

Article

Not peer-reviewed version

# Identification and validation of blood miRNAs as diagnostic markers of gastric cancer

Qingyu Wang , Zhouhan Xu , Xuesi Yang , Ruiqi Zhao , Xueling Li , Shujie Li , Xize Li , [Jiping Zeng](#) \*

Posted Date: 8 August 2023

doi: 10.20944/preprints202308.0543.v1

Keywords: Gastric cancer; Blood-based miRNAs; Differentially expressed genes; Biomarkers; Bioinformatics



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Article

# Identification and Validation of Blood miRNAs as Diagnostic Markers of Gastric Cancer

Qingyu Wang <sup>1,†</sup>, Zhouhan Xu <sup>2,†</sup>, Xuesi Yang <sup>2</sup>, Ruiqi Zhao <sup>2</sup>, Xueling Li <sup>1</sup>, Shujie Li <sup>1</sup>, Xize Li <sup>1</sup> and Jiping Zeng <sup>1,\*</sup>

<sup>1</sup> Department of Biochemistry and Molecular Biology; wqingy7788@163.com (Q.W.); 17853509565@163.com (X.L.); lxlyj33@163.com (S.L.); a1164224335@163.com (X.L.)

<sup>2</sup> Clinical Five-Year Program, School of Basic Medical Sciences, Cheeloo College of Medicine, Shandong University, Jinan, Shandong, China 250012; xzh013581@163.com (Z.X.); 15064051695@163.com (X.Y.); 863295184@qq.com (R.Z.)

\* Correspondence: zengjp@sdu.edu.cn; Tel.: 0086-531-88382092.

† The authors contributed equally to this work.

‡ Current address: School of Health and Life Sciences, University of Health and Rehabilitation Sciences, Qingdao, Shandong, China 266113.

**Abstract:** In order to provide a new direction for the diagnosis and treatment of gastric cancer, the differentially expressed blood-based miRNAs in gastric carcinoma were obtained and analyzed from the datasets GSE113486, GSE112264 and GSE113740 by the SangerBox, the miRNA - target genes prediction, GO and KEGG functional enrichment analyses with DAVID 6.8, and the STRING database with Cytoscape, respectively. The correlation between the key miRNAs expression and the prognosis and diagnostic value of gastric cancer patients were detected by survival prognosis and ROC curve. QRT-PCR was used to detect the possible key miRNAs expression in human blood species. The dataset GSE108307 which shows the miRNAs affected by *H. pylori* infection was analyzed by GEO2R. The expression level of CagA and the clinically used gastric cancer markers were detected in human blood species for contrast. Five key blood-based miRNAs were determined, including hsa-miR-124-3p, hsa-miR-125a-3p, hsa-miR-29b-3p, hsa-miR-4276, and hsa-miR-575. The detection in human blood species as well as the cross analysis with *H. pylori* infection and clinically used markers expression suggested the importance and efficacy of the specific miRNAs in the early diagnosis and monitoring of gastric carcinoma. The study identified potential novel blood miRNA biomarkers of gastric cancer by bioinformatics and human blood species detection, which may give a new possibility of the important biomarkers for the diagnosis and prognosis of gastric cancer.

