancer, previously reported in Human Protein Atlas (22), and those with glioblastoma. An exception was FGD3 (51,52), with its overexpression being predictive of a favorable outcome for patients with head and neck cancer (Fig. 3).

miRNAs represent important players in the post-transcriptional regulation of gene expression. In this manner, they affect signaling pathways and cellular processes with implication in cancer progression and development (53-56). miR-137 has shown to control tumorigenesis, invasion and metastasis in pancreatic neuroendocrine tumors (56). Moreover, miR-137 exhibits crucial developmental roles in neuronal differentiation (57). In addition, miR-137, by modulating SLC1A5-dependent glutamine uptake, is involved in the progression of head and neck squamous cell carcinoma (58,59). The significant role of miR-137 in the progression, diagnosis and prognosis of hepatocellular carcinoma has also been documented (60). The therapeutic potential of targeting miR-137 in non-small cell lung cancer (NSCLC) has been recently proposed (61). By retrospectively analyzing tumor patient data, the task of identifying potential druggable genes and miRNAs related to the process of epithelial mesenchymal plasticity and exhibiting deregulated expression levels in different malignancies has been set forth. Moreover, the data presented in the present study justify the affordability of identifying common druggable cancer biomarkers applicable to various tumors, in order to proceed thereafter, through the pharmacological assessment, to the development of successful anticancer therapeutics. Of note, HOXB9 and miR-137 were found to be deregulated in all types of malignancies that were analyzed. However, both were expressed in very low levels compared to other genes and/or miRNAs. Therefore, careful examination is required upon attempting to further clinically validate their usefulness and therapeutic applicability, as several DE genes and miRNAs related to EMT are also shared by different types of cancer that were analyzed herein. To this end, the regulatory network of miRNAs in EMT plasticity has been previously evaluated in breast cancer (62). Moreover, the EMT regulatory network includes a number of EMT-related transcription factors (e.g., the Snail and Zeb family) and epigenetic collaborative regulators (e.g., miR-34 for Snail and miR-200s for Zeb). Such interplay drives the well-orchestrated epithelial-mesenchymal transcriptional program, thus mediating the downstream biological effects (6). In the present study, the bioinformatics analysis focused on EMT-related miRNAs and genes dysregulated in various origin tumor patient samples and thus the findings obtained highlight only such conclusions. Whether or not there exists any functional involvement between the miRNAs and the genes identified, needs to be experimentally validated. It is interesting, however, to note that recent data highlight the yet unexplored role of miRNAs as regulators of Hox genes in hematopoiesis, through the elucidation of the role of miR-708 as a novel regulator of the Hoxa9 program in leukemia myeloid cells (63).

Overall, the analysis approach, is further strengthening the previously published data regarding the modulation of EMT in tumors and proposes that targeted research efforts focused on identifying common biomarkers could provide effective anticancer drugs. The results obtained support the notion that as such, druggable biomarkers could be considered the HOXB9 gene and/or miR-137, irrespective of the cured tumor type, although further clinical and experimental studies are also

needed. Importantly, however, such a direction is expected to provide valuable therapeutic interventions in malignancies by contributing toward overcoming the existed cellular and genomic heterogeneity (inter- and intra-tumoral) and the differential pharmacological response seen among patients. In particular, the data provided herein support the notion of identifying drug-gable biomarkers that impinge on the fundamental cancer cell traits that provide the needed advantageous capacity of tumor cell metabolism to abrogate the molecular balance, as well as the existing physiological restriction signals between differentiation, apoptosis and proliferation. This notion of the pan-cancer clinical intervention and the implementation of informed clinical decisions are based on the profiling of genomic signatures and molecular biomarkers in cancer patients; the origin and type of histology of the tumor have already begun to be left. Indeed, the development of therapeutics showing pan-cancer capabilities present a new revolutionary therapeutic era, an approach mentioned as ‘tumor-agnostic therapies’. Complementary to this, the ability to identify and clinically validate cancer biomarkers working irrespectively of the tumor type, permits the implementation of personalized cancer therapy in the clinical setting. The already marketed anticancer drugs, pembrolizumab, larotrectinib and entrectinib, belong to the class of tumor-agnostic therapies, by successfully receiving approval and being clinical used in patients with various types of tumor bearing common molecular features (64). Furthermore, the ability to implement cancer therapy with pharmacogenomics-guided therapeutic decisions offers the needed precision in clinical practice for the practical utilization of molecular profiling and biomarkers, as well as the outcome improvement in patients (65).

