Gastric cancer-associated microRNA expression signatures: integrated bioinformatics analysis, validation, and clinical significance

Xiaoguang Liu[1,2,3#^], Ke Pu[1,3#^], Yuping Wang[1,3^], Yanfei Chen[2], Yongning Zhou[1,3^]

[1]Key Laboratory for Gastrointestinal Diseases of Gansu Province, Lanzhou University, Lanzhou, China; [2]Department of Rheumatology and Immunology, The First Hospital of Lanzhou University, Lanzhou, China; [3]Department of Gastroenterology, The First Hospital of Lanzhou University, Lanzhou, China

Contributions: (I) Conception and design: X Liu; (II) Administrative support: Y Zhou; (III) Provision of study materials or patients: K Pu; (IV) Collection and assembly of data: Y Wang, Y Chen, Y Zhou; (V) Data analysis and interpretation: Y Wang, Y Chen, Y Zhou; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed to this article equally.

Correspondence to: Yongning Zhou. Department of Gastroenterology, The First Hospital of Lanzhou University, Lanzhou 730000, China; Key Laboratory for Gastrointestinal Diseases of Gansu Province, Lanzhou University, Lanzhou 730000, China. Email: zhouyn@lzu.edu.cn.

Background: Gastric cancer (GC) is one of the common gastrointestinal malignancy worldwide and exhibits a poor prognosis. Increasing studies have indicated that microRNAs play critical roles in the cancer progression and have shown great potential as useful biomarkers. The search for potential diagnostic and prognostic biomarkers of gastric cancer (GC) with integrated bioinformatics analyses has been undertaken in previous studies.

Methods: In this study, the robust rank aggregation (RRA) method was used to perform an integrated analysis of differentially expressed miRNAs (DEMs) from five microarray datasets in the Gene Expression Omnibus (GEO) database to find robust biomarkers for GC. Ultimately, seven miRNAs were filtered from fourteen primary miRNAs using the validation set of The Cancer Genome Atlas (TCGA) database. Based on these results, diagnostic and survival analyses were performed, and logistic regression and Cox regression were used to determine the clinicopathological characteristics of the DEM expression and overall survival.

Results: Nine eligible miRNA datasets related to GC were selected from the GEO database for integrated analysis in this study. Diagnostic analysis implied that these miRNAs could be regarded as promising candidate diagnostic biomarkers in GC tissues, but whether the results of the tissue analysis are consistent with those of peripheral blood analysis requires further validation. The logistic regression indicated that the ectopic expression of these DEMs was relevant to the histological type, anatomical region, and pathological grade of GC. However, the survival and Cox regression analyses suggested that the poor prognosis of GC patients was not strongly dependent on the ectopic expression of the seven miRNAs, but rather, a poor prognosis was associated with age, metastasis, and histological grade.

Conclusions: Based on the results presented in this study it can be concluded that these miRNAs (miR-455-3p, miR-135b-5p, let-7a-3p, miR-195-5p, miR-204-5p, miR-149-5p, and miR-143-3p) might be potential biomarkers for the early diagnosis of GC patients, but this finding should be regarded with caution. A large-scale, prospective, and multicenter cohort study should be performed.

Keywords: miRNA; gastric cancer (GC); robust rank aggregation; biomarkers

Submitted Mar 10, 2021. Accepted for publication Apr 29, 2021.
doi: 10.21037/atm-21-1631
View this article at: http://dx.doi.org/10.21037/atm-21-1631

^ ORCID: Xiaoguang Liu, 0000-0001-7335-6364; Ke Pu, 0000-0002-9627-4887; Yuping Wang, 0000-0003-0087-4771; Yongning Zhou, 0000-0002-7304-929X.

Introduction

Gastric cancer (GC) is the fifth most common cancer worldwide after lung cancer, breast cancer, rectal cancer, and prostate cancer, and it is the second major cause of cancer-related deaths worldwide (1). The pathogenesis of GC is complex and involves various factors such as diet, environment, infection, and genetics (2). Symptoms of epigastric pain and weight loss may occur in patients in the early stages of GC (3). Due to the lack of specificity of early symptoms, most patients are diagnosed at an advanced stage, which has led to a high mortality rate (4). Therefore, there is an urgent need to find new biomarkers for early diagnosis and effective symptomatic treatment of GC.

MicroRNAs (miRNAs) are endogenous noncoding RNAs, 17–25 nucleotides in length, that regulate gene expression at the posttranscriptional level. The ability to bind complementary sequences in 3'-untranslated regions (3'-UTRs) of target mRNAs promotes direct mRNA degradation or translational repression (5). MiRNAs play an important role in a variety of cellular biological processes, including development, differentiation, angiogenesis, and growth control (6). As posttranscriptional regulators, miRNAs affect the biological processes of cancer, including angiogenesis, tumor proliferation, and metastasis (7). Growing evidence indicates that miRNAs are differentially expressed in GC and are related to survival prognosis (8).

Many recent studies have revealed miRNAs as potential diagnostic or prognostic indicators of GC, but results have been inconsistent. Factors such as artifacts in the quality control of samples, different profiling platforms, and test methods have accounted for these inconclusive results. Moreover, intrinsic reasons such as the heterogeneity of the GC type (9), CpG island methylation (10), and single nucleotide polymorphisms (SNPs) (11,12) have also contributed to the indefinite findings.

To acquire robust and stable results, the robust rank aggregation (RRA) approach (13), which is an accurate and effective method to integrate differentially expressed signatures (14), was used to compare the different datasets. All miRNAs were assigned and reranked according to their P value. Wang *et al.* (15) found robust and strong prognostic signatures of GC from differentially expressed genes of eight Gene Expression Omnibus (GEO) microarrays using the RRA method. However, the search for promising diagnostic and prognostic miRNA biomarkers for GC remains worthy of investigation. Therefore, the aim of this study is to improve the understanding of the molecular mechanism of these miRNAs and identify more specific tumor biomarkers for GC. We present the following article in accordance with the MDAR checklist (available at http://dx.doi.org/10.21037/atm-21-1631).

Methods

Microarray data processing and differentially expressed microRNA (DEM) identification

Microarray repositories that provide miRNA expression profiles of GC were retrieved from the GEO database from December 2018 onwards. The following search terms were applied: stomach neoplasms OR "stomach neoplasms" OR "gastric cancer" and "microRNAs" or "miRNA" and "expression profiling by array" OR "expression". Eligible criteria for selection were that the data contained the microarray expression profiles of miRNAs for both GC and para-carcinoma tissues in humans. Of the datasets retrieved, 204 datasets were found, and of these, nine (GSE23739, GSE26595, GSE26645, GSE28700, GSE33743, GSE54397, GSE63121, GSE78091, and GSE93415) were finally retained after excluding repeated studies and studies of non-whole-genome expression. All raw datasets were normalized individually based on log2-transformation by the Robust Multi-Array Average (RMA). Next, Linear Models for Microarray (LIMMA), a Bioconductor package, was applied to determine the DEMs between GC and adjacent normal samples. A P value <0.05 and a fold change (FC) >1.0 (16) were established as the cut-off points to screen the significant DEMs. Various datasets of microarray chips were integrated and ranked in an unbiased manner with RRA methods (R package) (13). In addition, various probe IDs from the microarray were converted to gene IDs, and all miRNA names were normalized in accordance with the miRBase online database. If any probe could not be matched with the corresponding gene ID, it was regarded as a viral miRNA or a non-miRNA probe and was then removed. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Target gene prediction of common DEMs

The potential target genes of the DEMs were predicted with the starBase web tool (http://starbase.sysu.edu.cn), which includes seven databases: Targetscan, miRanda,

miRmap, PITA, RAN22, PicTar, and microT. To obtain more reliable target genes, target genes were determined by the consensus genes of the four online tools, and the cross-linking immunoprecipitation (CLIP) data was set with a high stringency (>5).

