

1 **Title: Comprehensive analysis of differentially expressed miRNAs in**
2 **Hepatocellular carcinoma: prognostic significance and pathway**
3 **insights**

4

5 **Authors:** Kayleigh Smith ^{1*}, Dan Beach ^{2*}, Roger Silva ³, Gyorffy Balazs ⁴, Francesca Salani ^{1,5,6§+}
6 and Francesco Crea ^{§1}.

7 * contributed equally; co-first authors

8 § co-last authors

9 + corresponding author

10 **Affiliations**

11 ¹ Cancer Research Group-School of Life Health and Chemical Sciences, The Open University, Milton
12 Keynes, UK

13 ² School of Life Sciences, University of Sussex

14 ³ Department of Medicine - Cancer Research Program Research Institute of the McGill University Health
15 Centre - McGill University

16 ⁴ Department of Bioinformatics, Semmelweis University, Tüzoltó u. 7-9, 1094, Budapest, Hungary; and
17 Institute of Molecular Life Sciences, Research Centre for Natural Sciences, Magyar tudosok korutja 2,
18 1117, Budapest, Hungary

19 ⁵ Department of Translational Research and New Technologies in Medicine and Surgery, University of
20 Pisa, Italy

21 ⁶ Institute of Interdisciplinary Research "Health Science", Scuola Superiore Sant'Anna, Pisa, Italy

22

23 **Abstract:**

24 Robust prognostic and predictive factors for hepatocellular carcinoma, a leading cause of cancer-related deaths
25 worldwide, have not yet been identified. Previous studies have identified potential HCC determinants such as
26 genetic mutations, epigenetic alterations, and pathway dysregulation. However the clinical significance of these
27 molecular alterations remains elusive. MicroRNAs are major regulators of protein expression. MiRNA functions are
28 frequently altered in cancer. In this study, we aimed to explore the prognostic value of differentially expressed
29 miRNAs in HCC and elucidate their associated pathways. To this aim, bioinformatics techniques and clinical dataset
30 analyses were employed to identify differentially expressed miRNAs in HCC compared to normal hepatic tissue.
31 We validated known associations and identified novel miRNAs with potential prognostic significance and proposed
32 new targeting pathways based on our comprehensive analysis.

33

34

35 **1. Introduction**

36 Primary liver cancer is the second most common cause of cancer mortality in the world (1). Hepatocellular
37 carcinoma (HCC) is the most common histology identified in primary liver cancers. HCC prognosis is
38 primarily determined by stage, liver function, patient conditions and, indirectly, by HCC aetiology (2). The
39 five-year survival ranges from more than 50% in operable diseases to less than 5% in the most advanced
40 stages (3). This is due to the lack of effective treatments, especially in the setting of high tumour burden,
41 where systemic treatments can lead to a median survival of less than two years. Despite extensive
42 molecular studies which have identified potential determinants among genetic mutations, epigenetic
43 alterations and pathways dysregulation (4), HCC pathogenesis is not entirely elucidated. A better
44 understanding of the molecular processes driving HCC could lead to more effective treatments.

45 MicroRNAs (miRNAs) are short non-coding RNAs that regulate mRNA translation and stability (5). The
46 human genome encodes approximately 2000 miRNAs, but each of these transcripts can regulate the

47 expression of hundreds of mRNAs. It has been estimated that miRNAs regulate the expression of
48 approximately 60% of the protein-coding genes. The expression of miRNAs is deregulated in many
49 cancers (6), making them attractive biomarkers and therapeutic targets. The role of specific miRNAs in
50 HCC has been studied (7), but there are only limited experiences on comprehensive studies assessing
51 the prognostic function of miRNAs in HCC and the downstream pathways in this malignancy (8), (9),
52 leading to the identification of non-univocal signatures.

53 In this study, we used a combination of bio-informatic techniques and clinical dataset analyses to identify
54 miRNAs that are differentially expressed in HCC *versus* normal hepatic tissue and whose expression is
55 associated with HCC prognosis. This approach allowed us to confirm known associations, identify several
56 uncharacterised miRNAs and propose new potential targeting pathways based on our analysis.

57

58 **2. Material and Methods**

59 **2.1 Differential expression analysis**

60 A data-led approach was chosen over a literature-based candidate gene-identifying approach for the sake
61 of analysis comprehensiveness. OncoMir Cancer Database (OMCD) (10) was used to access data on
62 miRNA expression from Liver hepatocellular carcinoma (referred to as LIHC) dataset. LIHC dataset
63 comprises The Cancer Genome Atlas (TCGA) data from 426 samples, of which 375 HCC patients and 51
64 non-HCC liver tissue. Differential expression between HCC and non-HCC liver tissue was tested via *t-test*
65 on OMCD portal (Bonferroni-adjusted p values of < 0.05) and miRNAs whose expression in HCC was at
66 least 2 folds different (ratios > 2 or < 0.5) from non-HCC were retained for further analyses.

67 **2.2 Survival analysis**

68 The shortlisted differentially expressed miRNAs were cross-referenced with miRNAs RNA-seq TCGA
69 dataset provided by Kaplan-Meier Plotter (KMP) tool (11). Only RNA-seq data from TCGA samples of
70 patients who had not undergone any other treatment (either radio- or chemotherapy or systemic drugs)

71 than surgery were comprised in the survival analysis. Patients were dichotomized according to each
72 differentially expressed miRNA median values, into high- and low-expression cohorts. Overall survival
73 (OS) was chosen as the prognostic end-point of interest and compared among high- and low-expression
74 patients through log-rank test p-value. Each comparison is presented as median OS (mOS), Hazard Ratio
75 (HR) with 95% Interval of Confidence (95%IC) and p-value.

76 **2.3 Construction of a prognostic miRNA signature**

77 MiRNAs identified in 2.2 bearing a significant prognostic value were used to generate a prognostic
78 signature through miRpower tool in the KMP. “*Multiple genes*” and “*mean gene expression*” options were
79 set; patients were dichotomized according to the median value of the signature and the OS of the two
80 cohorts was compared via log-rank test.

