RESEARCH ARTICLE

Glypican 6 is a putative biomarker for metastatic progression of cutaneous melanoma

Yuanyuan Li[1☺], Melissa Li[1☺], Igor Shats[2], Juno M. Krahn[3], Gordon P. Flake[4†], David M. Umbach[1], Xiaoling Li[2], Leping Li[1]*

1 Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, Durham, North Carolina, United States of America, **2** Laboratory of Signal Transduction, National Institute of Environmental Health Sciences, Durham, North Carolina, United States of America, **3** Genome Integrity & Structural Biology Laboratory, National Institute of Environmental Health Sciences, Durham, North Carolina, United States of America, **4** Cellular and Molecular Pathology Branch, National Institute of Environmental Health Sciences, Durham, North Carolina, United States of America

☺ These authors contributed equally to this work.
† Deceased.
* li3@niehs.nih.gov

Abstract

Due to the poor prognosis of advanced metastatic melanoma, it is crucial to find early biomarkers that help identify which melanomas will metastasize. By comparing the gene expression data from primary and cutaneous melanoma samples from The Cancer Genome Atlas (TCGA), we identified GPC6 among a set of genes whose expression levels can distinguish between primary melanoma and regional cutaneous/subcutaneous metastases. Glypicans are thought to play a role in tumor growth by regulating the signaling pathways of Wnt, Hedgehogs, fibroblast growth factors (FGFs), and bone morphogenetic proteins (BMPs). We showed that GPC6 expression was up-regulated in a melanoma cell line compared to normal melanocytes and in metastatic melanoma compared to primary melanoma. Furthermore, GPC6 expression was positively correlated with genes largely involved in cell adhesion and migration in both melanoma samples and in RNA-seq samples from other TCGA tumors. Our results suggest that GPC6 may play a role in tumor metastatic progression. In TCGA melanoma samples, we also showed that GPC6 expression was negatively correlated with miR-509-3p, which has previously been shown to function as a tumor suppressor in various cancer cell lines. We overexpressed miR-509-3p in A375 melanoma cells and showed that GPC6 expression was significantly suppressed. This result suggested that GPC6 was a putative target of miR-509-3p in melanoma. Together, our findings identified GPC6 as an early biomarker for melanoma metastatic progression, one that can be regulated by miR-509-3p.

Introduction

The Cancer Genome Atlas (TCGA) project has generated a large amount of data using several platforms, including RNA-seq, applied to the same tissue specimens for a variety of tumors

like melanoma [1]. These data provide unprecedented information about the molecular map of tumors. Previously, using TCGA data, we showed that primary melanomas are heterogeneous at the gene expression level, and that the degree of loss of epithelium-characteristic gene expression in those melanomas is correlated with predicted metastatic progression [2].

Despite its relatively low incidence rate, skin cutaneous melanoma (SKCM) is the deadliest type of skin cancer due to its invasiveness. For patients with stage IV melanoma with a metastasis spread to the lymph nodes or other organs, the median survival is 8–9 months [3]. The development of BRAF inhibitors, such as vemurafenib and dabrafenib [4], have dramatically altered the paradigm of melanoma treatment and improved patient survival. Unfortunately, many patients eventually develop resistance to these drugs. To address this challenge, new targeted drugs are being developed and immunotherapy has shown great promise (for a recent review, see [5]).

Although it can be deadly, melanoma is often curable when diagnosed early. Thus, early detection and intervention can be lifesaving and markers for melanoma progression would be informative. Herein, we analyzed the RNA-seq gene expression data from TCGA for 74 primary melanomas and 66 melanomas that had metastasized to regional cutaneous/subcutaneous tissues (including satellite and in-transit metastases). We aimed to identify genes whose expression levels distinguish primary melanoma from cutaneous/subcutaneous melanoma, hoping to identify biomarkers that are indicative of early melanoma metastatic progression. In this report, we focused on a new putative melanoma gene, glypican 6 (GPC6).

Glypicans are a family of heparan sulfate proteoglycans that are linked to the extracellular cell surface of the plasma membrane. They are thought to regulate the signaling of Wnt, Hedgehogs, fibroblast growth factors (FGFs), and bone morphogenetic proteins (BMPs) [6–12]. Six glypicans have been identified in mammals (GPC1 to GPC6) and two in Drosophila, where they play important roles in development (reviewed in [11–13]). GPC6 is the newest member of the family, is most homologous to GPC4, and is ubiquitously expressed [14].

GPC6 has also been implicated in many cancers and other diseases. Exome sequencing identified GPC6 among genes that were recurrently mutated across individual prostate tumors from different patients [15]. In a large-scale genome-wide association study (GWAS), Amankwah et al. identified a single-nucleotide polymorphism (SNP) (rs17702471) in GPC6/GPC5 that was associated with an increased risk of invasive epithelial ovarian carcinoma [16]. Loftus et al. [17] demonstrated that HIF1α signaling and hypoxia in melanocytes directly regulate the expression of GPC6 and that increased expression of GPC6 was positively associated with poor survival in melanoma. GPC6 was among a group of genes whose expression levels were significantly correlated with reduced time of disease-free status in melanoma [17]. Taken together, these studies demonstrated that GPC6 functions in tumor progression.

Using RNA-seq gene expression data from TCGA, we identified GPC6 among a set of genes whose expression levels can distinguish primary skin melanoma from melanoma metastasized to cutaneous/subcutaneous tissue. In this report, we used both computational and experimental approaches to try to understand the role of GPC6 in the progression of melanoma.

Materials and methods

Data

The Cancer Genome Atlas (TCGA) collected melanoma tissue samples from one of the four tissue sites: skin (primary tumor), and metastases to regional cutaneous/subcutaneous tissues, lymph nodes, or more distant sites (the latter three referred to collectively as metastatic tumors). For simplicity, we refer to the four categories of melanoma samples simply as:

primary, cutaneous, lymph, and distant metastases with 74, 66, 195, and 54 RNA-seq samples (389 total), respectively. We downloaded (May 2015) UNC RNASeqV2 and SKCM level 3 expression data from the TCGA data portal (https://portal.gdc.cancer.gov/). We $\log_2$-transformed the normalized read counts (per million reads mapped) for RNA-seq data (all values less than 1 were assigned value 1 before transformation), but we carried out no further normalization. We also downloaded RNA-seq data for 1,156 cancer cell lines from the Cancer Cell Line Encyclopedia (CCLE) (https://portals.broadinstitute.org/ccle).