Functional analysis of target genes

Functional annotation of the Gene Ontology (GO) analysis and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was performed using the online STRING software tool (https://string-db.org/) and the Cytoscape software plug-in ClueGo (17), which provides a comprehensive set of functional annotation information of genes and proteins to allow users to extract biological information. Target genes with P values less than 0.05 from the GO analysis and KEGG pathway enrichment were defined as statistically significant.

Construction of the miRNA-gene network

A DEM and target gene network was constructed and visualized with the Cytoscape 3.6.1 software (18).

Statistical analysis

The differential expression analysis of nine microarrays was analyzed with R-software, and the differential expression validation using the TCGA database was performed with an unpaired t-test and GraphPad Prism software. The DEMs associated with clinical characteristics were analyzed by chi-square and t-tests, and diagnostic performance was assessed by receiver operating characteristic (ROC) analysis; both analyses were performed with IBM SPSS version 19.0 (SPSS, Chicago, IL, USA), and $P<0.05$ was considered statistically significant. Logistic regression and multivariate Cox regression analyses, which were performed with STATA software version 12.0 (Stata Corporation, College Station, TX, USA), were used to evaluate the relevant influence of DEMs and survival time.

Results

Characteristics of the included miRNA expression profiling of microarrays

In this study, nine eligible miRNA datasets related to GC [GSE23739 (19), GSE26595 (20), GSE26645, GSE28700, GSE33743 (21), GSE54397 (22), GSE63121 (23), GSE78091, and GSE93415] met the inclusion criteria and were selected from the GEO database for integrated analysis (*Table 1*). In the GSE2645 and GSE93415 datasets, a total of 24 paired samples were from GC and adjacent normal tissue, while the remaining samples from seven microarrays were from GC and normal gastric tissue. In GSE54397, an investigation of the miRNAs that are differentially expressed in the intestinal type of GC using a miRNA microarray was conducted. Moreover, the miRNA expression according to the microarray confirmed that the pathogenesis of GC differed between *H. pylori*-positive and *H. pylori*-negative patients. In addition, among the four microarrays including GSE26595, GSE26645, GSE33743, and GSE63121 the fold change (FC) value of miRNA differential expression treated with log2 transformation was less than 1.0 (logFC <1.0).

Identification of common DEMs in GC microarrays

To determine common DEMs for GC, a multi-step strategy was adopted to acquire valuable DEMs for the prognosis of GC patients. First, nine microarray datasets containing tumor and normal samples were downloaded from the GEO database. Significant DEMs with a FC >1.0 and a P value <0.05 were established as screening criteria in the individual microarray analysis (Figure S1). Next, from the original nine microarrays, five (GSE23739, GSE28700, GSE54397, GSE78091, and GSE93415) were screened for integration. Then, RRA methods were utilized to integrate and rank the DEMs, and from 1,128 DEMs, we obtained fifteen DEMs with an adjusted P value <0.05 (*Figure 1*). These included five upregulated miRNAs (miR-455-3p, miR-135b-5p, let-7a-3p, miR-218-5p, and miR-548au-3p), six downregulated miRNAs (miR-195-5p, miR-148a-3p, miR-204-5p, miR-149-5p, miR-143-3p, and miR-193b-3p), and three miRNAs (miR-1-3p, miR-199a-5p, and miR-3910) with unclear expression changes in GC tissues compared with normal tissues or adjacent normal tissues (*Table 2*).

Common miRNA putative target genes and GO/pathway enrichment analysis

The primary function of miRNAs is the repression of target gene expression via binding to specific target sites (24). Hence, the online tool starBase v3.0 was utilized to forecast

Table 1 The datasets characteristic of nine microarray chips

GEO accession	Platform of dataset	Region	Submission year	Type of platform	Sample	All miRNAs	DEMs	References
GSE23739	GPL7731	Switzerland	2011	Agilent-019118 Human miRNA Microarray 2.0 G4470B	40 paired of GC *vs.* GNT	720	122	Oh HK
GSE26595	GPL8179	South Korea	2011	Illumina Human v2 MicroRNA expression beadchip	60 paired of GC *vs.* GNT	360	5	Lim JY
GSE26645	GPL11487	China	2011	Agilent-021827 Human miRNA Microarray	4 paired of GC *vs.* ANT	851	3	–
GSE28700	GPL9081	Taiwan	2011	Agilent-016436 Human miRNA Microarray 1.0 G4472A	22 paired of GC *vs.* GNT	470	25	–
GSE33743	GPL14895	Portugal	2011	miRNAChip_human_V2	37 primary GC *vs.* 10 GNT	703	7	Carvalho J
GSE54397	GPL15159	South Korea	2014	Agilent-031181 Unrestricted_Human_miRNA_V16.0_ Microarray 030840	16 paired of GC *vs.* GNT	1205	50	Chang H
GSE63121	GPL8786	China	2014	Affymetrix Multispecies miRNA-1 Array	15 paired of GC *vs.* GNT	848	24	Zhang X
GSE78091	GPL21439	China	2016	miRCURY LNA microRNA Array, 7th generation - hsa, mmu & rno	3 paired of GC *vs.* GNT	1,921	385	–
GSE93415	GPL19071	Poland	2017	Exiqon miRCURY LNA microRNA array; 7th generation REV - hsa, mmu & rno; batch 208520-22; lot 35101-35101	20 paired of GC *vs.* ANT	891	107	–

GEO, Gene Expression Omnibus; DEMs, differentially expressed miRNAs; GC, gastric cancer; GNT, gastric normal tissue; ANT, adjacent normal tissue; LNA, Locked Nucleic Acid; REV, Revision.

the potential 3'-UTR (3'-untranslated region) of target genes with a strict stringency (>5) of CLIP data and to illustrate the biological function and pathway enrichment of DEMs using the STRING and DAVID software.

As a result, 672 consensus target genes with fourteen common DEMs were obtained from the abovementioned online software, and the GO and pathway enrichment analysis of these target genes was conducted individually. As shown in Table S1, the correlation between the high expression of miRNAs and clinicopathological features of GC patients illustrated that the target genes of common DEMs mainly regulated the following biological processes (BP): macromolecule metabolic process, cellular metabolic process, and protein location. The molecular function (MF) was particularly associated with protein activity, heterocyclic compound binding, and transferase activity. The cellular component (CC) of DEMs was distributed in the intracellular part, cytosol, nucleus, and cytoplasm. Moreover, microRNAs in cancer (GC, hepatocellular carcinoma, colorectal cancer, and breast cancer), and in cancer-related pathways (mTOR signaling pathway, Hippo signaling pathway, PI3K/AKT signaling pathway, FoxO signaling pathway, and Wnt signaling pathway) and proteoglycans in the cancer pathway primarily had roles in the KEGG pathways (*Table 3*).