81

82 **2.4 Pathway analysis**

83 Pathway analysis was conducted to explore the mechanisms underpinning the prognostic effects of these
84 differentially expressed miRNAs. This involved the identification of target genes for up- and down-
85 regulated miRNAs, and KEGG pathway analysis using the Over Representation Analysis (ORA) method.

86 **2.4.1 Selection of mature miRNAs**

87 The TCGA miRNA expression data used by the OMCD and KM Plotter databases do not differentiate
88 between mature 3p and 5p miRNAs (12). Both miRNA forms are capable of binding to target mRNAs
89 through a sequence-specific base pairing mechanism and regulate gene expression by interacting with
90 distinct sets of target genes. However, only one mature strand is typically incorporated into the RNA-
91 induced silencing complex (RISC) and involved in regulating mRNA transcription. Thus, for each miRNA,
92 it was necessary to determine whether the 3p or 5p form should be used for the analysis of downstream
93 target genes and pathways. The predominant mature form of each miRNA was selected based on the
94 number of experimental reads in the miRBase database (13) (**Table S1, supplementary**). The *hsa-miR-*

95 3653 identifier was subsequently excluded as the available experimental data no longer supports its
96 annotation as a miRNA (13).

97 **2.4.2 Target gene identification**

98 The first step in conducting the pathway analysis was to choose a database that provides miRNA-target
99 interaction data. MirTarBase is a manually curated database of experimentally validated miRNA-target
100 interactions (MTI) collated from reporter assay, western blot and next generation sequencing (NGS)
101 studies (14). The database includes more than 360,000 MTIs and represents the largest and most up to
102 date source of experimentally validated MTI data. MirTarBase was initially selected due its size and
103 rigorous inclusion criteria. An alternative MTI database, miRPathDB was found to include the latest
104 version mirTarBase data and was selected instead. The prognostic mature miRNA identifiers were
105 queried in the mirRPathDB database, retrieving a comprehensive list of target genes. This list was cross
106 validated using the miRTargetLink 2.0 database which provided a second mirror of the most recent
107 mirTarBase data.

108 A Python script was written to remove duplicates and genes with no experimental evidence of miRNA-
109 interaction. Two genes could not be mapped to current Ensembl/Entrez IDs and were excluded from
110 further analysis. The remaining target genes were relevant and aligned with the objective of the analysis.

111

112 **2.4.3 Functional annotation and enrichment analysis**

113 Functional annotation and enrichment analysis was conducted using the Database for Annotation,
114 Visualization and Integrated Discovery (DAVID) tool. Alternative online functional enrichment tools
115 WebGestalt and g:Profiler were evaluated, but DAVID was selected because it is highly cited, regularly
116 updated, and well documented (15). The target gene lists for up-regulated miRNAs and down-regulated
117 miRNAs were uploaded to DAVID which identified them as being for the *Homo Sapiens* species.
118 Functional annotation enrichment was performed for targets of upregulated miRNAs and downregulated
119 miRNAs separately, selecting only the KEGG pathway database option. Statistical analysis tested the

120 hypothesis that the overlap between genes in the target gene list and a given pathway or functional
121 category was greater than expected by chance. The statistical test used was Fisher's exact test and
122 Benjamini-Hochberg correction was applied to p-values to correct for multiple tests. Text-based result files
123 for each gene list were downloaded from DAVID. A Python script was written to filter pathways with a
124 Benjamini-Hochberg significance threshold of ≤ 0.05 .

125

126 **2.4.4 Analysis of enriched KEGG pathways**

127 To focus on the most relevant pathways, KEGG analysis results were filtered to include only pathways
128 that are putatively involved in HCC. Relevant pathways were defined as those listed as "related" in the
129 KEGG HCC Pathway (*hsa05225*) (16). A Python script was used to filter the KEGG results according to a
130 predefined set of keywords derived from these maps. The keyword list used to filter results was:
131 Hepatocellular, Hepatitis, Alcohol, NAFLD, *PI3K*, *AKT*, *P53*, *WNT*, *MAPK*, Cell cycle, calcium signaling,
132 *TGF-beta*.

133

134 **2.4.5 Validation of the functional enrichment analysis.**

135 A validation of the DAVID functional enrichment analysis was performed using the WebGeSTALT tool
136 (17). While the results from both analyses were largely concordant, WebGeSTALT identified a greater
137 number of enriched pathways, suggesting that the algorithm employed by DAVID was more conservative.

138

139 **2.4.6 Identifying direct miRNA interactions with cancer driver genes**

140 A previously reported pan-cancer analysis of more than 9000 tumor exomes from 33 TCGA projects
141 identified 299 oncogenic driver genes and 3473 associated driver missense mutations (18). Subsequent
142 functional validation utilizing an independent dataset of experimentally validated mutations indicated that

143 60-85% of the predicted mutations were likely to be cancer drivers. A list of 299 HCC and pan-cancer
144 associated genes was retrieved from TCGA and a Python script was used to return intersection between
145 the prognostic miRNA target genes and the 299 putative cancer driver genes.

146 3.0 Results

147 3.1 Some differentially expressed miRNAs are prognostic in resected HCC

148 Using OMCD, levels of expression of miRNAs in HCC *versus* healthy liver tissue were compared. Out of
149 the 106 significantly differentially expressed miRNAs, 59 had a significant, higher than 2-fold differential
150 expression (HCC/healthy liver either >2 or <0.5), compared to normal liver: respectively, 23 were
151 increased and 36 decreased in HCC (**Table S2, supplementary**). Forty of the differentially expressed
152 miRNAs were also present in the KMP RNA-seq dataset which we have employed for prognostic
153 stratification (21 of which up-regulated and 19 down-regulated in HCC *versus* healthy liver).

154 Of the 40 shortlisted miRNAs, 10 showed a statistically significant correlation with OS (p for log-rank test
155 <0.05); in particular, 5 among the HCC up-regulated (*hsa-miR-501*, *hsa-miR-877*, *hsa-miR-1180*, *hsa-*
156 *miR-3127*, *hsa-miR-3677*) and 5 among the HCC down-regulated miRNAs (*hsa-miR-3653*, *hsa-miR-99a*,
157 *hsa-miR-145*, *hsa-miR-326*, *hsa-miR-139*) (**Table 1**).