Computational method

In preliminary work, we concluded that it is not feasible to identify a set of genes whose expression levels can correctly classify all four classes of melanoma samples (data not shown). Thus, we decided to focus on distinguishing primary melanoma (74 samples) from cutaneous/ subcutaneous (66 samples) melanoma. We randomly divided the data into a training (75% of the samples) and a testing set (25% of the samples). We used the training set to train the classification algorithm (GA/KNN) [18, 19] and evaluated training performance through a leave-one-out cross-validation procedure. The prediction accuracies may vary depending on which samples are assigned to the training and testing sets. To avoid idiosyncrasies from use of a single random assignment, we carried out 100 independent random training and testing assignments and obtained training-set and testing-set prediction accuracies from each. On average, each sample was placed 75 times in the training set and 25 times in the testing set and predicted the corresponding number of times. We used the average prediction accuracies from the ~75 training predictions and from the ~25 testing predictions for each sample as the overall prediction accuracies for that sample in training and testing sets, respectively. In our analysis, for a given training/testing partition, we collected 5,000 near-optimal feature sets, resulting in 5,000 classifications of the training- and testing-set samples.

The details of the GA/KNN algorithm can be found in [18, 19]. The chromosome length was set to 20 (a 20-gene set) and the population size was set to 300. The maximal number of generations of "genetic evolution" was set to be 1000. For the k-nearest neighbor (KNN) classification, k was set to 5 with a majority "voting" rule.

To identify putative miR-509-3p binding sites in the 3'-untranslated region (UTR) of human GPC6 gene, we downloaded the 3'-UTR sequence of human GPC6 (NM_005708) from the UCSC genome browser (build hg38) (https://genome.ucsc.edu/). We used a custom C code to search for segments in the GPC6 sequence that are complementary to the 3'- to 5'- seed sequence (seed length = 8) of human miR-509-3p. A putative target site was declared when seven or more of the eight nucleotides were complementary (one G:U pairing was allowed).

Experimental method

A375 melanoma cells (150,000)/well were reverse transfected in 6-well plates with 4nM negative control mimic (catalog # 4464077, Thermo Fisher) or miR-509-3p mimic (assay ID MC12984, Thermo Fisher) with Lipofectamine RNAiMax reagent (catalog # 13778030, Thermo Fisher) according to the manufacturer's protocol. After 48 hours of transfection, cells were collected in TRIzol (catalog # 15596018, Thermo Fisher) and total RNA was extracted according to the manufacturer's protocol. cDNA was synthesized with the High-Capacity cDNA Reverse Transcription Kit (catalog # 4368814, Thermo Fisher) according to the manufacturer's protocol. qPCR was performed using iQ SYBR Green Supermix (catalog # 1708880, Bio-Rad) on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad) with Taqman Assays for miR-509-3p (assay ID: 002236, Thermo Fisher), GPC6 (assay ID: Hs00170677_m1,

Thermo Fisher), and Rplp0 (assay ID: HS99999902_m1, Thermo Fisher). Relative quantitation of target transcript expression was calculated using the ddCT method using Rplp0 as the endogenous control.

Results

Classification accuracies

The mean and median prediction accuracies were 93.5% and 97.8% for primary and cutaneous samples in the training sets, and 77.5% and 83.0% for samples in the testing sets (Table 1). The prediction accuracies for the training set were higher than those for the testing set, suggesting some training bias existed. Nonetheless, 80% of the time, the gene signatures (sets of 20 genes) selected by the GA/KNN algorithm could correctly assign an RNA-seq sample as either a primary melanoma or a metastasis to a regional cutaneous/subcutaneous tissue.

Top genes whose expression levels distinguish between primary and cutaneous melanoma

The ten most frequently selected genes from all 100 training/testing partitions were FABP4, SFRP4, CILP, SLITRK4, EBF2, PRG4, KRT6B, GPC6, KRT17, and OGN. All genes except the two keratin genes (KRT6B and KRT17) were up-regulated in cutaneous melanoma compared to primary melanoma (S1 Fig). Gene ontology (GO) analysis [20] showed that the top 300 genes were significantly enriched in GO terms: signal transduction, cell communication, cell structure and motility, and others (Table 2). Some of the other notable genes include genes in cell adhesion (ADAMTS6, ADAMTS9, ADAMTS12, FAT2, FAT3, COL17A1, DSC2, LAMC2, LARRC15, NTN1, ODZ2, PVRL4, PCDH17, SEMA3G, SCUBE3, and SLIT2), epithelial-mesenchymal transition (EMT) (ANGPTL1 and HMGA2), Hedgehog signaling (BMP8B, DHH, WNT4, WNT5A, WNT7B, and WNT11), MAPK pathway (DUSP4, DUSP5, IL17D, and RAPGEF5), receptor tyrosine kinase pathway (FLT1, PDGFRL, SH2B2 and TIE1), transcription factors (GLIS1), and Wnt signaling pathway (APCDD1, WNT4, WNT5A, WNT7B, and WNT11).

Among the top ten genes whose expression levels distinguish primary from metastatic melanoma, we focused our subsequent analyses on a putative novel biomaker, GPC6, whose expression level was up-regulated in metastatic melanoma compared to primary melanoma (Fig 1A).

Up-regulation of GPC6 expression in melanoma

Using public array gene expression data from the Encyclopedia of DNA Elements (ENCODE), we found that GPC6 was also up-regulated in a melanoma cell line (mel_2183) compared to

Table 1. Summary statistics of classification performances.

Melanoma tissue site	Min.	1st Quartile	Median	Mean	3rd Quartile	Max.
			Training set performance			
primary	0.428	0.935	0.984	0.934	0.996	1.000
cutaneous	0.420	0.943	0.975	0.936	0.990	0.999
Overall	0.420	0.937	0.978	0.935	0.994	1.000
			Testing set performance			
primary	0.124	0.688	0.862	0.781	0.962	0.997
cutaneous	0.144	0.689	0.819	0.768	0.912	0.993
Overall	0.124	0.686	0.830	0.775	0.943	0.997

Table 2. Significant GO terms for the top 300 genes that distinguish primary from cutaneous melanoma.

Biological process	Number of genes in GO term	Multiple-testing-adjusted p-value
Signal transduction	80	0.0040
Cell communication	38	0.0021
Developmental processes	56	0.0022
Cell structure and motility	33	0.013
Mesoderm development	21	0.011
Cell adhesion-mediated signaling	16	0.018
Other neuronal activity	9	0.029
Muscle development	9	0.029
Cell structure	21	0.045

https://doi.org/10.1371/journal.pone.0218067.t002

normal melanocytes (GSE15805) (Fig 1B) and is overexpressed in metastatic melanoma compared to primary melanoma (Fig 1A). Next, we investigated a possible mechanism for GPC6 overexpression in melanoma.