Acknowledgements

The statistical methods used in the present study were reviewed by Angelis Eleftherios, Professor of Statistics and Information Systems, Head of the School of Informatics, Faculty of Sciences, Aristotle University of Thessaloniki, Greece.

Funding

The present study was funded in the context of the project ‘Molecular signatures analysis of three-dimensional cell cultures and circulating tumor cells in the treatment of cancer’ (MIS 5004622) under the call for proposals ‘Supporting researchers with emphasis on new researchers’ (EDULLL 34). The project was co-financed by Greece and the European Union (European Social Fund-ESF) by the Operational Programme Human Resources Development, Education and Lifelong Learning 2014-2020.

Availability of data and materials

All data generated or analyzed during this study are included in this published article or are available from the corresponding author on reasonable request.

Authors' contributions

KAK was involved in the acquisition of data, analysis and interpretation of data, drafting the article and providing final

approval. MGA was involved in the interpretation of the data, cell culture experiments and providing final approval. LPNG was involved in the interpretation of the data, drafting the article and providing final approval. NGG was involved in the interpretation of the data, revising the article and providing final approval. ISV conceived and designed the study, drafting the article, critical revision and provided final approval.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

NGG declares ownership of Biogenea Pharmaceuticals Ltd. The company had no role in the design or outcomes of the study. The remaining authors declare that they have no competing interests.