158 **Table 1: OS correlation of differentially expressed miRNAs (> 2-fold increase or decrease in HCC**
159 **vs normal livers), according to KMP; miRNAs in bold are significantly correlated with OS.**

160

MiRNA identifier	P value	HR	Median survival -	Median survival -
			Low expression	High expression
			(Months)	(Months)
<i>Increased expression in HCC</i>				
<i>hsa-miR-660</i>	0.6338	1.09 (0.77 - 1.54)	58.88	55.4
<i>hsa-miR-581</i>	0.4568	1.14 (0.81 - 1.61)	55.69	53.33

<i>hsa-miR-93</i>	0.455	1.14 (0.81 - 1.62)	55.69	80.75
<i>hsa-miR-10b</i>	0.2934	1.21 (0.85 - 1.71)	55.4	58.88
<i>hsa-miR-532</i>	0.2351	1.24 (0.87 - 1.75)	55.4	83.24
<i>hsa-miR-34a</i>	0.2247	0.81 (0.57 - 1.14)	55.4	58.88
<i>hsa-miR-183</i>	0.1846	1.27 (0.89 - 1.79)	70.06	45.93
<i>hsa-miR-500b</i>	0.1732	1.27 (0.9 - 1.81)	70.06	45.57
<i>hsa-miR-96</i>	0.1296	1.31 (0.92 - 1.85)	70.06	45.93
<i>hsa-miR-224</i>	0.1179	1.32 (0.93 - 1.88)	70.06	46.78
<i>hsa-miR-222</i>	0.0882	1.35 (0.95 - 1.92)	70.06	51.29
<i>hsa-miR-589</i>	0.0865	1.36 (0.96 - 1.92)	60.89	51.29
<i>hsa-miR-221</i>	0.0696	1.38 (0.97 - 1.96)	70.06	51.29
<i>hsa-miR-500a</i>	0.0666	1.38 (0.98 - 1.96)	69.57	58.88
<i>hsa-miR-452</i>	0.0657	1.39 (0.98 - 1.97)	80.75	45.93
<i>hsa-miR-21</i>	0.0552	1.41 (0.99 - 2)	70.06	46.78
<i>hsa-miR-501</i>	0.05	1.42 (1 - 2.01)	69.57	55.4
<i>hsa-miR-877</i>	0.0158	1.54 (1.08 - 2.18)	80.75	45.57
<i>hsa-miR-1180</i>	0.013	1.56 (1.1 - 2.23)	69.57	45.93
<i>hsa-miR-3127</i>	0.0055	1.65 (1.15 - 2.35)	83.57	45.93
<i>hsa-miR-3677</i>	0.0000015	2.42 (1.67 - 3.51)	80.75	37.32
<i>Decreased expression in HCC</i>				
<i>hsa-miR-542</i>	0.8858	0.97 (0.69 - 1.38)	60.89	55.69
<i>hsa-miR-214</i>	0.8781	1.03 (0.73 - 1.45)	70.06	53.33
<i>hsa-miR-29c</i>	0.8026	0.96 (0.68 - 1.35)	81.73	55.4
<i>hsa-miR-199b</i>	0.7972	1.05 (0.74 - 1.48)	60.89	55.4
<i>hsa-miR-33b</i>	0.5578	1.11 (0.78 - 1.57)	69.57	53.33
<i>hsa-miR-3614</i>	0.5502	0.9 (0.64 - 1.27)	60.89	55.69

<i>hsa-miR-10a</i>	0.4714	0.88 (0.62 - 1.25)	70.06	55.4
<i>hsa-miR-497</i>	0.4714	0.88 (0.62 - 1.25)	60.89	55.4
<i>hsa-miR-26b</i>	0.4599	1.14 (0.8 - 1.62)	107.11	48.99
<i>hsa-miR-424</i>	0.3057	1.2 (0.85 - 1.7)	81.73	51.29
<i>hsa-let-7c</i>	0.166	0.78 (0.55 - 1.11)	40.41	69.57
<i>hsa-miR-195</i>	0.1263	0.76 (0.54 - 1.08)	53.39	58.88
<i>hsa-miR-130a</i>	0.1166	0.76 (0.53 - 1.07)	51.29	69.57
<i>hsa-miR-3607</i>	0.1029	0.75 (0.53 - 1.06)	45.57	70.06
<i>hsa-miR-3653</i>	0.0376	0.69 (0.49 - 0.98)	41.79	80.75
<i>hsa-miR-99a</i>	0.0333	0.68 (0.48 - 0.97)	41.79	70.06
<i>hsa-miR-145</i>	0.0326	0.68 (0.48 - 0.97)	46.78	69.57
<i>hsa-miR-326</i>	0.0002	1.97 (1.37 - 2.83)	81.73	45.57
<i>hsa-miR-139</i>	0.00000057	0.41 (0.28 - 0.59)	30.61	80.75

161 Abbreviations: miRNA, microRNA; HCC, hepatocellular carcinoma; HR, hazard ratio; OS, overall survival.

162 In keeping with our hypothesis, we found that for the majority of miRNAs, expression in HCC *versus*
 163 normal tissue and prognostic significance were highly correlated. All the miRNAs up-regulated in HCC
 164 were associated with negative prognosis, 4 out of 5 miRNAs that are down-regulated in HCC were
 165 associated with better prognosis. *Hsa-miR-326* was the only transcript decreased in HCC showing an
 166 unexpected negative prognostic role (HR:1.97(1.37–2.83), p 0.0002).

167 To increase the robustness of our analysis, a 9-miRNA signature was generated from prognostically
 168 significant miRNAs (removing *Hsa-miR-326*). The signature yielded an HR for OS of 2.15 (1.52 – 3.05),
 169 with log-rank p of 0.00005 (**Figure 1A**), consistent across different clinical sub-groups, as shown in
 170 **Figure 1B-F**.

171 **Figure 1: Nine-miRNA prognostic signature predicting overall survival in the overall HCC cohort**
172 **(A) and sub-groups: White-Caucasians (B), Asians (C), well-differentiated (D) and poorly-**
173 **differentiated HCCs (E), high-TMB HCCs (F).**

174

175 Taken together, these results suggest that the miRNAs up-regulated in HCC and associated with worse
176 prognosis may be oncogenic, whilst the ones that are down-regulated in HCC and associated with better
177 prognosis may act as tumour suppressors.