GPC6 is a putative target of miR-509-3p

There are three microRNA-509 genes (miR-509-1, miR-509-2, and miR-509-3) in TCGA samples that all produce two mature forms of microRNAs (miR-509-5p and miR-509-3p). The expression levels of the three microRNAs were highly correlated (r = 0.997 − 0.999, Spearman correlation). Because of their high correlations, we used miR-509-1 as the representative of the three. The expression level of miR-509-1 in normal human melanocytes is not known. However, we found that miR-509-1 had higher expression level in melanoma distant metastases than in primary melanoma in TCGA samples (p = 3.53E-05, t test, two-sided) (Fig 2A). miR-509-1 expression was negatively correlated with that of GPC6 in TCGA melanoma samples (r = -0.41, Spearman correlation) (Fig 2B). Previously, Pan et al. [21] showed that GPC6 was among a few genes that were down-regulated in two ovarian cancer cell lines (OVCAR8 and HEYA8) transfected with miR-509-3p mimics relative to untreated controls. We identified two putative miR-509-3p binding sites in the 3'-UTR of human GPC6 gene (Fig 3A) and hypothesized that GPC6 is a putative target of miR-509-3a in melanoma cells. To test this hypothesis, we transfected A375 melanoma cells with negative control mimic or miR-509-3p mimic. After 48 hours, we collected total RNA and measured levels of miR-509-3p and GPC6. We found that miR-509-3p levels were significantly overexpressed with miR-509-3p mimic
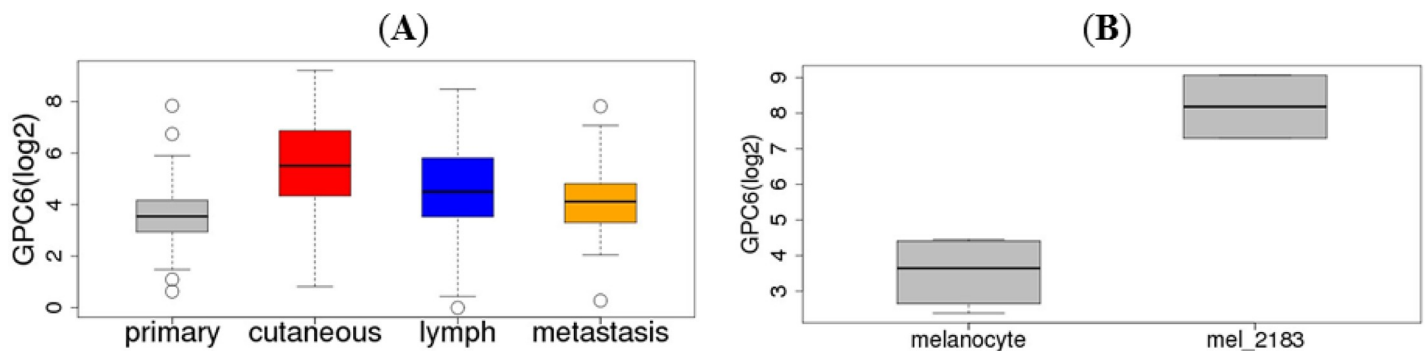

Fig 1. GPC6 expression in normal and melanoma cell lines and TCGA melanoma tumors. Data from the two platforms were not normalized for direct comparison. (A) GPC6 expression in TCGA melanoma samples from four tissue sites (RNA-seq from TCGA). (B) Boxplots of the expression of GPC6 in normal melanocytes (four replicates) and melanoma cell line (mel-2183, two replicates) measured by Affymetrix arrays (GSE15805).
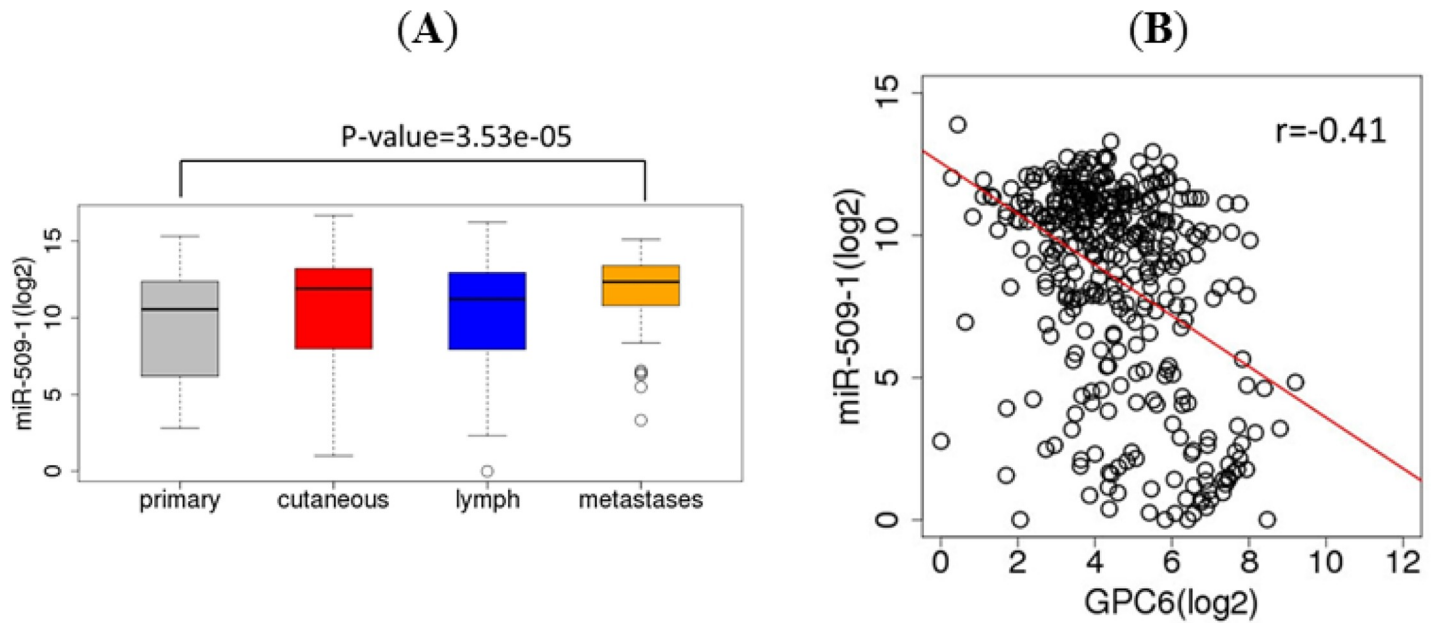
https://doi.org/10.1371/journal.pone.0218067.g001

(A)

P-value=3.53e-05

(B)

r=-0.41

Fig 2. miR-509-1 expression and correlation with GPC6. (A) Boxplots of miR-509-1 expression levels in TCGA melanoma samples for the four tissue sites. (B) Scatter plot of miR-509-1 and GPC6 expression levels in TCGA melanoma samples (N = 389). The red line indicates the regression line.