However, Path-net analysis with ClueGo software was performed to delineate the interaction among 14 significant pathways from 175 enrichment pathways (*Figure 2*), which showed that these pathways including cancer pathways (degree =45), cellular senescence (degree =21), TGF-beta signing pathway (degree =12), HIF-1 signaling pathway (degree =13), and mTOR signaling pathway (degree =18) with the highest degree might play a key role in GC tumorigenesis. Therefore, on the basis of the significant GO and KEGG pathways, miR-target gene networks, miR-pathway networks, and miR-GO-networks were

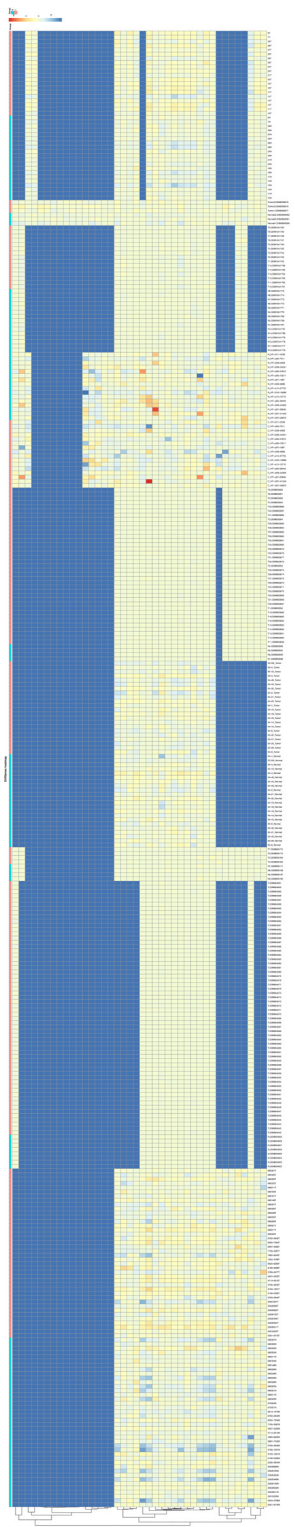

Figure 1 Heatmap of DEMs integrated from five miRNA microarrays of GC using the RRA method. DEMs, differentially expressed miRNAs; GC, gastric cancer; RRA, robust rank aggregation.

constructed to determine the vital regulatory roles of the miRNAs on their target genes, biological processes, and pathways, respectively. *Figure 2A*,*B* illustrate the interaction of miRNAs and target genes enriched in significant cancer-associated pathways. The target gene numbers of each miRNA ranged from 1 to 18, and notably, ROCK1, PPP1CB, PTEN, ARHGEF12, MAP3K3, CCDC6, RAP1B, E2F5, and KRAS were targeted by more than one miRNA. Additionally, as shown in *Figure 2C*, all fourteen miRNAs were involved in cancer pathways; among them, eleven miRNAs were associated with the adhesion pathway, and twelve miRNAs participated in the mTOR signaling pathway, which implied these miRNAs were closely associated with GC proliferation and invasion.

Validation of common DEMs in GC tissues

To confirm whether the expression levels of fifteen common DEMs from five microarrays were consistent with the independent cohort of stomach adenocarcinoma data from the TCGA database, the GC dataset of miRNAs from GC tissue and adjacent normal tissue of GC patients were acquired from an online website (https://xenabrowser.net). Compared with the DEMs in GC from the TCGA cohort, it was revealed that three upregulated miRNAs (miR-455-3p, miR-135b-5p, and let-7a-3p) and four downregulated miRNAs (miR-195-5p, miR-204-5p, miR-149-5p, and miR-143-3p) from the microarrays had the same expression direction as those from the TCGA cohort (*Figure 3*). In contrast, other miRNAs (miR-218-5p, miR-548au-3p, miR-193b-3p, and miR-148a-3p) were not consistent with the microarray results. Among them, miR-218-5p was inversely expressed and miR-548au-3p, miR-193b-3p, and miR-148a-3p were undetectable in GC samples. Moreover, the three miRNAs (miR-1-3p, miR-199a-5p, and miR-3910) with an indefinite expression direction in the microarrays were undetectable in the TCGA GC dataset. Finally, seven miRNAs were included in a further investigation.

Common DEMs were associated with the malignant behavior of GC

The association between the expression of seven common DEMs and the clinicopathological parameters are illustrated in Tables S2 and S3. In the upregulated common DEMs, compared with the low expression groups, the high expression of miR-455-3p was closely associated with

Table 2 Rank scores of DEMs from the five microarray profiling

Name	Score	P value	Corrected P value	Up/down	Count
hsa-miR-195-5p	0.001577824	0.002190703	0.975997555	Down	–5/+3
hsa-miR-455-3p	0.00270464	0.004715457	0.975997555	Up	–1/+6
hsa-miR-135b-5p	0.000634852	0.005811592	0.975997555	Up	–3/+6
hsa-miR-148a-3p	0.001315335	0.006934263	0.975997555	Down	–6/+2
hsa-miR-204-5p	4.74E-05	0.026336541	1	Down	–7/+2
hsa-let-7a-3p	0.030963273	0.030415077	1	Up	–1/+5
hsa-miR-218-5p	0.006761493	0.039897372	1	Up	–2/+6
hsa-miR-149-5p	0.000562649	0.040691271	1	Down	–4/+3
hsa-miR-1-3p	0.005634377	0.042473656	1	Down/up	–4/+4
hsa-miR-199a-5p	0.033672208	0.044637584	1	Down/up	–4/+4
hsa-miR-143-3p	0.035413603	0.047791055	1	Down	–5/+3
hsa-miR-3910	0.030963273	0.048588366	1	Down/up	–1/+1
hsa-miR-548au-3p	0.06183182	0.049078584	1	Up	1
hsa-miR-193b-3p	0.043176309	0.049635999	1	Down	–6/+2

DEMs, differentially expressed miRNAs.

the intestinal type of gastric adenocarcinoma (120/55 *vs.* 98/113, P<0.001) and histological grade G1–G2 (97/49 *vs.* 122/119, P=0.002); miR-135b-5p expression was correlated with male sex (82/50 *vs.* 192/63, P=0.007), intestinal type adenocarcinoma (143/33 *vs.* 131/80, P<0.001), histological grade G1–G2 (119/27 *vs.* 155/86, P<0.001), and diseases located in the proximal anatomical region of the cardia (83/15 *vs.* 94/44 *vs.* 92/51 *vs.* 5/3, P=0.005). However, the overexpression of let-7a-3p was related to pathological T1–2 classification (74/24 *vs.*147/142, P<0.001), pathological stage I–II (117/61 *vs.* 104/105, P=0.002), and survival (145/91 *vs.* 76/75, P=0.03).

In the four downregulated common DEMs, we found that a low expression of miR-195-5p (104/72 *vs.* 101/110, P=0.028) and miR-204-5p (79/79 *vs.* 64/129, P=0.001) was only associated with the intestinal type of gastric adenocarcinoma. In addition, miR-149-5p expression was correlated with GC in the distal stomach (0/98 *vs.* 58/80 *vs.* 51/92 *vs.* 1/7), and miR-135b-5p was associated with the intestinal type of gastric adenocarcinoma (140/36 *vs.* 142/69, P=0.007) and histological grade G1–G2 (118/28 *vs.* 164/77, P=0.006).