178 **3.2 Functional pathways of miRNA-targeted genes**

179 To understand which molecular pathways are modulated by differential miRNA expression in HCC, we
180 used a dataset of validated miRNA/mRNA target pairs. A total of 982 unique target protein-coding genes
181 for the 10 prognostic miRNAs were retrieved from MiRPathDB. Among these, 430 genes were targeted
182 by up-regulated miRNAs, while 585 genes were targeted by down-regulated ones. We then run pathway
183 analyses on the two separate lists of protein coding genes.

184 Target genes of up-regulated miRNAs were significantly enriched in 7 KEGG pathways ($p < 0.05$), relating
185 mainly to neurodegenerative diseases. After filtering for HCC-specific pathways, cell cycle was the only
186 pathway associated with targets of up-regulated miRNAs (**Table 2**).

187 **Table 2: HCC-specific pathways relative to up-regulated miRNAs**

Term (KEEG)	Genes	%	P-Value	Fold Enrichment	Benjamini	FDR
hsa04110:Cell cycle	12	2.790698	0.000356	3.716564	0.019756	0.019392

188 Abbreviations: FDR, false discovery rate.

189 Targets of down-regulated miRNAs were significantly enriched in 93 KEGG pathways relating to
190 numerous cancers, cancer related processes and associated signalling pathways. Ten HCC-specific

191 pathways were significantly enriched ($p < 0.05$), including HCC, Hepatitis B, MAPK signalling pathway,
192 PI3K-Akt signalling pathway, Cell cycle, Hepatitis C, TGF-beta signalling and p53 signalling pathway
193 (Table 3).

194 **Table 3: HCC-specific pathways relative to down-regulated miRNAs**

Term (KEGG)	Genes	%	P-Value	Fold Enrichment	Benjamini	FDR
hsa05200:Pathways in cancer	58	9.7643	3.41E-12	2.71604428	2.11E-10	1.44E-10
hsa05225:Hepatocellular carcinoma	30	5.0505	1.73E-11	4.44033101	7.66E-10	5.22E-10
hsa05161:Hepatitis B	25	4.2088	2.39E-08	3.83732310	3.91E-07	2.66E-07
hsa04010:MAPK signaling pathway	34	5.7239	6.48E-08	2.87564294	8.74E-07	5.95E-07
hsa04151:PI3K-Akt signaling pathway	36	6.0606	5.86E-07	2.52873088	6.27E-06	4.27E-06
hsa05206:MicroRNAs in cancer	32	5.3872	2.11E-06	2.56679780	2.05E-05	1.39E-05
hsa04110:Cell cycle	16	2.6936	0.000146	3.15756872	8.54E-04	5.81E-04
hsa05160:Hepatitis C	18	3.0303	0.000174	2.85086220	9.65E-04	6.57E-04
hsa04350:TGF-beta signaling pathway	12	2.0202	0.001279	3.17436430	5.43E-03	3.70E-03
hsa04115:p53 signaling pathway	9	1.5152	0.008519	3.06565319	3.01E-02	2.05E-02

195 Abbreviations: FDR, false discovery rate.

196 A graphical representation of the results of the above mentioned enrichment analyses is presented in
197 **Figure 2**.

198 **Figure 2: Enrichment analysis for miRNA-targeted genes in KEGG pathways.** The shared x-axis
199 represents fold enrichment and False Discovery Rates (FDR) for both plots. All listed KEGG pathways
200 represent the enrichment of down-regulated miRNAs (left plot) and up-regulated miRNAs (right plot)
201 targeting genes in different pathways.

202

203

204

205

206 A Comprehensive overview of KEGG HCC-specific pathways can be found in **Figure S1**.

207 **Figure S1: KEGG HCC pathways. Enriched genes from target up- or down-regulated miRNAs are**
208 **red-highlighted.**

209

210 The process of intersecting the miRNA target genes with the HCC and pan cancer associated genes revealed that the
211 list of 982 miRNA target genes contained 10 HCC driver genes and 11 pan-cancer driver genes (**Table 4**).

212 **Table 4: HCC- and pan-cancer driver genes among miRNA target genes.**

213

Symbol	Ensembl Gene ID	Chr	Position (Mbp)	Description
<i>THRAP3</i>	ENSG00000054118	1	36.2244	thyroid hormone receptor associated protein 3
<i>NRAS</i>	ENSG00000213281	1	114.7045	NRAS proto-oncogene, GTPase
<i>NOTCH2</i>	ENSG00000134250	1	119.9116	notch receptor 2
<i>TBL1XR1</i>	ENSG00000177565	3	177.0193	TBL1X receptor 1
<i>PIK3CA</i>	ENSG00000121879	3	179.1481	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha
<i>MSH3</i>	ENSG00000113318	5	80.6547	mutS homolog 3
<i>H1-4</i>	ENSG00000168298	6	26.1563	H1.4 linker histone, cluster member
<i>CDKN1A</i>	ENSG00000124762	6	36.6765	cyclin dependent kinase inhibitor 1A
<i>EEF1A1</i>	ENSG00000156508	6	73.4893	eukaryotic translation elongation factor 1 alpha 1
<i>ESR1</i>	ENSG00000091831	6	151.6567	estrogen receptor 1

<i>MYC</i>	ENSG00000136997	8	127.7354	MYC proto-oncogene, bHLH transcription factor
<i>KMT2A</i>	ENSG00000118058	11	118.4365	lysine methyltransferase 2A
<i>KRAS</i>	ENSG00000133703	12	25.2052	KRAS proto-oncogene, GTPase
<i>RB1</i>	ENSG00000139687	13	48.3037	RB transcriptional corepressor 1
<i>CHD8</i>	ENSG00000100888	14	21.3852	chromodomain helicase DNA binding protein 8
<i>BRD7</i>	ENSG00000166164	16	50.3135	bromodomain containing 7
<i>BCL2</i>	ENSG00000171791	18	63.1233	BCL2 apoptosis regulator
<i>SMARCA4</i>	ENSG00000127616	19	10.9609	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4
<i>RPS6KA3</i>	ENSG00000177189	X	20.1499	ribosomal protein S6 kinase A3
<i>USP9X</i>	ENSG00000124486	X	41.0854	ubiquitin specific peptidase 9 X-linked
<i>AR</i>	ENSG00000169083	X	67.544	androgen receptor

214 Thus, there is experimental evidence that the 9 prognostic miRNAs may also directly interact with 21 putative HCC
215 and pan-cancer driver genes.