(Fig 3B), as expected. Furthermore, with miR-509-3p overexpression, GPC6 levels were significantly reduced (by 83–88%) (Fig 3C). GPC6 expression was up-regulated in TCGA metastatic melanoma samples compared to normal samples, however, its expression levels showed a metastasis stage-dependent decrease (Fig 1A) with an overall lowest expression in tumors metastasized to distant organs. Conversely miR-509-3p had the highest expression in distant
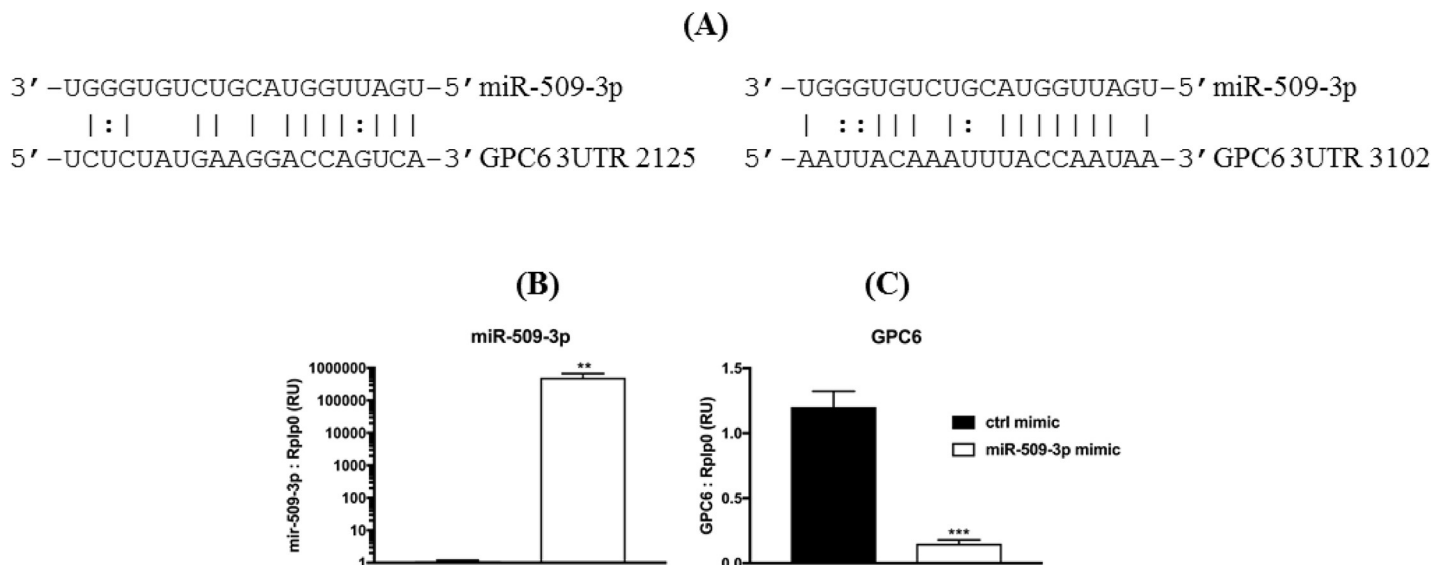
(A)

```
3'-UGGGUGUCUGCAUGGUUAGU-5' miR-509-3p
    |:|      ||  |  ||||:|||
5'-UCUCUAUGAAGGACCAGUCA-3' GPC6 3UTR 2125
```

```
3'-UGGGUGUCUGCAUGGUUAGU-5' miR-509-3p
    |  ::|||  |:  |||||||  |
5'-AAUUACAAAUUUACCAAUAA-3' GPC6 3UTR 3102
```

(B)

miR-509-3p

(C)

GPC6

Fig 3. GPC6 is a putative target of miR-509-3p and overexpression of miR-509-3p downregulates GPC6 mRNA levels. (A) Predicted binding sites of miR-509-3p on the 3'-UTR of human GPC6 gene (NM005708). (B-C) A375 cells were transfected with 4nM negative control mimic (mimic ctrl) or miR-509-3p mimic for 48 hours. RNA was collected and RNA levels of miR-509-3p (B) and GPC6 (C) were measured using Taqman assays relative to Rplp0 from n = 3 experiments. Means ± S.D. **, $p < 0.001$; ***, $p < 0.0001$ for miR-509-3p mimic versus mimic control.

metastases compared to primary melanoma tumors (Fig 2A). Those data are consistent with the notation that miR-509-3p may suppress GPC6 expression.

Genes coordinately expressed with GPC6 are largely involved in cell adhesion

Analysis of all 389 melanoma samples revealed 82 genes whose expression levels were strongly positively correlated (r ≥ 0.5, Spearman correlation), and seven that were strongly negatively correlated (r ≤ -0.5, Spearman correlation) with GPC6. Among the 82 genes, 14 (EDIL3, CHD11, COL3A1, COL8A1, COL12A1, ECM2, LAMA2, NRP1, NTM, POSTN, PCDH17, ROR2, TNFAIP6, and VCAN) are known to be involved in cell adhesion.

ZEB1 exhibited the highest correlation with that of GPC6 (r = 0.63, Spearman correlation) when all 389 melanoma samples were combined. It was also among the ten most positively correlated genes when samples from each of the four categories were analyzed separately (Table 3 and Fig 4A). ZEB1, a zinc finger transcription factor, is an important regulator of the EMT by suppressing the expression of cadherin along with Snail-related zinc-finger transcriptional repressors (SNAIL and SLUG) and other factors such as TWIST [22]. ZEB1 expression was higher in metastatic melanoma (cutaneous/subcutaneous, lymph node, or distant metastases) compared to primary melanoma (Fig 4B). The correlation in expression between GPC6 and ZEB1 is evident within each of the four categories of melanoma. This commonality of correlation within separate categories of melanoma is also true for most of the remaining 82 genes (data not shown).