Logistic regression analysis of the upregulated miRNAs demonstrated that miR-455-3p was only associated with the intestinal type of gastric adenocarcinoma (OR: 0.43, 95% CI: 0.27–0.68, P=0.001) (Table S3). Moreover, miR-135b-5p was still associated with male sex (OR: 1.86, 95% CI: 1.15–3.03, P=0.01), the intestinal type of gastric adenocarcinoma (OR: 0.41, 95% CI: 0.24–0.69, P=0.001), and GC in the proximal region of the stomach (OR: 0.64, 95% CI: 0.48–0.86, P=0.003), while let-7a-3p was related to pathological T1–2 classification (OR: 0.42, 95% CI: 0.24–0.76, P=0.003). Of the downregulated miRNAs, logistic regression analysis showed that miR-195-5p (OR: 1.69, 95% CI: 1.08–2.64, P=0.02), miR-204-5p (OR: 2.12, 95% CI: 1.38–3.55, P=0.001), and miR-143-3p (OR: 1.68, 95% CI: 1.02–2.77, P=0.04) were associated with the diffuse type of gastric adenocarcinoma (Table S4). In addition, miR-195-5p also had a significant relationship with distal GC (OR: 1.35, 95% CI: 1.05–1.75, P=0.02).

Diagnostic performance of common DEMs

According to the clinicopathological parameters of common DEMs, although the signatures were infrequently associated with early pathological stage and TNM stage, the diagnostic

Table 3 The functional annotation (GO and KEGG) of common DEMs

Characteristic	GO-term	Description	Count in gene set	False discovery rate
Biological process	GO:0044237	Cellular metabolic process	391 of 8,797	1.56e-11
	GO:0043170	Macromolecule metabolic process	346 of 7,453	9.98e-12
	GO:0034613	Cellular protein localization	80 of 1,367	8.20e-05
Molecular function	GO:0004672	Protein kinase activity	50 of 635	1.08e-05
	GO:1901363	Heterocyclic compound binding	244 of 5,305	6.35e-06
	GO:0016740	Transferase activity	125 of 2,250	6.35e-06
Cellular component	GO:0044424	Intracellular part	577 of 13,996	5.19e-21
	GO:0005829	Cytosol	262 of 4,958	1.05e-13
	GO:0005634	Nucleus	332 of 6,892	1.05e-13
	GO:0005737	Cytoplasm	470 of 11,238	3.48e-12
KEGG pathways	hsa05206	MicroRNAs in cancer	18 of 149	0.0011
	hsa05226	Gastric cancer	13 of 147	0.0176
	hsa05225	Hepatocellular carcinoma	17 of 163	0.0027
	hsa05210	Colorectal cancer	12 of 85	0.0027
	hsa05224	Breast cancer	14 of 147	0.0082
	hsa04350	TGF-beta signaling pathway	12 of 83	0.0025
	hsa04150	mTOR signaling pathway	16 of 148	0.0027
	hsa04151	PI3K-Akt signaling pathway	24 of 348	0.0105
	hsa04310	Wnt signaling pathway	12 of 143	0.0269
	hsa04390	Hippo signaling pathway	15 of 152	0.0064
	hsa04068	FoxO signaling pathway	13 of 130	0.0082
	hsa05205	Proteoglycans in cancer	20 of 195	0.0015

GO, gene ontology; KEGG, Kyoto encyclopedia of genes and genomes; DEMs, differentially expressed miRNAs; mTOR Mammalian Target of Rapamycin; TGF, transforming growth factor; PI3K, Phosphoinositide 3-kinase.

performance of these DEMs was analyzed by a ROC curve to avoid the aforementioned results error (*Figure 4*). Based on the different combination styles of miRNAs, we divided them into three groups (single miRNA, two-miRNA, and three-miRNA signatures) of upregulated miRNAs and four groups (single miRNA, two-miRNA, three-miRNA, and four-miRNA signatures) of downregulated miRNAs to detect the diagnostic performance for GC.

As illustrated in *Table 4*, 89% of the area under the curve (AUC) of miR-135b-5p indicated a higher diagnostic accuracy compared with miR-455-3p and let-7a-3p in the upregulated miRNA groups. The combination of miR-135b-5p with miR-455-3p or let-7a-3p acquired 89% and 92% of the AUC for the diagnosis of early GC, respectively. When the three miRNAs were combined for GC diagnosis, the accuracy reached 91%. As shown in *Table 4*, miR-204-5p and miR-143-3p functioned as a single molecule for GC diagnosis with 79% of the AUC individually, and demonstrated better diagnostic accuracy than miR-149-5p (70%). However, when both of these miRNAs were combined to diagnose early GC, the AUC was increased to 93%. Furthermore, upon the addition of miR-149-5p, the AUC of ROC reached 94%. In addition, the AUC of the four-miRNA signatures was also 94%, which represented a

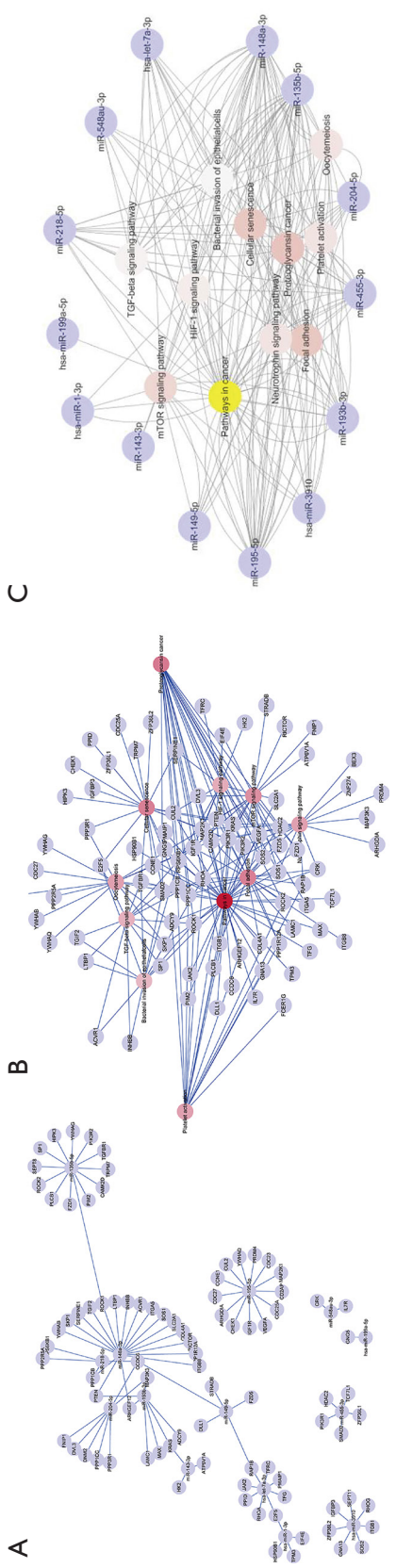

Figure 2 Common miRNA putative target genes and GO/Pathway enrichment analysis. The network analysis of DEMs between target genes (A), pathway (B), and biological process (C) of GC tumorigenesis with ClueGO software. The red and yellow nodes represent the biological pathway. The nodes of the light blue colors around represent the genes and miRNAs involved in modules and the lines represent the interaction between two nodes. The more forward ranking is represented by a redder color. GO, gene ontology; DEMs, differentially expressed miRNAs; GC, gastric cancer.

good diagnostic accuracy for GC.