216

217 **4.0 Discussion**

218 In this study, we have identified a 9-miRNA prognostic signature for HCC. If confirmed in clinical datasets,
219 this signature could enable patient stratification and treatment optimisation, especially for low- and
220 intermediate-risk diseases. At these stages, several therapeutic options are available, but currently these
221 treatments cannot be tailored based on individual molecular profiles. Interestingly, some miRNAs are well
222 detectable in biological fluids, such as plasma (19). If the 9-miRNA signature proves to be detectable in
223 plasma samples in the HCC population, it can be used as a minimally invasive biomarker for patient
224 stratification and dynamic monitoring.

225 Several experimental studies have dissected the pathogenic role of specific miRNAs in HCC. Our results
226 are in line with mechanistical evidence. Of the five up-regulated/negatively prognostic miRNAs, four have

227 been shown to act as oncogenes in HCC cells: *mir-501* by promoting epithelial to mesenchymal transition
228 (20); *mir-3127* and *mir-1180* by promoting proliferation and tumorigenesis (21), (22); *mir-3677* by
229 promoting proliferation and drug resistance (23). Notably, our pathway analysis has identified cell cycle
230 regulation as the main molecular pathway controlled by these miRNAs, further confirming the link
231 between bioinformatic and experimental findings.

232 Our results on *mir-877* are apparently in contrast with previous reports, showing that this transcript inhibits
233 HCC proliferation (24). We found that this miRNA is up-regulated in HCC *versus* normal tissue, and that
234 higher expression of this miRNA predicts poor prognosis. This implies an oncogenic function. Notably,
235 another study showed that a genetic variant of this miRNA predicts HCC risk (25). It is not known if this
236 genetic variant has a prognostic role in advanced disease settings. It is therefore conceivable that this
237 miRNA could play different roles at different stages of disease progression (e.g. onco-suppressive for
238 early stage, oncogenic for advanced stage), as demonstrated for other cancer-related genes (26). Further
239 studies are therefore warranted to dissect the role of this miRNA at different stages of HCC progression,
240 and to confirm its prognostic role in other clinical datasets.

241 In parallel with the results on putative oncogenic miRNAs, our results on putative tumour suppressors
242 agree with pre-clinical results. *Mir-3653*, *mir-326* and *mir-139* have been shown to inhibit metastasis and
243 progression of HCC cells (27), (28), (29). Similarly, *mir-145* has been implicated in G2/M phase arrest by
244 inhibiting cyclin B1 (30). Finally, a recent study has shown that *mir-99a* is a direct target of the epigenetic
245 silencer EZH2 (31). EZH2 promotes cell proliferation, metastasis and drug resistance in several cancers
246 by silencing the expression of tumour suppressors (32). We have shown that higher EZH2 activity
247 (measured via a liquid biopsy approach) predicts poorer prognosis in HCC patients treated with sorafenib
248 (33). Hence our finding on *mir-99a* paves the way for the combined use of miRNA and epigenetic
249 biomarkers for patient stratification and drug efficacy prediction in HCC.

250 In conclusion, we have identified for the first time a 9-miRNA signature that is differentially expressed in
251 HCC and robustly predicts prognosis. This signature includes transcripts whose function has been
252 independently validated in pre-clinical studies. Future studies on clinical samples (primary tumours and
253 plasma) are warranted to confirm the clinical usefulness of this potential new biomarker panel.

254

255 References

1. McGlynn KA, Petrick JL, El-Serag HB. Epidemiology of Hepatocellular Carcinoma. *Hepatology*. 2021 Jan;73(Suppl 1):4–13.
2. Ito T, Nguyen MH. Perspectives on the Underlying Etiology of HCC and Its Effects on Treatment Outcomes. *Journal of Hepatocellular Carcinoma*. 2023;10:413.
3. Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado Á, et al. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. *Journal of Hepatology*. 2022 Mar 1;76(3):681–93.
4. Tümen D, Heumann P, Gülow K, Demirci CN, Cosma LS, Müller M, et al. Pathogenesis and Current Treatment Strategies of Hepatocellular Carcinoma. *Biomedicines*. 2022 Dec;10(12):3202.
5. Naeli P, Winter T, Hackett AP, Alboushi L, Jafarnejad SM. The intricate balance between microRNA-induced mRNA decay and translational repression. *The FEBS Journal*. 2023;290(10):2508–24.
6. Smolarz B, Durczyński A, Romanowicz H, Szyłto K, Hogendorf P. miRNAs in Cancer (Review of Literature). *Int J Mol Sci*. 2022 Mar 3;23(5):2805.
7. Li J, Bao H, Huang Z, Liang Z, Wang M, Lin N, et al. Little things with significant impact: miRNAs in hepatocellular carcinoma. *Frontiers in Oncology* [Internet]. 2023 [cited 2023 Nov 29];13. Available from: <https://www.frontiersin.org/articles/10.3389/fonc.2023.1191070>
8. Chen Z, Zhang L, Ding C, Ren K, Wan D, Lin S. A six-miRNA signature as a novel biomarker for improving prediction of prognosis and patterns of immune infiltration in hepatocellular carcinoma. *Am J Transl Res*. 2022 Jun 15;14(6):3610–37.