We also computed the pair-wise correlation in gene expression between GPC6 and all other genes in each of the 32 TCGA tumor types (9,511 samples) (S1 Table). The overall rank of correlation between each gene and GPC6 across all 32 TCGA tumors was computed using a robust ranking method [23]. The top 200 genes with the highest positive correlation are listed in S2 Table. Fifty (S3 Table) of the top 200 genes are highly significantly associated with cell adhesion (p = 1.7E-22) (Table 4). Genes associated with angiogenesis (THY1, CXCL12, COL15A1, CTGF, ENDRA, ENPEP, MMP14, NRP1, ROBO1, SLIT2, and THBS1) and protein kinase signaling (HTR2A, F2R, EDNRA, HGF, ITGA1, LRP1, PDGFC, THBS1, TGFB3, and

Table 3. Spearman correlation coefficients between GPC6 and the top 10 positively and 5 negatively correlated genes in melanoma RNA-seq samples.

Gene	Melanoma tissue site				Average
	Primary	Cutaneous	Lymph node	Distant metastases	
VCAN	0.57	0.54	0.54	0.68	0.58
SRPX2	0.42	0.70	0.61	0.57	0.58
INHBA	0.51	0.42	0.67	0.71	0.58
TNFAIP6	0.41	0.64	0.65	0.60	0.57
PPAP2B	0.48	0.61	0.55	0.65	0.57
EDIL3	0.43	0.62	0.63	0.56	0.56
PDGFC	0.58	0.49	0.53	0.64	0.56
TCF4	0.50	0.56	0.54	0.63	0.56
COL8A1	0.51	0.48	0.57	0.67	0.56
ZEB1	0.40	0.56	0.65	0.57	0.55
PMEL	-0.36	-0.65	-0.51	-0.47	-0.50
RPUSD3	-0.46	-0.53	-0.46	-0.57	-0.50
UCK1	-0.53	-0.51	-0.43	-0.54	-0.50
TRIM63	-0.35	-0.64	-0.51	-0.57	-0.52
TSPAN10	-0.45	-0.74	-0.49	-0.50	-0.54

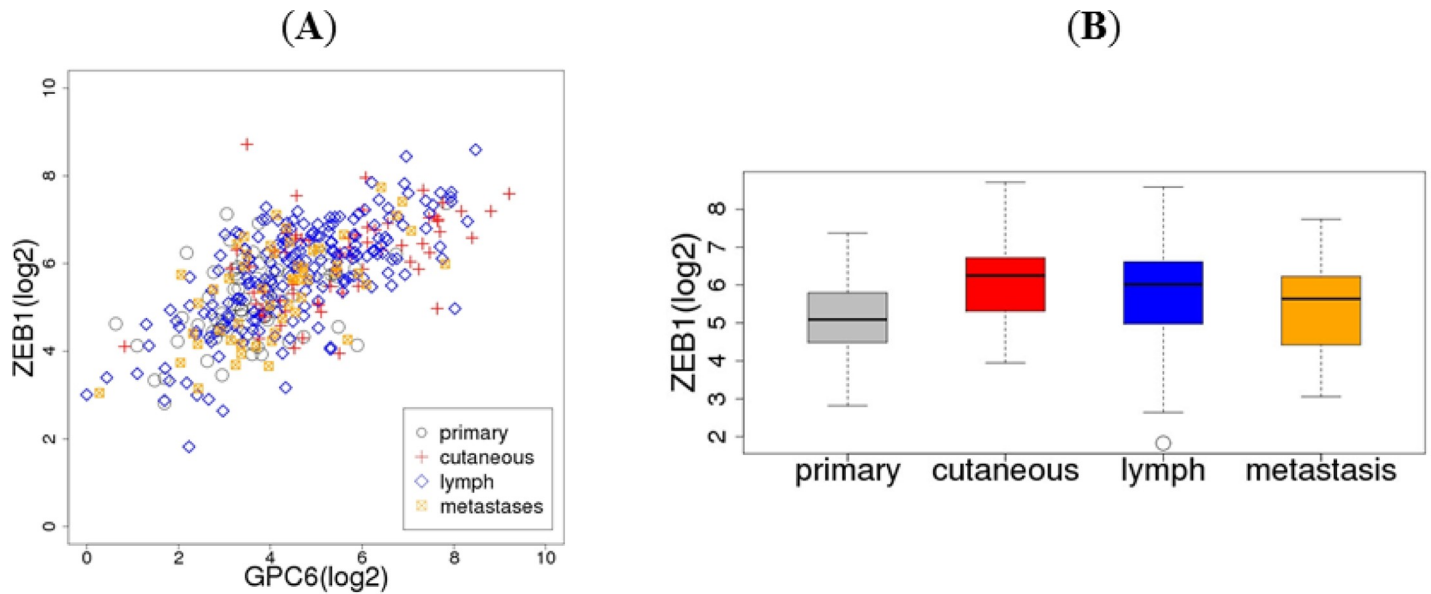https://doi.org/10.1371/journal.pone.0218067.t003

Fig 4. Correlation between GPC6 and ZEB1 expression. (A) GPC6 expression is correlated with ZEB1 in TCGA cutaneous skin melanoma samples. (B) Boxplots of RNA-seq expression levels of ZEB1 in TCGA cutaneous melanoma samples from primary site (skin), and metastases to cutaneous/subcutaneous tissue, lymph node, and distant sites.

https://doi.org/10.1371/journal.pone.0218067.g004

ZEB2) were also significantly enriched (p = 3.3E-04 and 2.3E-02, respectively). Key EMT transcription factors were ranked among the top 500 (ZEB1 ranked 88th; ZEB2, 148th; SNAI1, 422nd; SNAI2, 224th; and TWIST1, 385th).

Similarly, we computed the pair-wise Spearman correlation between the expression levels of GPC6 and those of all other genes in the 1,156 CCLE RNA-seq samples [24]. GO analysis [20] suggested that the top 200 most positively correlated genes were enriched with terms such as development and cell adhesion (S4 Table). Forty-five of the top 200 genes were also among the top 200 identified from the TCGA pan-cancer tumor samples.