**Keywords:** gastric cancer; blood-based miRNAs; differentially expressed genes; biomarkers; bioinformatics

## 1. Introduction

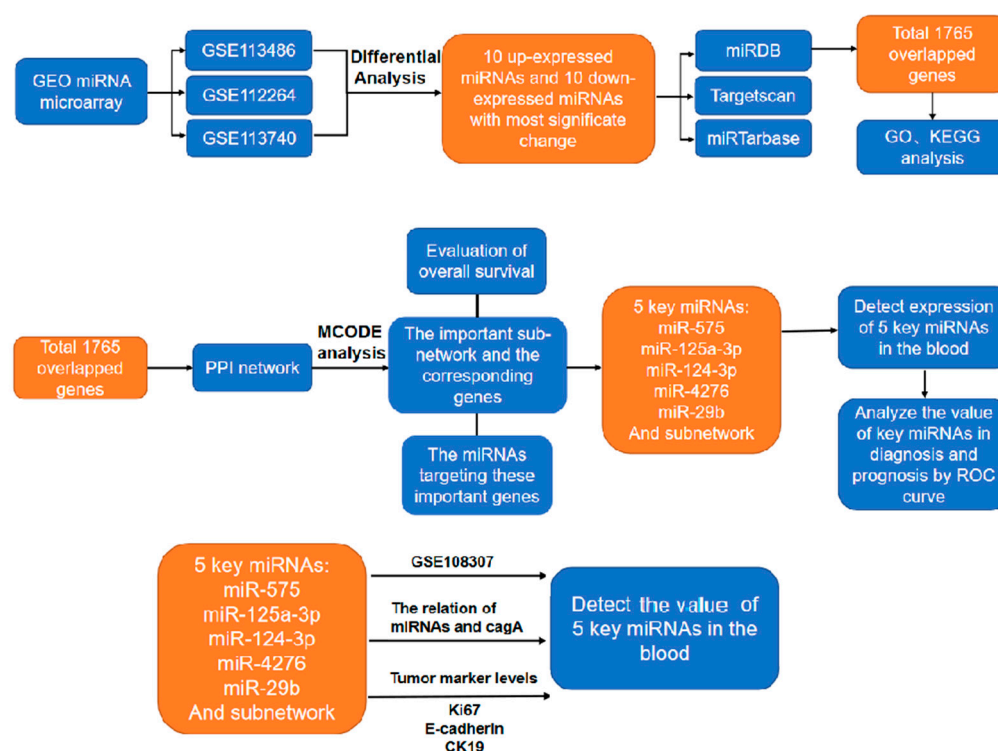
Gastric cancer is a kind of malignant tumors which organizes from the gastric mucosal epithelium. There are more than 1,000,000 new patients suffering from gastric cancer every year. At the same time, about 769,000 people lose their lives in 2021 due to it [1, 2]. Gastric cancer has been bringing enormous mental and financial burden to individuals and the whole society. There is progress on the early diagnosis and treatment of gastric cancer, however it is needed of breakthrough on the new molecular mechanism of gastric carcinoma to achieve efficient, reliable and measurable biomarkers.

MiRNA is a class of non-coding single-stranded RNA molecules of approximately 22 nucleotides in length encoded by endogenous gene [3]. It plays an important role in the regulation of the post-transcriptional gene expression in plants and animals [4]. MiRNA acts through base pairing with complementary sequences within the mRNA molecule, resulting in silencing of these mRNA molecules. This disorder of negative expression regulation is closely related to the most of

tumorigenesis. In the past few years, manipulation of miRNA clusters expression is attempted for diagnosis, treatment and prognosis in cancers [5].

As a common test material, blood is easy to obtain, easy to detect and minimally invasive. At the same time, since the tumor microenvironment has a great influence on the carcinogenesis, invasion and metastasis, the components in the blood might also affect the tumorigenesis. Therefore, the study of miRNAs in blood may have great effect on the prediction and diagnose of tumors [6]. After searching for the most cutting-edge data, we found that the research on differentially expressed miRNAs in the blood of patients with gastric cancer is not perfect. Therefore, the study on differentially expressed miRNAs in the gastric cancer patients' blood is promising to promote the discovery of new tumor markers and determine new research directions.

In recent years, with the establishment of tumors databases, sequencing technology and bioinformatics analysis have been widely used in the research of cancer at the molecular level [7]. Drug designed based on bioinformatics analysis is also emerging [8]. This provides new ideas and tools for us to find out differentially expressed miRNAs from patients' blood samples and explore their functional pathways that might affect gastric carcinoma. Therefore, this study searched for the corresponding dataset through the GEO database [9]. Next the bioinformatics tools were used to analyze the obtained data for screening out key miRNAs which were predicted biological mechanism through functional enrichment and protein network construction. The results were detected in human blood species with analysis of patient prognosis and diagnostic value. These determined crucial blood miRNAs and the corresponding target genes might become biomarkers for early screening and diagnosis of gastric cancer. *Helicobacter pylori* (*H. pylori*) is recognized Class I carcinogen by the World Health Organization/International Agency for Research on Cancer (WHO/IARC). According to epidemiological statistics, *H. pylori* and its key virulence factor Cytotoxin-associated gene A (*CagA*) are the main environmental factors in gastric carcinogenesis. Thus the prediction results was checked by the relevance analysis with *H. pylori* infection and *CagA* expression by the utilization of clinically used markers. The flow chart of the methods used in this study is shown in Figure 1.