9. A five-miRNA expression signature predicts survival in hepatocellular carcinoma - Liu - 2017 - APMIS - Wiley Online Library [Internet]. [cited 2023 Nov 29]. Available from: <https://onlinelibrary.wiley.com/doi/10.1111/apm.12697>
10. Sarver AL, Sarver AE, Yuan C, Subramanian S. OMCD: OncomiR Cancer Database. *BMC Cancer*. 2018 Dec 6;18(1):1223.
11. Györfy B. Discovery and ranking of the most robust prognostic biomarkers in serous ovarian cancer. *GeroScience*. 2023 Mar 1;45(3):1889–98.
12. Kuo WT, Su MW, Lee YL, Chen CH, Wu CW, Fang WL, et al. Bioinformatic Interrogation of 5p-arm and 3p-arm Specific miRNA Expression Using TCGA Datasets. *J Clin Med*. 2015 Sep 15;4(9):1798–814.
13. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res*. 2019 Jan 8;47(Database issue):D155–62.
14. Huang HY, Lin YCD, Cui S, Huang Y, Tang Y, Xu J, et al. miRTarBase update 2022: an informative resource for experimentally validated miRNA–target interactions. *Nucleic Acids Res*. 2021 Nov 30;50(D1):D222–30.
15. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research*. 2022 Jul 5;50(W1):W216–21.
16. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Research*. 2021 Jan 8;49(D1):D545–51.
17. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Research*. 2019 Jul 2;47(W1):W199–205.

18. Bailey MH, Tokheim C, Porta-Pardo E, Sengupta S, Bertrand D, Weerasinghe A, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell*. 2018 Apr 5;173(2):371-385.e18.
19. Watahiki A, Macfarlane RJ, Gleave ME, Crea F, Wang Y, Helgason CD, et al. Plasma miRNAs as Biomarkers to Identify Patients with Castration-Resistant Metastatic Prostate Cancer. *International Journal of Molecular Sciences*. 2013 Apr;14(4):7757–70.
20. Yu W, Deng W, Zhao Q, Zhuang H, Zhang C, Jian Z. miR-501 acts as an independent prognostic factor that promotes the epithelial–mesenchymal transition through targeting JDP2 in hepatocellular carcinoma. *Human Cell*. 2019 Jul 1;32(3):343–51.
21. Jiang J, Zhang Y, Guo Y, Yu C, Chen M, Li Z, et al. MicroRNA-3127 promotes cell proliferation and tumorigenicity in hepatocellular carcinoma by disrupting of PI3K/AKT negative regulation. *Oncotarget*. 2015 Jan 31;6(8):6359–72.
22. Zhou X, Zhu H qiang, Ma C qun, Li H guang, Liu F feng, Chang H, et al. MiR-1180 promoted the proliferation of hepatocellular carcinoma cells by repressing TNIP2 expression. *Biomedicine & Pharmacotherapy*. 2016 Apr 1;79:315–20.
23. He H, Zhou J, Cheng F, Li H, Quan Y. MiR-3677-3p promotes development and sorafenib resistance of hepatitis B-related hepatocellular carcinoma by inhibiting FOXM1 ubiquitination. *Human Cell*. 2023 Sep 1;36(5):1773–89.
24. Huang X, Qin J, Lu S. Up-regulation of miR-877 induced by paclitaxel inhibits hepatocellular carcinoma cell proliferation though targeting FOXM1. *Int J Clin Exp Pathol*. 2015 Feb 1;8(2):1515–24.
25. Wang H, Wang B, Wang T, Fan R. A genetic variant in the promoter region of miR-877 is associated with an increased risk of hepatocellular carcinoma. *Clinics and Research in Hepatology and Gastroenterology*. 2020 Oct 1;44(5):692–8.

26. Qin Y, Ashrafizadeh M, Mongiardini V, Grimaldi B, Crea F, Rietdorf K, et al. Autophagy and cancer drug resistance in dialogue: Pre-clinical and clinical evidence. *Cancer Letters*. 2023 Aug 28;570:216307.
27. Zhang L, Zhang T, Deng Z, Sun L. MicroRNA-3653 inhibits the growth and metastasis of hepatocellular carcinoma by inhibiting ITGB1. *Oncology Reports*. 2019 Mar 1;41(3):1669–77.
28. Wong CC, Wong C, Tung EK, Au SL, Lee JM, Poon RT, et al. The MicroRNA miR-139 Suppresses Metastasis and Progression of Hepatocellular Carcinoma by Down-regulating Rho-Kinase 2. *Gastroenterology*. 2011 Jan 1;140(1):322–31.
29. Liu C, Xu K, Liu J, He C, Liu P, Fu Q, et al. LncRNA RP11-620J15.3 promotes HCC cell proliferation and metastasis by targeting miR-326/GPI to enhance glycolysis. *Biol Direct*. 2023 Apr 5;18:15.
30. Lai Y, Liu J, Hu X, Zeng X, Gao P. Modifications of The Human Liver Cancer Cells through microRNA-145-Mediated Targeting of CDCA3. *Cell J*. 2023 Aug;25(8):546–53.
31. Wu CS, Chien YC, Yen C, Wu JY, Bai LY, Yu YL. EZH2-mediated epigenetic silencing of tumor-suppressive let-7c/miR-99a cluster by hepatitis B virus X antigen enhances hepatocellular carcinoma progression and metastasis. *Cancer Cell International*. 2023 Sep 9;23(1):199.
32. Crea F, Fornaro L, Bocci G, Sun L, Farrar WL, Falcone A, et al. EZH2 inhibition: targeting the crossroad of tumor invasion and angiogenesis. *Cancer Metastasis Rev*. 2012 Dec 1;31(3):753–61.
33. Salani F, Latarani M, Casadei-Gardini A, Gangadharannambiar P, Fornaro L, Vivaldi C, et al. Predictive significance of circulating histones in hepatocellular carcinoma patients treated with sorafenib. *Epigenomics*. 2022 May;14(9):507–17.