Identification of the common DEMs of GC for prognosis

The prognostic value of the seven miRNAs was assessed because of dysregulation in their expression, which was correlated with the malignant behavior of GC. As shown in *Figure 5*, the log-rank test and Kaplan-Meier analysis were performed to assess the overall survival (OS) associated with the seven miRNAs. We found that upregulated miR-135b-5p was negatively and significantly associated with the OS of GC patients, whereas the interaction between aberrant miRNA expression and the OS of GC patients was not observed in the remaining six miRNAs. As illustrated in *Figure 6*, the association of progression-free survival (PFS) with the seven miRNAs was also examined to further evaluate their survival value for GC patients. However, no significant difference was observed between the seven miRNAs and PFS as determined by the log-rank test.

Next, univariate and multivariate analyses were performed to assess the risk factors of OS, and included the clinical characteristics of patients including age at diagnosis, gender, TNM stage, pathological stage, histologic type, histologic grade, anatomical subdivision, and dysregulated expression of miRNAs. The detailed results are shown in Tables S5 and S6. In the univariate analysis, we found that TNM stage, pathological stage, and histologic grade were associated with OS. After adjusting for other confounding factors, the multivariate analysis showed that age at diagnosis, pathological metastasis, and histological grade were independent predictors for poor OS. Notably, the aberrant expression of the seven miRNAs was not predominantly linked to the poor prognosis of GC patients, which corresponded to the results of the analysis of OS.

Discussion

In this study, using the RRA method, we attempted to integrate nine publicly available miRNA expression profiling datasets from 190 GC samples and 190 noncancerous tissue samples (normal gastric mucosa and adjacent normal tissue). The RRA method is an algorithm that specifically detects the overlapping variables and compares and integrates similar data from eligible studies (25). After the application of this method, all the results trended toward being robust and accurate. Finally, the integrated analysis showed that a panel of

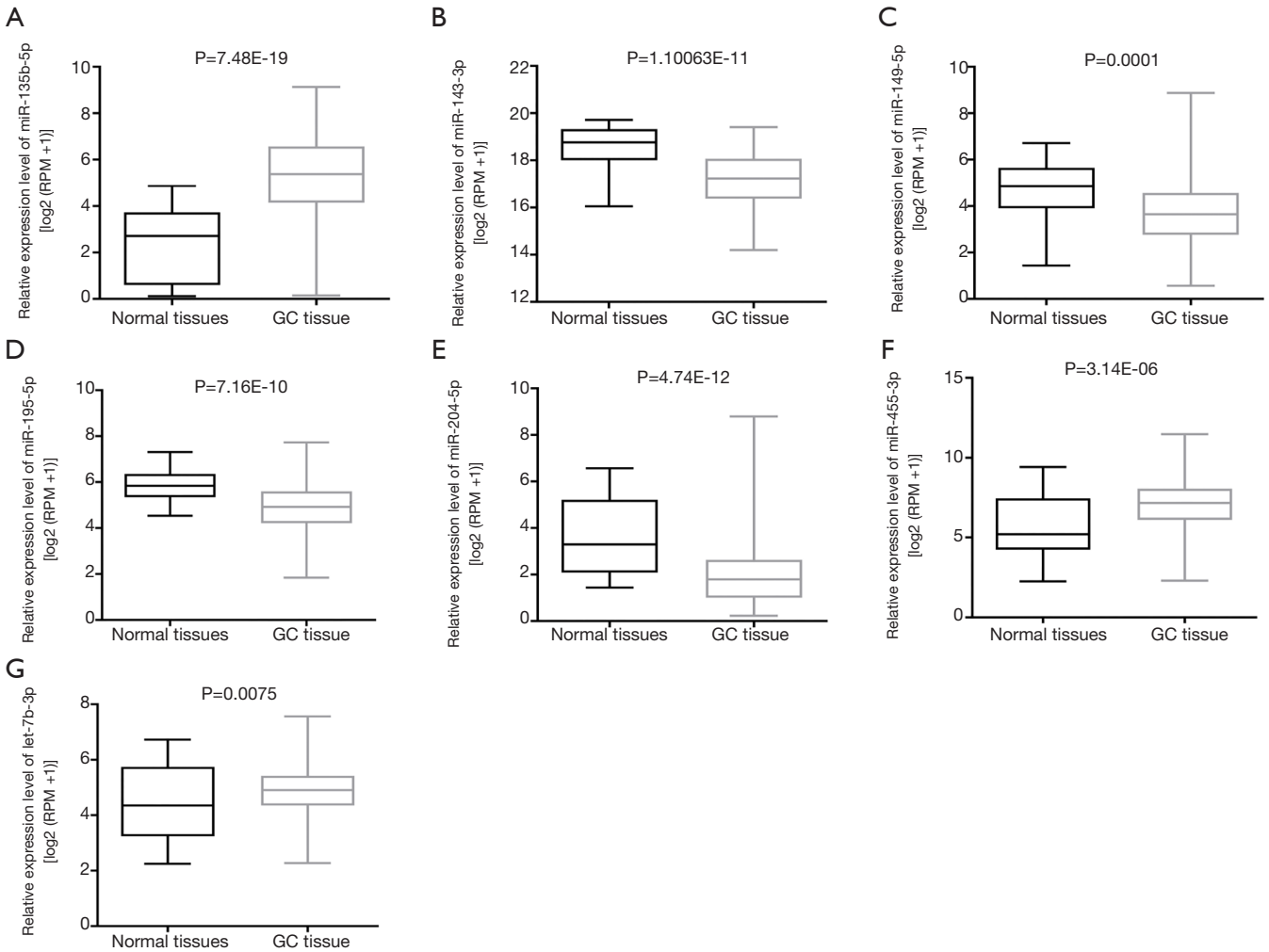

Figure 3 The differential expression validation of common DEMs with TCGA datasets. (A) miR-135b-5p, (B) miR-143-3p, (C) miR-149-5p, (D) miR-195-5p, (E) miR-204-5p, (F) miR-455-3p, (G) let-7b-3p. DEMs, differentially expressed miRNAs; TCGA, The Cancer Genome Atlas.

fourteen ectopically expressed miRNAs was obtained in GC and noncancerous tissues, which included five upregulated miRNAs, six downregulated miRNAs, and three miRNAs with an indefinite direction of regulation. From the enrichment analysis and path-net analyses of the target genes of the fourteen miRNAs, it was shown that the cancer pathways involved in the tumorigenesis of many cancers, such as GC, and many tumor signaling pathways play a core role in tumor development. The GC cohort from the TCGA database was applied to identify the miRNA expression as the validation set, and seven miRNAs (miR-455-3p, miR-135b-5p, let-7a-3p, miR-195-5p, miR-204-5p, miR-149-5p, and miR-143-3p) were selected for further analysis. Shao *et al.* (26) reported that miR-135b-5p promoted the proliferation, invasion, and cell cycle progression of GC cells. Yang *et al.* (27) found that let-7a-3p was correlated with the differentiation stage of GC. Moreover, they demonstrated that RAB40C was a target of let-7a and played a pivotal role in the tumorigenesis of GC. Low expression of miR-195-5p (28) inhibits the migration and invasiveness of GC cells by downregulating bFGF. Chen *et al.* (29) demonstrated that miR-204-5p was distinctly expressed at low levels in GC and that the Linc01234/miR-204-5p/CBFB axis played