References

- Bocci F, Gearhart-Serna L, Boareto M, Ribeiro M, Ben-Jacob E, Devi GR, Levine H, Onuchic JN and Jolly MK: Toward understanding cancer stem cell heterogeneity in the tumor microenvironment. *Proc Natl Acad Sci USA* 116: 148-157, 2019.
- Vizirianakis IS, Miliotou AN, Mystridis GA, Andriotis EG, Andreadis II, Papadopoulou LC and Fatouros DG: Tackling pharmacological response heterogeneity by PBPK modeling to advance precision medicine productivity of nanotechnology and genomics therapeutics. *Expert Rev Precis Med Drug Dev* 4: 139-151, 2019.
- Brabletz T, Kalluri R, Nieto MA and Weinberg RA: EMT in cancer. *Nat Rev Cancer* 18: 128-134, 2018.
- Roche J: The epithelial-to-mesenchymal transition in cancer. *Cancers (Basel)* 10: 52, 2018.
- Culig Z: Epithelial mesenchymal transition and resistance in endocrine-related cancers. *Biochim Biophys Acta Mol Cell Res* 1866: 1368-1375, 2019.
- Lu W and Kang Y: Epithelial-mesenchymal plasticity in cancer progression and metastasis. *Dev Cell* 49: 361-374, 2019.
- Williams ED, Gao D, Redfern A and Thompson EW: Controversies around epithelial-mesenchymal plasticity in cancer metastasis. *Nat Rev Cancer* 19: 716-732, 2019.
- Kyrodinou M, Andreadis D, Drougou A, Amanatiadou EP, Angelis L, Barbatis C, Epivatianos A and Vizirianakis IS: Desmoglein-3/ γ -catenin and E-cadherin/ β -catenin differential expression in oral leukoplakia and squamous cell carcinoma. *Clin Oral Invest* 18: 199-210, 2014.
- Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C and Stuart JM: The cancer genome atlas pan-cancer analysis project. *Nat Genet* 45: 1113-1120, 2013.
- Zhao M, Kong L, Liu Y and Qu H: dbEMT: An epithelial-mesenchymal transition associated gene resource. *Sci Rep* 5: 11459, 2015.
- Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, *et al*: TCGAAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* 44: e71, 2016.
- Anders S, Pyl PT and Huber W: HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31: 166-169, 2015.
- Love MI, Huber W and Anders S: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15: 550, 2014.
- Shaffer JP: Multiple hypothesis testing. *Annu Rev Psychol* 46: 561-584, 1995.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK: limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43: e47, 2015.
- Therneau TM and Grambsch PM: Modeling survival data: Extending the Cox model. Springer-Verlag New York, 2000.
- Therneau T: A package for survival analysis in R. R package version 3.2-7, 2020.
- Kassambara A, Kosinski M and Biecek P: survminer: Drawing survival curves using 'ggplot2'. 2017.
- Tseligka ED, Rova A, Amanatiadou EP, Calabrese G, Tsiouklis J, Fatouros DG and Vizirianakis IS: Pharmacological development of target-specific delocalized lipophilic cation-functionalized carboranes for cancer therapy. *Pharm Res* 33: 1945-1958, 2016.
- Akrivou MG, Demertzidou VP, Theodoroula NF, Chatzopoulou FM, Kyritsis KA, Grigoriadis N, Zografos AL and Vizirianakis IS: Uncovering the pharmacological response of novel sesquiterpene derivatives that differentially alter gene expression and modulate the cell cycle in cancer cells. *Int J Oncol* 53: 2167-2179, 2018.
- Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25: 402-408, 2001.
- Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, Benfante R, Arif M, Liu Z, Edfors F, *et al*: A pathology atlas of the human cancer transcriptome. *Science* 357: eaan2507, 2017.
- Liang H, Fu Z, Jiang X, Wang N, Wang F, Wang X, Zhang S, Wang Y, Yan X, Guan WX, *et al*: miR-16 promotes the apoptosis of human cancer cells by targeting FEAT. *BMC Cancer* 15: 448, 2015.
- Qu Y, Liu H, Lv X, Liu Y, Wang X, Zhang M, Zhang X, Li Y, Lou Q, Li S and Li H: MicroRNA-16-5p overexpression suppresses proliferation and invasion as well as triggers apoptosis by targeting VEGFA expression in breast carcinoma. *Oncotarget* 8: 72400-72410, 2017.
- Ruan L and Qian X: MiR-16-5p inhibits breast cancer by reducing AKT3 to restrain NF- κ B pathway. *Biosci Rep* 39: BSR20191611, 2019.
- Nilsson S, Möller C, Jirstrom K, Lee A, Busch S, Lamb R and Landberg G: Downregulation of miR-92a is associated with aggressive breast cancer features and increased tumour macrophage infiltration. *PLoS One* 7: e36051, 2012.
- Zearo S, Kim E, Zhu Y, Zhao JT, Sidhu SB, Robinson BG and Soon PS: MicroRNA-484 is more highly expressed in serum of early breast cancer patients compared to healthy volunteers. *BMC Cancer* 14: 200, 2014.
- Volinia S and Croce CM: Prognostic microRNA/mRNA signature from the integrated analysis of patients with invasive breast cancer. *Proc Natl Acad Sci USA* 110: 7413-7417, 2013.
- Ye FG, Song CG, Cao ZG, Xia C, Chen DN, Chen L, Li S, Qiao F, Ling H, Yao L, *et al*: Cytidine deaminase axis modulated by miR-484 differentially regulates cell proliferation and chemoresistance in breast cancer. *Cancer Res* 75: 1504-1515, 2015.
- Thompson EW, Reich R, Shima TB, Albini A, Graf J, Martin GR, Dickson RB and Lippman ME: Differential regulation of growth and invasiveness of MCF-7 breast cancer cells by antiestrogens. *Cancer Res* 48: 6764-6768, 1988.
- Gjerdum C, Tiron C, Høiby T, Stefansson I, Haugen H, Sandal T, Collett K, Li S, McCormack E, Gjertsen BT, *et al*: Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proc Natl Acad Sci USA* 107: 1124-1129, 2010.
- Theodossiou TA, Ali M, Grigalavicius M, Grallert B, Dillard P, Schink KO, Olsen CE, Wälchli S, Inderberg EM, Kubin A, *et al*: Simultaneous defeat of MCF7 and MDA-MB-231 resistances by a hypericin PDT-tamoxifen hybrid therapy. *NPJ Breast Cancer* 5: 13, 2019.
- Sha L, Dong L, Lv L, Bai L and Ji X: HOXB9 promotes epithelial-to-mesenchymal transition via transforming growth factor- β 1 pathway in hepatocellular carcinoma cells. *Clin Exp Med* 15: 55-64, 2015.
- Bhatlekar S, Fields JZ and Boman BM: Role of HOX genes in stem cell differentiation and cancer. *Stem Cells Int* 2018: 3569493, 2018.
- Hayashida T, Takahashi F, Chiba N, Brachtel E, Takahashi M, Godin-Heymann N, Gross KW, Vivanco Md, Wijendran V, Shioda T, *et al*: HOXB9, a gene overexpressed in breast cancer, promotes tumorigenicity and lung metastasis. *Proc Natl Acad Sci USA* 107: 1100-1105, 2010.