256 **Supporting information**

257 **Table S1. Number of experimental reads of the predominant mature form of each miRNA.**

258 **Table S2. Fifty-nine differentially expressed miRNAs (2-fold) in HCC vs normal livers according to**
259 **OMCD.**

260 **Figure S1. KEGG HCC pathways. Enriched genes from target up- or down-regulated miRNAs are**
261 **red-highlighted.**

262 **Figure 1. Nine-miRNA prognostic signature predicting overall survival in the overall HCC cohort**
263 **(A) and sub-groups: White-Caucasians (B), Asians (C), well-differentiated (D) and poorly-**
264 **differentiated HCCs (E), high-TMB HCCs (F).**

265 **Figure 2. Enrichment analysis for miRNA-targeted genes in KEGG pathways.**

266

267

268

269

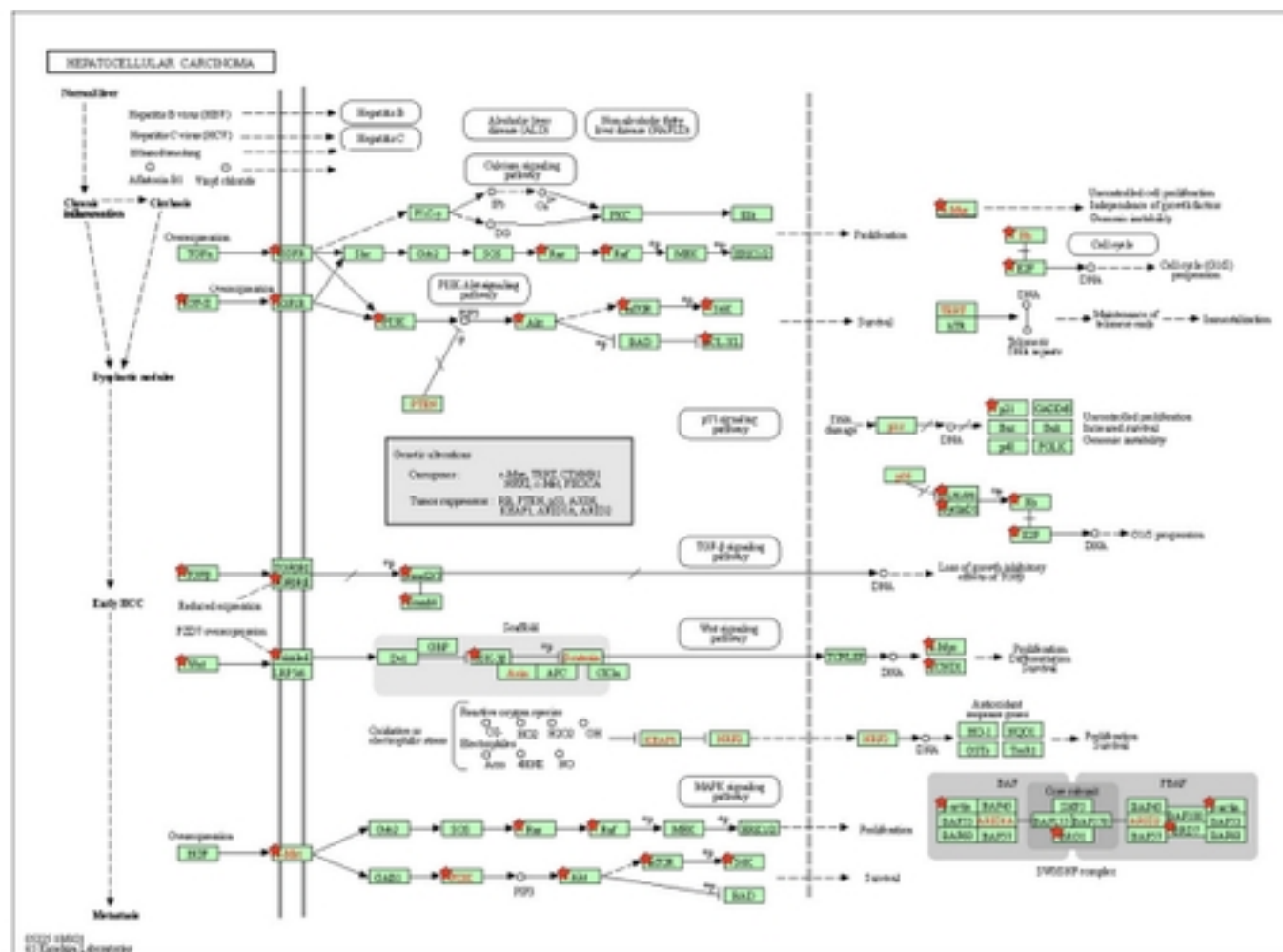


Figure S1. KEGG HCC pathways. Enriched genes from target up- or down-regulated miRNAs are red-highlighted.