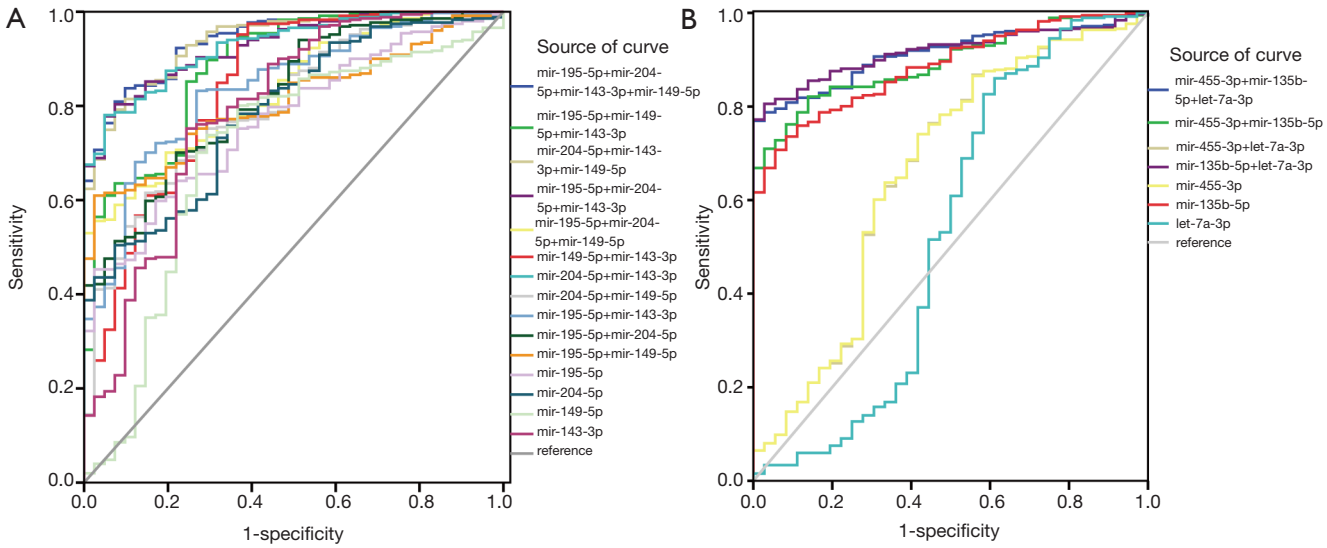

Figure 4 The diagnostic performance of seven DEMs with ROC curve. (A) The diagnostic analysis of downregulated miRNAs including single miRNA (mir-143-3p, miR-149-5p, miR-204-5p, and miR-195-5p), two-miRNAs, three-miRNAs, and four-miRNAs. (B) The diagnostic analysis of upregulated miRNAs including single miRNA (miR-455-3p, miR-135b-5p, and let-7a-3p), two-miRNAs, and three-miRNAs. DEMs, differentially expressed miRNAs; ROC, receiver operating characteristic.

an important role in GC tumorigenesis. Zhang *et al.* (30) discovered that knockdown of miR-149-5p promoted the proliferation, migration, and invasiveness of GC cells, and they also illustrated that the circNRIP1/miR-149-5p/AKT axis functioned in GC development. Wu *et al.* (31) observed that miR-143-3p was associated with advanced malignant tumors and lymph node metastasis. However, only a few studies have focused on the function of miR-455 in GC.

Logistic regression analysis indicated that the upregulated miRNAs (miR-455-3p and miR-135b-5p) were relevant to the intestinal type of gastric adenocarcinoma. However, the downregulated miRNAs (miR-195-5p, miR-204-5p, and miR-143-3p) were related to the diffuse type of gastric adenocarcinoma. Moreover, miR-135b-5p had a close correlation with proximal region GC, while miR-195-5p was correlated with distal region GC. Additionally, let-7a-3p was observed to be relevant to the pathological T1–2 classification. However, whether these signatures can be regarded as indicators to predict the histological type and anatomical region of GC requires additional studies to support our findings, and therefore, further research needs to be performed to explore the potential mechanism.

The evaluation of the diagnostic accuracy of these miRNAs indicated that miR-135b-5p, miR-204-5p, and miR-143-3p possessed a higher diagnostic value as single molecules. Notably, two-miRNA, three-miRNA, and four-miRNA signatures that were used to diagnose GC were superior to single signatures. However, the diagnostic result was based on GC tissue, and the diagnostic ability of these miRNAs in the blood still needs to be confirmed. Tsujiura *et al.* (32) determined that the ratio of miR-106a/let-7a in the blood had a better diagnostic accuracy (87.9% of the AUC), but for other miRNAs, not enough evidence was produced to support our results. Additionally, the prognostic analyses of OS and RFS and the Cox regression analysis implied that the poor prognosis of GC patients was not dependent on the aberrant expression of the seven miRNAs, but rather, the miRNAs were associated with age, metastasis, and histological grade. Current relevant studies have primarily focused on the mechanism of GC tumorigenesis, and no studies have reported on the prognostic value of these miRNAs; therefore, it is possible that there was no direct linkage between the seven miRNAs and their prognostic value.

In summary, in this study we performed an integrated analysis using the RRA method to merge five miRNA

Table 4 The diagnostic analyses of upregulated and downregulated miRNAs

MiRNAs	AUC	95% CI	P value
Upregulated			
miR-455-3p	0.66	0.55–0.77	0.001
miR-135b-5p	0.89	0.85–0.93	<0.001
let-7a-3p	0.53	0.40–0.66	0.57
miR-455-3p+miR-135b-5p	0.89	0.86–0.93	<0.001
miR-455-3p+let-7a-3p	0.66	0.55–0.77	0.002
miR-135b-5p+let-7a-3p	0.92	0.89–0.95	<0.001
miR-455-3p+miR-135b-5p+let-7a-3p	0.91	0.88–0.94	<0.001
Downregulated			
miR-195-5p	0.77	0.71–0.83	<0.001
miR-204-5p	0.79	0.72–0.85	<0.001
miR-149-5p	0.70	0.61–0.80	<0.001
miR-143-3p	0.79	0.71–0.87	<0.001
miR-195-5p+miR-204-5p	0.82	0.76–0.88	<0.001
miR-195-5p+miR-149-5p	0.81	0.76–0.86	<0.001
miR-195-5p+miR-143-3p	0.84	0.78–0.90	<0.001
miR-204-5p+miR-149-5p	0.80	0.73–0.86	<0.001
miR-204-5p+miR-143-3p	0.93	0.90–0.96	<0.001
miR-149-5p+miR-143-3p	0.84	0.76–0.91	<0.001
miR-195-5p+miR-204-5p+miR-149-5p	0.83	0.78–0.89	<0.001
miR-195-5p+miR-204-5p+miR-143-3p	0.93	0.90–0.96	<0.001
miR-204-5p+miR-149-5p+miR-143-3p	0.94	0.91–0.97	<0.001
miR-195-5p+miR-149-5p+miR-143-3p	0.89	0.84–0.94	<0.001
miR-195-5p+miR-204-5p+miR-149-5p+miR-143-3p	0.94	0.91–0.97	<0.001

AUC, Area Under Curve; CI, confidence interval.

microarray datasets from the GEO database. Fourteen signatures with aberrant miRNA expression were included. Additionally, the GC datasets from the TCGA database were selected as the validation set for the differential expressions, and finally, seven miRNAs consistent with the TCGA and GEO databases were subjected to diagnostic and prognostic analyses. These seven miRNAs were regarded as promising candidate diagnostic biomarkers in GC tissues, but whether the results obtained in tissues are consistent with those found in peripheral blood requires further validation. Furthermore, the poor prognosis of GC patients was not closely correlated with the ectopic expression of the seven miRNAs. However, we believe that these miRNAs might be potential biomarkers for the early diagnosis of GC, and future studies should be performed using a large, prospective, and multicenter cohort.