36. De Braekeleer E, Douet-Guilbert N, Basinko A, Le Bris MJ, Morel F and De Braekeleer M: Hox gene dysregulation in acute myeloid leukemia. *Futur Oncol* 10: 475-495, 2014.
37. Eoh KJ, Kim HJ, Lee JY, Nam EJ, Kim S, Kim SW and Kim YT: Dysregulated expression of homeobox family genes may influence survival outcomes of patients with epithelial ovarian cancer: Analysis of data from the cancer genome atlas. *Oncotarget* 8: 70579-70585, 2017.
38. Song J, Wang T, Xu W, Wang P, Wan J, Wang Y, Zhan J and Zhang H: HOXB9 acetylation at K27 is responsible for its suppression of colon cancer progression. *Cancer Lett* 426: 63-72, 2018.
39. Zhan J, Niu M, Wang P, Zhu X, Li S, Song J, He H, Wang Y, Xue L, Fang W and Zhang H: Elevated HOXB9 expression promotes differentiation and predicts a favourable outcome in colon adenocarcinoma patients. *Br J Cancer* 111: 883-893, 2014.
40. Chiba N, Comaills V, Shiotani B, Takahashi F, Shimada T, Tajima K, Winokur D, Hayashida T, Willers H, Brachtel E, *et al*: Homeobox B9 induces epithelial-to-mesenchymal transition-associated radioresistance by accelerating DNA damage responses. *Proc Natl Acad Sci USA* 109: 2760-2765, 2012.
41. Carbone C, Piro G, Simionato F, Ligorio F, Cremolini C, Loupakis F, Ali G, Rossini D, Merz V, Santoro R, *et al*: Homeobox B9 mediates resistance to anti-VEGF therapy in colorectal cancer patients. *Clin Cancer Res* 23: 4312-4322, 2017.
42. Zhang G, Fan E, Yue G, Zhong Q, Shuai Y, Wu M, Feng G, Chen Q and Gou X: Five genes as a novel signature for predicting the prognosis of patients with laryngeal cancer. *J Cell Biochem*, Oct 31, 2019 (Online ahead of print).
43. Xu H, Wu S, Shen X, Wu D, Qin Z, Wang H, Chen X and Sun X: Silencing of HOXB9 suppresses cellular proliferation, angiogenesis, migration and invasion of prostate cancer cells. *J Biosci* 45: 40, 2020.
44. Fu D, He C, Wei J, Zhang Z, Luo Y, Tan H and Ren C: PGK1 is a potential survival biomarker and invasion promoter by regulating the HIF-1 α -mediated epithelial-mesenchymal transition process in breast cancer. *Cell Physiol Biochem* 51: 2434-2444, 2018.
45. He Y, Luo Y, Zhang D, Wang X, Zhang P, Li H, Ejaz S and Liang S: PGK1-mediated cancer progression and drug resistance. *Am J Cancer Res* 9: 2280-2302, 2019.
46. Sambri I, Capasso R, Pucci P, Perna AF and Ingrosso D: The microRNA 15a/16-1 cluster down-regulates protein repair isoaspartyl methyltransferase in hepatoma cells: Implications for apoptosis regulation. *J Biol Chem* 286: 43690-43700, 2011.
47. Dong L, Li Y, Xue D and Liu Y: PCMT1 is an unfavorable predictor and functions as an oncogene in bladder cancer. *IUBMB Life* 70: 291-299, 2018.
48. Schosserer M, Minois N, Angerer TB, Amring M, Dellago H, Harreither E, Calle-Perez A, Pircher A, Gerstl MP, Pfeifenberger S, *et al*: Methylation of ribosomal RNA by NSUN5 is a conserved mechanism modulating organismal lifespan. *Nat Commun* 6: 6158, 2015.
49. Heissenberger C, Liendl L, Nagelreiter F, Gonskikh Y, Yang G, Stelzer EM, Krammer TL, Micutkova L, Vogt S, Kreil DP, *et al*: Loss of the ribosomal RNA methyltransferase NSUN5 impairs global protein synthesis and normal growth. *Nucleic Acids Res* 47: 11807-11825, 2019.