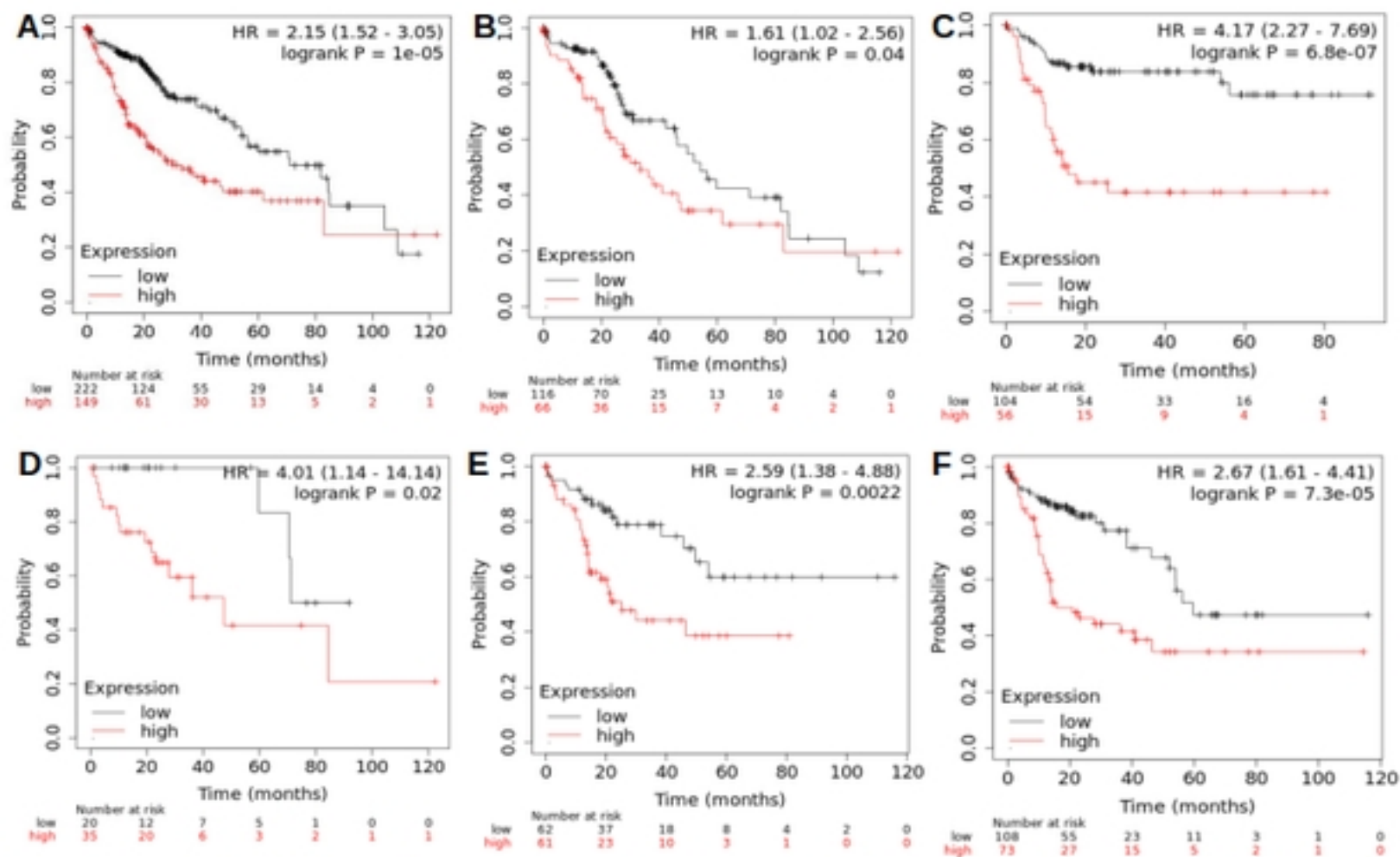


Figure 1. Nine-miRNA prognostic signature predicting overall survival in the overall HCC cohort (A) and sub-groups: White-Caucasians (B), Asians (C), well-differentiated (D) and poorly-differentiated HCCs (E), high-TMB HCCs (F).

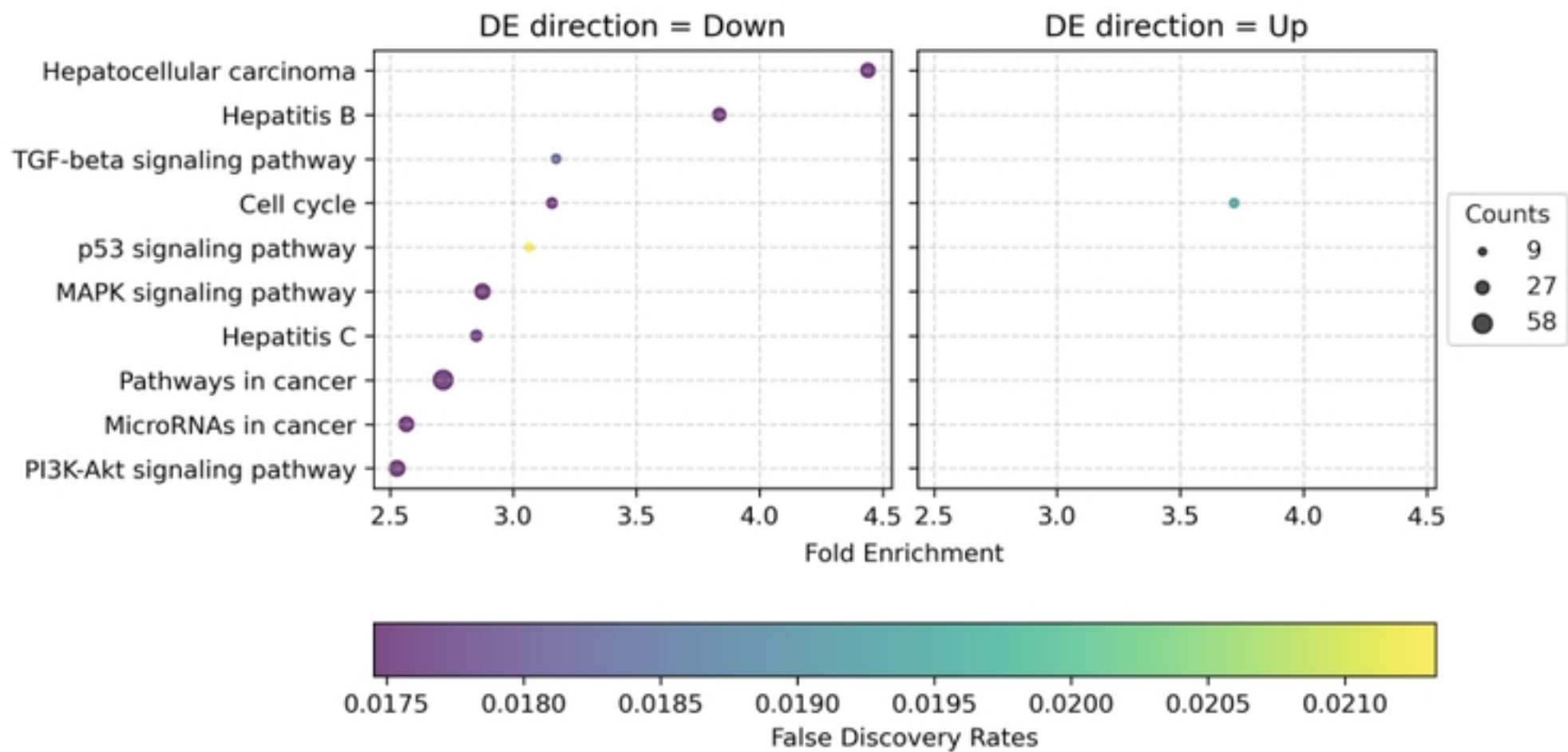


Figure 2. Enrichment analysis for miRNA-targeted genes in KEGG pathways. The shared x-axis represents fold enrichment and False Discovery Rates (FDR) for both plots. All listed KEGG pathways represent the enrichment of down-regulated miRNAs (left plot) and up-regulated miRNAs (right plot) targeting genes in different pathways.