Discussion

Tumor metastasis is the main driver of cancer-related death. It is a complex process that is thought to involve several steps including EMT, invasion, and angiogenesis [25–27]. Many key

Table 4. Significant GO terms (GOTERM_BP_ALL) that are associated with the top 200 genes mostly correlated with GPC6 across all 33 TCGA tumors.

Biological process	Number of genes in GO term	Multiple-testing-adjusted p-value
Cell adhesion	50	1.7E-22
Biological adhesion	50	9.0E-23
Extracellular matrix organization	22	5.1E-18
Extracellular structure organization	24	2.5E-16
Developmental process	86	1.1E-14
Multicellular organismal development	79	4.7E-13
Blood vessel development	24	1.4E-12
Vasculature development	24	2.1E-12
Anatomical structure development	72	2.9E-12
System development	67	2.7E-11

https://doi.org/10.1371/journal.pone.0218067.t004

signaling pathways have been implicated in EMT [22] including those associated with receptor tyrosine kinase [28], the transforming growth factor β (TGFB) superfamily [29], Wnt [30], NOTCH [31], and Hedgehog [32].

GPC6 was up-regulated in a melanoma cell line compared to normal melanocytes. Its expression was also up-regulated in melanoma that had metastasized to cutaneous/subcutaneous tissues, lymph nodes or distant metastases compared to expression in primary melanoma. In melanomas, genes coordinately expressed with GPC6 were largely those involved in cell adhesion and migration including INHBA, PDGFC, PDGFRA, PPAP2B, SPRX2, TCF4, and TNFAIP6 and ZEB1. Pan-cancer analysis using ~9,500 RNA-seq samples from 32 different tumors also confirmed that genes with the highest correlation in expression levels with GPC6 across the 32 tumors were enriched with genes related to cell adhesion and migration. The correlation between GPC6 expression and the expression of key EMT transcription factors (ZEB1/2, TWIST1, and SNAI1/2) among the top 500 (out of ~20,000 genes) suggests that GPC6 may be associated with EMT.

Glypicans are thought to regulate the signaling of Wnt, Hedgehogs, fibroblast growth factors (FGFs) and bone morphogenetic proteins (BMPs) [6–12]. The most studied glypican is GPC3. GPC3 promotes the growth of hepatocellular carcinoma through the Wnt pathway [33, 34] and the activated extracellular signal-regulated kinase (ERK) pathway [35]. GPC3 protein expression is elevated in several neoplastic tissues including melanoma compared to nonneoplastic and normal tissues [36]. Contrarily, down-regulation of GPC3 in ovarian cancer [37] and breast cancer [38] promotes tumor migration and invasion. GPC1, another member of the glypican family, is overexpressed in human breast cancer, and its overexpression may promote tumorigenesis [39]. Although little is known about the role of GPC6 in tumor progression and metastasis in melanoma and other tumors, one would speculate that GPC6 functions also through Wnt, FGF, Hedgehog, and BMP signaling.

To determine a potential mechanism of GPC6 regulation during cancer progression, we hypothesized that GPC6 was targeted by miR-509-3p. It has previously been shown that miR-509-3p was downregulated in gastric cancer, while it was upregulated in gastrointestinal stromal tumors of epithelioid and mixed histological types compared to spindle type [40], metastatic melanoma compared to skin samples from normal healthy donors [41], and in metastatic (CRL-1676 and Sk-mel-28) cells compared to primary (CRL-1675) melanoma cells [42]. It has also previously been shown that miR-509-3p functions in cell growth, proliferation, and migration in gastric cancer [43], renal cell carcinoma [44], and ovarian cancer [21].

Conclusions

GPC6 expression was elevated in melanoma samples compared to normal melanocytes and elevated in melanomas that had metastasized to regional cutaneous/subcutaneous tissue, lymph node, or distant organs compared to primary melanomas. GPC6 expression was positively correlated with expression of many genes that are involved in cell adhesion and migration in melanoma samples as well as in samples from other tumors from TCGA. We showed that overexpression of miR-509-3p mimic in A375 melanoma cells suppressed GPC6 expression. It has previously been shown that miR-509-3p mimics also suppressed expression of GPC6 and other transcripts (e.g., SNAI2 and TWIST) associated with EMT in ovarian cancer cell lines [21]. Taken together, these results suggest that GPC6 may play a role in and serve as a biomarker for tumor metastatic progression and GPC6 expression may be regulated by miR-509-3p.

Supporting information

S1 Fig. Boxplots of the expression of the ten most frequently selected genes. Both RNA-seq and clinical classification data were obtained from TCGA.
(TIF)

S2 Fig. Boxplots of the expression of miR-205 in melanoma samples. Both RNA-seq and clinical classification data were obtained from TCGA.
(TIF)

S1 Table. TCGA tumor samples analyzed by RNA-seq (May 2015).
(DOCX)

S2 Table. Top 200 most correlated genes with GPC6. Each cell lists the Pearson correlation coefficient between the expression level of GPC6 and that of the gene (row) in a TCGA tumor type (column).
(XLSX)

S3 Table. The 50 cell adhesion genes among the top 200 most highly correlated with GPC6 across the 32 TCGA tumors.
(DOCX)

S4 Table. Top 10 significant GO terms (GOTERM_BP_ALL) that are associated with the top 200 genes mostly correlated with GPC6 in 1,156 CCLE samples.
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: Yuanyuan Li, Melissa Li, Igor Shats, Gordon P. Flake, David M. Umbach, Xiaoling Li, Leping Li.

Data curation: Yuanyuan Li, Gordon P. Flake, David M. Umbach, Leping Li.

Formal analysis: Yuanyuan Li, Melissa Li, Gordon P. Flake, David M. Umbach, Leping Li.

Funding acquisition: Leping Li.

Investigation: Melissa Li, Igor Shats, Leping Li.

Methodology: Yuanyuan Li, David M. Umbach, Leping Li.

Resources: Gordon P. Flake, Leping Li.

Software: Yuanyuan Li, Juno M. Krahn, Leping Li.

Supervision: Gordon P. Flake, David M. Umbach, Xiaoling Li, Leping Li.

Validation: Gordon P. Flake, David M. Umbach, Leping Li.

Visualization: Yuanyuan Li, David M. Umbach, Leping Li.

Writing – original draft: Yuanyuan Li, Melissa Li, David M. Umbach, Leping Li.

Writing – review & editing: Yuanyuan Li, Melissa Li, Igor Shats, David M. Umbach, Xiaoling Li, Leping Li.

References

1. Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. Nature genetics. 2013; 45(10):1113–20. https://doi.org/10.1038/ng.2764 PMID: 24071849; PubMed Central PMCID: PMC3919969.