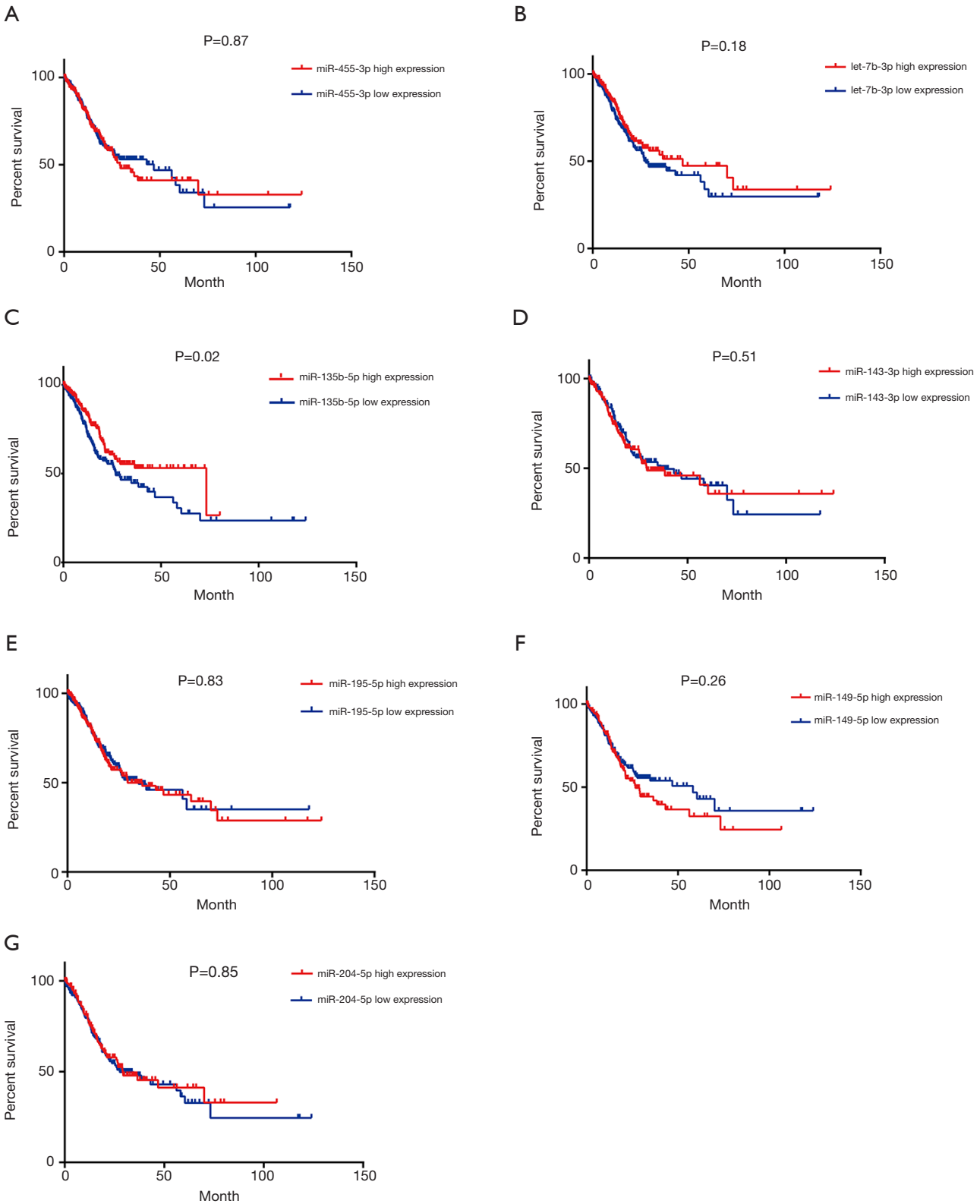

Figure 5 Identification of the seven DEMs of GC patients with overall survival. (A) miR-455-3p, (B) let-7b-3p, (C) miR-135b-5p, (D) miR-143-3p, (E) miR-195-5p, (F) miR-149-5p, (G) miR-204-5p. DEMs, differentially expressed miRNAs; GC, gastric cancer.

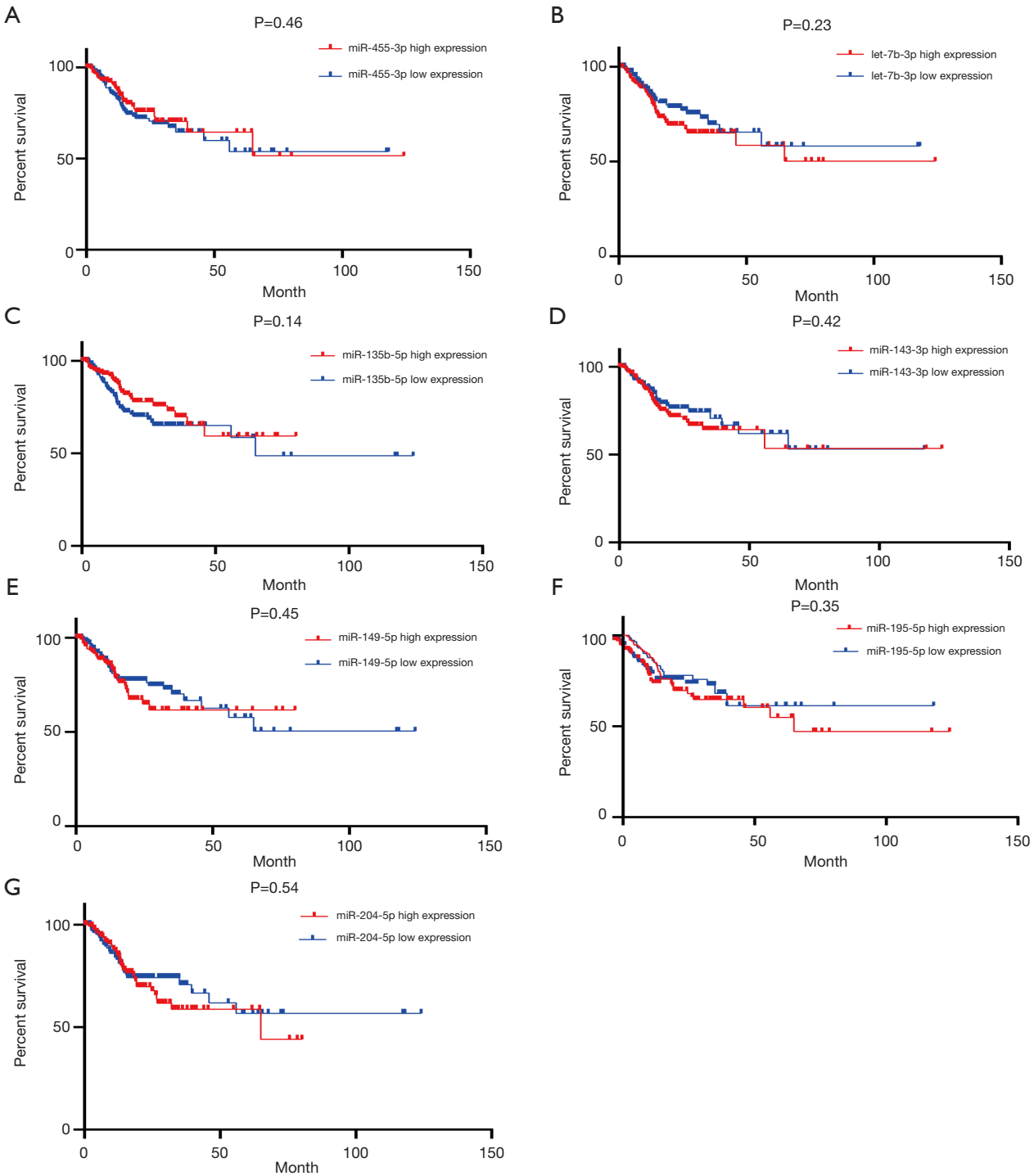

Figure 6 Identification of the seven DEMs of GC patients with progression-free survival. (A) miR-455-3p, (B) let-7b-3p, (C) miR-135b-5p, (D) miR-143-3p, (E) miR-195-5p, (F) miR-149-5p, (G) miR-204-5p. DEMs, differentially expressed miRNAs; GC, gastric cancer.