50. de Melo IS, Iglesias C, Benítez-Rondán A, Medina F, Martínez-Barberá JP and Bolívar J: NOA36/ZNF330 is a conserved cysteine-rich protein with proapoptotic activity in human cells. *Biochim Biophys Acta* 1793: 1876-1885, 2009.
51. Willis S, Sun Y, Abramovitz M, Fei T, Young B, Lin X, Ni M, Achua J, Regan MM, Gray KP, *et al*: High expression of FGD3, a putative regulator of cell morphology and motility, is prognostic of favorable outcome in multiple cancers. *JCO Precis Oncol* 1: PO.17.00009, 2017.
52. Renda I, Bianchi S, Vezzosi V, Nori J, Vanzi E, Tavella K and Susini T: Expression of FGD3 gene as prognostic factor in young breast cancer patients. *Sci Rep* 9: 15204, 2019.
53. Vidigal JA and Ventura A: The biological functions of miRNAs: Lessons from in vivo studies. *Trends Cell Biol* 25: 137-147, 2015.
54. Peng Y and Croce CM: The role of microRNAs in human cancer. *Signal Transduct Target Ther* 1: 15004, 2016.
55. Bartel DP: Metazoan MicroRNAs. *Cell* 173: 20-51, 2018.
56. Michael IP, Saghafeinia S and Hanahan D: A set of microRNAs coordinately controls tumorigenesis, invasion, and metastasis. *Proc Natl Acad Sci USA* 116: 24184-24195, 2019.
57. Cherone JM, Jorgji V and Burge CB: Cotargeting among microRNAs in the brain. *Genome Res* 29: 1791-1804, 2019.
58. Dong J, Xiao D, Zhao Z, Ren P, Li C, Hu Y, Shi J, Su H, Wang L, Liu H, *et al*: Epigenetic silencing of microRNA-137 enhances ASCT2 expression and tumor glutamine metabolism. *Oncogenesis* 6: e356, 2017.
59. Zhang Z, Liu R, Shuai Y, Huang Y, Jin R, Wang X and Luo J: ASCT2 (SLC1A5)-dependent glutamine uptake is involved in the progression of head and neck squamous cell carcinoma. *Br J Cancer* 122: 82-93, 2020.
60. Wu P, Xiao Y, Guo T, Wang Y, Liao S, Chen L and Liu Z: Identifying miRNA-mRNA pairs and novel miRNAs from hepatocellular carcinoma mirnomes and TCGA database. *J Cancer* 10: 2552-2559, 2019.
61. Nuzzo S, Catuogno S, Capuozzo M, Fiorelli A, Swiderski P, Boccella S, de Nigris F and Esposito CL: Axl-targeted delivery of the oncosuppressor miR-137 in non-small-cell lung cancer. *Mol Ther Nucleic Acids* 17: 256-263, 2019.
62. Drago-García D, Espinal-Enríquez J and Hernández-Lemus E: Network analysis of EMT and MET micro-RNA regulation in breast cancer. *Sci Rep* 7: 13534, 2017.
63. Schneider E, Pochert N, Ruess C, MacPhee L, Escano L, Miller C, Krowiorz K, Delsing Malmberg E, Heravi-Moussavi A, Lorzadeh A, *et al*: MicroRNA-708 is a novel regulator of the Hoxa9 program in myeloid cells. *Leukemia* 34: 1253-1265, 2020.
64. Looney AM, Nawaz K and Webster RM: Tumor-agnostic therapies. *Nat Rev Drug Discov* 19: 383-384, 2020.
65. Astras G, Papagiannopoulos CI, Kyritsis KA, Markitani C and Vizirianakis IS: Pharmacogenomic testing to guide personalized cancer medicine decisions in private oncology practice: A case study. *Front. Oncol* 10: 521, 2020.
66. Gu Z, Eils R and Schlesner M: Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 32: 2847-2849, 2016.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.