2. Li Y, Krahn JM, Flake GP, Umbach DM, Li L. Toward predicting metastatic progression of melanoma based on gene expression data. Pigment cell & melanoma research. 2015; 28(4):453–63. https://doi.org/10.1111/pcmr.12374 PMID: 25847062; PubMed Central PMCID: PMC4469521.

3. Kaufman HL, Kirkwood JM, Hodi FS, Agarwala S, Amatruda T, Bines SD, et al. The Society for Immunotherapy of Cancer consensus statement on tumour immunotherapy for the treatment of cutaneous melanoma. Nature reviews Clinical oncology. 2013; 10(10):588–98. https://doi.org/10.1038/nrclinonc.2013.153 PMID: 23982524.

4. Lo JA, Fisher DE. The melanoma revolution: from UV carcinogenesis to a new era in therapeutics. Science (New York, NY. 2014; 346(6212):945–9. https://doi.org/10.1126/science.1253735 PMID: 25414302.

5. Luke JJ, Flaherty KT, Ribas A, Long GV. Targeted agents and immunotherapies: optimizing outcomes in melanoma. Nature reviews Clinical oncology. 2017; 14(8):463–82. https://doi.org/10.1038/nrclinonc.2017.43 PMID: 28374786.

6. Ohkawara B, Yamamoto TS, Tada M, Ueno N. Role of glypican 4 in the regulation of convergent extension movements during gastrulation in Xenopus laevis. Development. 2003; 130(10):2129–38. https://doi.org/10.1242/dev.00435 PMID: 12668627.

7. Yan D, Lin X. Drosophila glypican Dally-like acts in FGF-receiving cells to modulate FGF signaling during tracheal morphogenesis. Dev Biol. 2007; 312(1):203–16. https://doi.org/10.1016/j.ydbio.2007.09.015 PMID: 17959166; PubMed Central PMCID: PMC2151973.

8. Song HH, Shi W, Xiang YY, Filmus J. The loss of glypican-3 induces alterations in Wnt signaling. The Journal of biological chemistry. 2005; 280(3):2116–25. https://doi.org/10.1074/jbc.M410090200 PMID: 15537637.

9. Lin X, Perrimon N. Dally cooperates with Drosophila Frizzled 2 to transduce Wingless signalling. Nature. 1999; 400(6741):281–4. https://doi.org/10.1038/22343 PMID: 10421372.

10. Topczewski J, Sepich DS, Myers DC, Walker C, Amores A, Lele Z, et al. The zebrafish glypican knypek controls cell polarity during gastrulation movements of convergent extension. Developmental cell. 2001; 1(2):251–64. PMID: 11702784.

11. Filmus J, Capurro M, Rast J. Glypicans. Genome biology. 2008; 9(5):224. https://doi.org/10.1186/gb-2008-9-5-224 PMID: 18505598; PubMed Central PMCID: PMC2441458.

12. Filmus J, Capurro M. The role of glypicans in Hedgehog signaling. Matrix biology: journal of the International Society for Matrix Biology. 2014; 35:248–52. https://doi.org/10.1016/j.matbio.2013.12.007 PMID: 24412155.

13. De Cat B, David G. Developmental roles of the glypicans. Seminars in cell & developmental biology. 2001; 12(2):117–25. https://doi.org/10.1006/scdb.2000.0240 PMID: 11292377.

14. Veugelers M, De Cat B, Ceulemans H, Bruystens AM, Coomans C, Durr J, et al. Glypican-6, a new member of the glypican family of cell surface heparan sulfate proteoglycans. The Journal of biological chemistry. 1999; 274(38):26968–77. https://doi.org/10.1074/jbc.274.38.26968 PMID: 10480909.

15. Kumar A, White TA, MacKenzie AP, Clegg N, Lee C, Dumpit RF, et al. Exome sequencing identifies a spectrum of mutation frequencies in advanced and lethal prostate cancers. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108(41):17087–92. https://doi.org/10.1073/pnas.1108745108 PMID: 21949389; PubMed Central PMCID: PMC3193229.

16. Amankwah EK, Lin HY, Tyrer JP, Lawrenson K, Dennis J, Chornokur G, et al. Epithelial-Mesenchymal Transition (EMT) Gene Variants and Epithelial Ovarian Cancer (EOC) Risk. Genet Epidemiol. 2015; 39 (8):689–97. https://doi.org/10.1002/gepi.21921 PMID: 26399219; PubMed Central PMCID: PMCPMC4721602.

17. Loftus SK, Baxter LL, Cronin JC, Fufa TD, Program NCS, Pavan WJ. Hypoxia-induced HIF1alpha targets in melanocytes reveal a molecular profile associated with poor melanoma prognosis. Pigment cell & melanoma research. 2017; 30(3):339–52. Epub 2017/02/09. https://doi.org/10.1111/pcmr.12579 PMID: 28168807; PubMed Central PMCID: PMCPMC5411287.

18. Li L, Darden TA, Weinberg CR, Levine AJ, Pedersen LG. Gene assessment and sample classification for gene expression data using a genetic algorithm/k-nearest neighbor method. Combinatorial chemistry & high throughput screening. 2001; 4(8):727–39. PMID: 11894805.

19. Li L, Weinberg CR, Darden TA, Pedersen LG. Gene selection for sample classification based on gene expression data: study of sensitivity to choice of parameters of the GA/KNN method. Bioinformatics (Oxford, England). 2001; 17(12):1131–42. https://doi.org/10.1093/bioinformatics/17.12.1131 PMID: 11751221.

20. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols. 2009; 4(1):44–57. https://doi.org/10.1038/nprot.2008.211 PMID: 19131956.

21. Pan Y, Robertson G, Pedersen L, Lim E, Hernandez-Herrera A, Rowat AC, et al. miR-509-3p is clinically significant and strongly attenuates cellular migration and multi-cellular spheroids in ovarian cancer. Oncotarget. 2016; 7(18):25930–48. Epub 2016/04/02. https://doi.org/10.18632/oncotarget.8412 PMID: 27036018; PubMed Central PMCID: PMCPMC5041955.

22. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nature reviews Molecular cell biology. 2014; 15(3):178–96. https://doi.org/10.1038/nrm3758 PMID: 24556840; PubMed Central PMCID: PMC4240281.