Acknowledgments

Funding: This study was supported by the National Natural Science Foundation of China (71964021), and the National Natural Science Foundation of Gansu Province (17JR5RA260), the Science and Technology Development Guiding Plan Project of Lanzhou (2019-ZD-39).

Footnote

Reporting Checklist: The authors have completed the MDAR checklist. Available at http://dx.doi.org/10.21037/atm-21-1631

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-21-1631). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

References

1. Fitzmaurice C, Dicker D, Pain A, et al. The global burden of cancer 2013. JAMA Oncol 2015;1:505-27. Erratum in: JAMA Oncol 2015;1:690.
2. Nagini S. Carcinoma of the stomach: A review of epidemiology, pathogenesis, molecular genetics and chemoprevention. World J Gastrointest Oncol 2012;4:156-69.
3. Fielding JW, Ellis DJ, Jones BG, et al. Natural history of "early" gastric cancer: results of a 10-year regional survey. Br Med J 1980;281:965-7.
4. Mabula JB, Mchembe MD, Koy M, et al. Gastric cancer at a university teaching hospital in northwestern Tanzania: a retrospective review of 232 cases. World J Surg Oncol 2012;10:257.
5. Wang C, Leng X, Zhang Y, et al. Transcriptome-wide analysis of dynamic variations in regulation modes of grapevine microRNAs on their target genes during grapevine development. Plant Mol Biol 2014;84:269-85.
6. Liu HS, Xiao HS. MicroRNAs as potential biomarkers for gastric cancer. World J Gastroenterol 2014;20:12007-17.
7. Iorio MV, Croce CM. MicroRNAs in cancer: small molecules with a huge impact. J Clin Oncol 2009;27:5848.
8. Yan Z, Xiong Y, Xu W, et al. Identification of recurrence-related genes by integrating microRNA and gene expression profiling of gastric cancer. Int J Oncol 2012;41:2166-74.
9. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature 2014;513:202.
10. Steponaitiene R, Kupcinskas J, Langner C, et al. Epigenetic silencing of miR-137 is a frequent event in gastric carcinogenesis. Mol Carcinog 2016;55:376-86.
11. Mishra PJ, Bertino JR. MicroRNA polymorphisms: the future of pharmacogenomics, molecular epidemiology and individualized medicine. Pharmacogenomics 2009;10:399-416.
12. Liu C, Rennie WA, Carmack CS, et al. Effects of genetic variations on microRNA: target interactions. Nucleic Acids Res 2014;42:9543-52.
13. Kolde R, Laur S, Adler P, et al. Robust rank aggregation for gene list integration and meta-analysis. Bioinformatics 2012;28:573-80.
14. Võsa U, Kolde R, Vilo J, et al. Comprehensive meta-analysis of microRNA expression using a robust rank aggregation approach. In: RNA Mapping. Springer. 361-73.
15. Wang J, Gao P, Song Y, et al. Prognostic value of gastric cancer-associated gene signatures: Evidence based on a meta-analysis using integrated bioinformatics methods. J Cell Mol Med 2018;22:5743-7.
16. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47
17. Bindea G, Mlecnik B, Hackl H, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 2009;25:1091-3.
18. Li M, Li D, Tang Y, et al. CytoCluster: A cytoscape plugin for cluster analysis and visualization of biological networks. Int J Mol Sci 2017;18:1880.
19. Oh HK, Tan AL, Das K, et al. Genomic loss of miR-486

regulates tumor progression and the OLFM4 antiapoptotic factor in gastric cancer. Clin Cancer Res 2011;17:2657-67.
20. Lim JY, Yoon SO, Seol SY, et al. Overexpression of miR-196b and HOXA10 characterize a poor-prognosis gastric cancer subtype. World J Gastroenterol 2013;19:7078-88.
21. Carvalho J, van Grieken NC, Pereira PM, et al. Lack of microRNA-101 causes E-cadherin functional deregulation through EZH2 up-regulation in intestinal gastric cancer. J Pathol 2012;228:31-44.
22. Chang H, Kim N, Park JH, et al. Different microRNA expression levels in gastric cancer depending on Helicobacter pylori infection. Gut Liver 2015;9:188-96.
23. Zhang X, Ni Z, Duan Z, et al. Overexpression of E2F mRNAs associated with gastric cancer progression identified by the transcription factor and miRNA co-regulatory network analysis. PLoS One 2015;10:e0116979.
24. Shah MY, Ferrajoli A, Sood AK, et al. microRNA therapeutics in cancer—an emerging concept. EBioMedicine 2016;12:34-42.
25. Shrestha S, Hsu SD, Huang WY, et al. A systematic review of microRNA expression profiling studies in human gastric cancer. Cancer Med 2014;3:878-88.
26. Shao L, Chen Z, Soutto M, et al. Helicobacter pylori-induced miR-135b-5p promotes cisplatin resistance in gastric cancer. FASEB J 2019;33:264-74.
27. Yang Q, Jie Z, Cao H, et al. Low-level expression of let-7a in gastric cancer and its involvement in tumorigenesis by targeting RAB40C. Carcinogenesis 2011;32:713-22.
28. Wang J, Li L, Jiang M, et al. MicroRNA-195 inhibits human gastric cancer by directly targeting basic fibroblast growth factor. Clin Transl Oncol 2017;19:1320-8.
29. Chen X, Chen Z, Yu S, et al. Long noncoding RNA LINC01234 functions as a competing endogenous RNA to regulate CBFB expression by sponging miR-204-5p in gastric cancer. Clin Cancer Res 2018;24:2002-14.
30. Zhang X, Wang S, Wang H, et al. Circular RNA circNRIP1 acts as a microRNA-149-5p sponge to promote gastric cancer progression via the AKT1/mTOR pathway. Mol Cancer 2019;18:20.
31. Wu X-L, Cheng B, Li P-Y, et al. MicroRNA-143 suppresses gastric cancer cell growth and induces apoptosis by targeting COX-2. World J Gastroenterol 2013;19:7758.
32. Tsujiura M, Ichikawa D, Komatsu S, et al. Circulating microRNAs in plasma of patients with gastric cancers. Br J Cancer 2010;102:1174-9.

(English Language Editor: D. Fitzgerald)

Cite this article as: Liu X, Pu K, Wang Y, Chen Y, Zhou Y. Gastric cancer-associated microRNA expression signatures: integrated bioinformatics analysis, validation, and clinical significance. Ann Transl Med 2021;9(9):797. doi: 10.21037/atm-21-1631