23. Kolde R, Laur S, Adler P, Vilo J. Robust rank aggregation for gene list integration and meta-analysis. Bioinformatics (Oxford, England). 2012; 28(4):573–80. https://doi.org/10.1093/bioinformatics/btr709 PMID: 22247279; PubMed Central PMCID: PMC3278763.

24. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012; 483 (7391):603–7. https://doi.org/10.1038/nature11003 PMID: 22460905; PubMed Central PMCID: PMCPMC3320027.

25. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nature medicine. 2013; 19(11):1423–37. https://doi.org/10.1038/nm.3394 PMID: 24202395; PubMed Central PMCID: PMC3954707.

26. Friedl P, Alexander S. Cancer invasion and the microenvironment: plasticity and reciprocity. Cell. 2011; 147(5):992–1009. https://doi.org/10.1016/j.cell.2011.11.016 PMID: 22118458.

27. Geiger TR, Peeper DS. Metastasis mechanisms. Biochimica et biophysica acta. 2009; 1796(2):293–308. https://doi.org/10.1016/j.bbcan.2009.07.006 PMID: 19683560.

28. Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell. 2010; 141(7):1117–34. https://doi.org/10.1016/j.cell.2010.06.011 PMID: 20602996; PubMed Central PMCID: PMC2914105.

29. Massague J. TGFbeta signalling in context. Nature reviews Molecular cell biology. 2012; 13(10):616–30. https://doi.org/10.1038/nrm3434 PMID: 22992590; PubMed Central PMCID: PMC4027049.

30. Cruciat CM, Niehrs C. Secreted and transmembrane wnt inhibitors and activators. Cold Spring Harbor perspectives in biology. 2013; 5(3):a015081. https://doi.org/10.1101/cshperspect.a015081 PMID: 23085770; PubMed Central PMCID: PMC3578365.

31. Andersson ER, Sandberg R, Lendahl U. Notch signaling: simplicity in design, versatility in function. Development. 2011; 138(17):3593–612. https://doi.org/10.1242/dev.063610 PMID: 21828089.

32. Briscoe J, Therond PP. The mechanisms of Hedgehog signalling and its roles in development and disease. Nature reviews Molecular cell biology. 2013; 14(7):416–29. https://doi.org/10.1038/nrm3598 PMID: 23719536.

33. Capurro M, Martin T, Shi W, Filmus J. Glypican-3 binds to Frizzled and plays a direct role in the stimulation of canonical Wnt signaling. J Cell Sci. 2014; 127(Pt 7):1565–75. https://doi.org/10.1242/jcs.140871 PMID: 24496449.

34. Capurro MI, Xiang YY, Lobe C, Filmus J. Glypican-3 promotes the growth of hepatocellular carcinoma by stimulating canonical Wnt signaling. Cancer research. 2005; 65(14):6245–54. https://doi.org/10.1158/0008-5472.CAN-04-4244 PMID: 16024626.

35. Wu Y, Liu H, Weng H, Zhang X, Li P, Fan CL, et al. Glypican-3 promotes epithelial-mesenchymal transition of hepatocellular carcinoma cells through ERK signaling pathway. International journal of oncology. 2015; 46(3):1275–85. https://doi.org/10.3892/ijo.2015.2827 PMID: 25572615.

36. Baumhoer D, Tornillo L, Stadlmann S, Roncalli M, Diamantis EK, Terracciano LM. Glypican 3 expression in human nonneoplastic, preneoplastic, and neoplastic tissues: a tissue microarray analysis of 4,387 tissue samples. American journal of clinical pathology. 2008; 129(6):899–906. https://doi.org/10.1309/HCQWPWD50XHD2DW6 PMID: 18480006.

37. Liu Y, Zheng D, Liu M, Bai J, Zhou X, Gong B, et al. Downregulation of glypican-3 expression increases migration, invasion, and tumorigenicity of human ovarian cancer cells. Tumour biology: the journal of

the International Society for Oncodevelopmental Biology and Medicine. 2015. https://doi.org/10.1007/s13277-015-3528-6 PMID: 25967456.

38. Peters MG, Farias E, Colombo L, Filmus J, Puricelli L, Bal de Kier Joffe E. Inhibition of invasion and metastasis by glypican-3 in a syngeneic breast cancer model. Breast cancer research and treatment. 2003; 80(2):221–32. https://doi.org/10.1023/A:1024549729256 PMID: 12908826.

39. Matsuda K, Maruyama H, Guo F, Kleeff J, Itakura J, Matsumoto Y, et al. Glypican-1 is overexpressed in human breast cancer and modulates the mitogenic effects of multiple heparin-binding growth factors in breast cancer cells. Cancer research. 2001; 61(14):5562–9. PMID: 11454708.

40. Gyvyte U, Juzenas S, Salteniene V, Kupcinskas J, Poskiene L, Kucinskas L, et al. MiRNA profiling of gastrointestinal stromal tumors by next-generation sequencing. Oncotarget. 2017; 8(23):37225–38. https://doi.org/10.18632/oncotarget.16664 PMID: 28402935; PubMed Central PMCID: PMCPMC5514905.

41. Streicher KL, Zhu W, Lehmann KP, Georgantas RW, Morehouse CA, Brohawn P, et al. A novel onco-genic role for the miRNA-506-514 cluster in initiating melanocyte transformation and promoting mela-noma growth. Oncogene. 2012; 31(12):1558–70. https://doi.org/10.1038/onc.2011.345 PMID: 21860416.

42. Saleiban A, Faxalv L, Claesson K, Jonsson JI, Osman A. miR-20b regulates expression of proteinase-activated receptor-1 (PAR-1) thrombin receptor in melanoma cells. Pigment cell & melanoma research. 2014; 27(3):431–41. https://doi.org/10.1111/pcmr.12217 PMID: 24405508.

43. Sun J, Li J, Zhang W, Zhang J, Sun S, Li G, et al. MicroRNA-509-3p Inhibits Cancer Cell Proliferation and Migration via Upregulation of XIAP in Gastric Cancer Cells. Oncol Res. 2017; 25(3):455–61. Epub 2016/10/05. https://doi.org/10.3727/096504016X14747283032017 PMID: 27697095.

44. Zhai Q, Zhou L, Zhao C, Wan J, Yu Z, Guo X, et al. Identification of miR-508-3p and miR-509-3p that are associated with cell invasion and migration and involved in the apoptosis of renal cell carcinoma. Biochemical and biophysical research communications. 2012; 419(4):621–6. Epub 2012/03/01. https://doi.org/10.1016/j.bbrc.2012.02.060 PMID: 22369946.