**Figure 1.** Flow chart of the methods utilized in the present study.

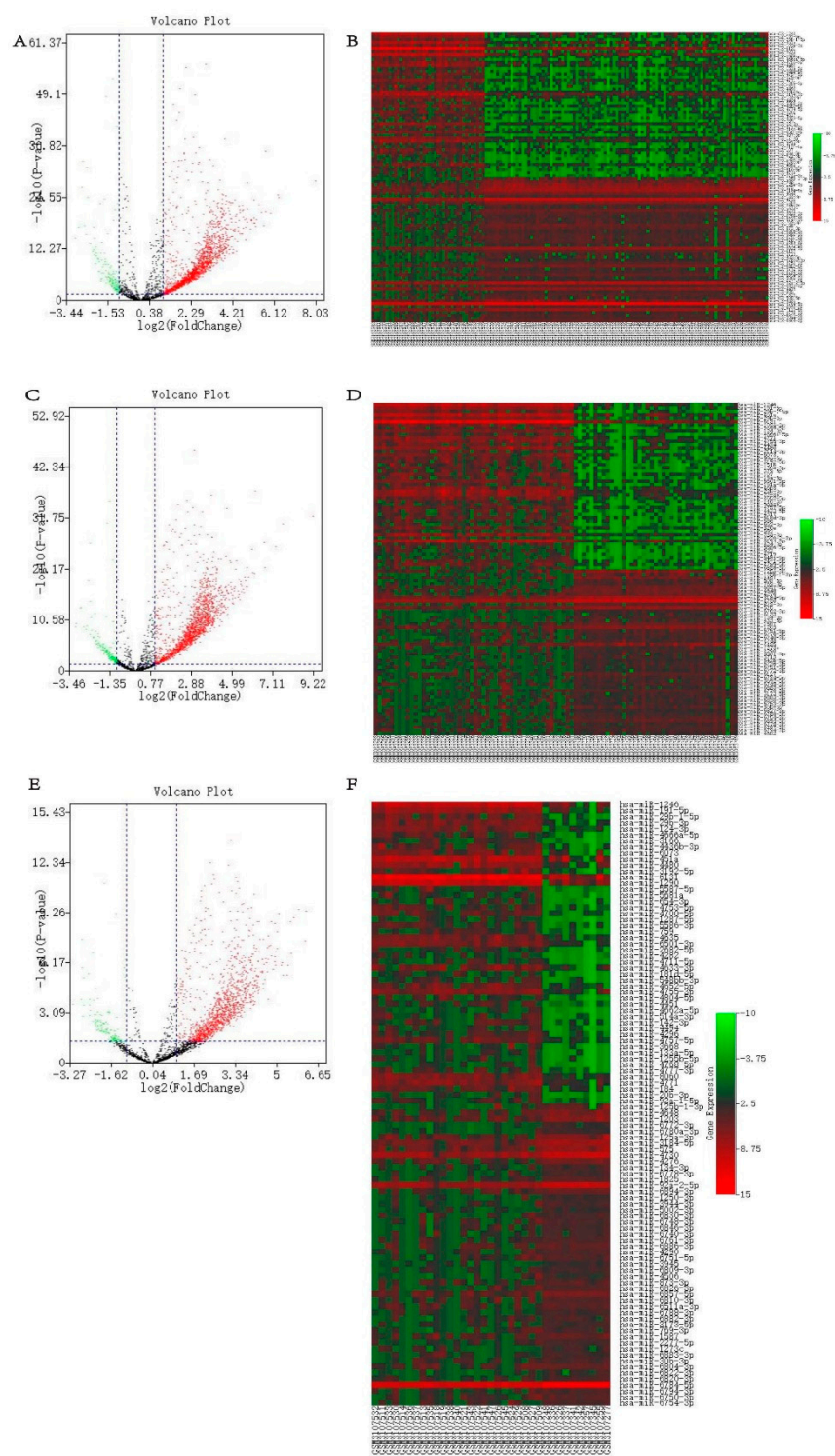
## 2. Results

### 2.1. The differentially expressed blood miRNAs in gastric carcinoma

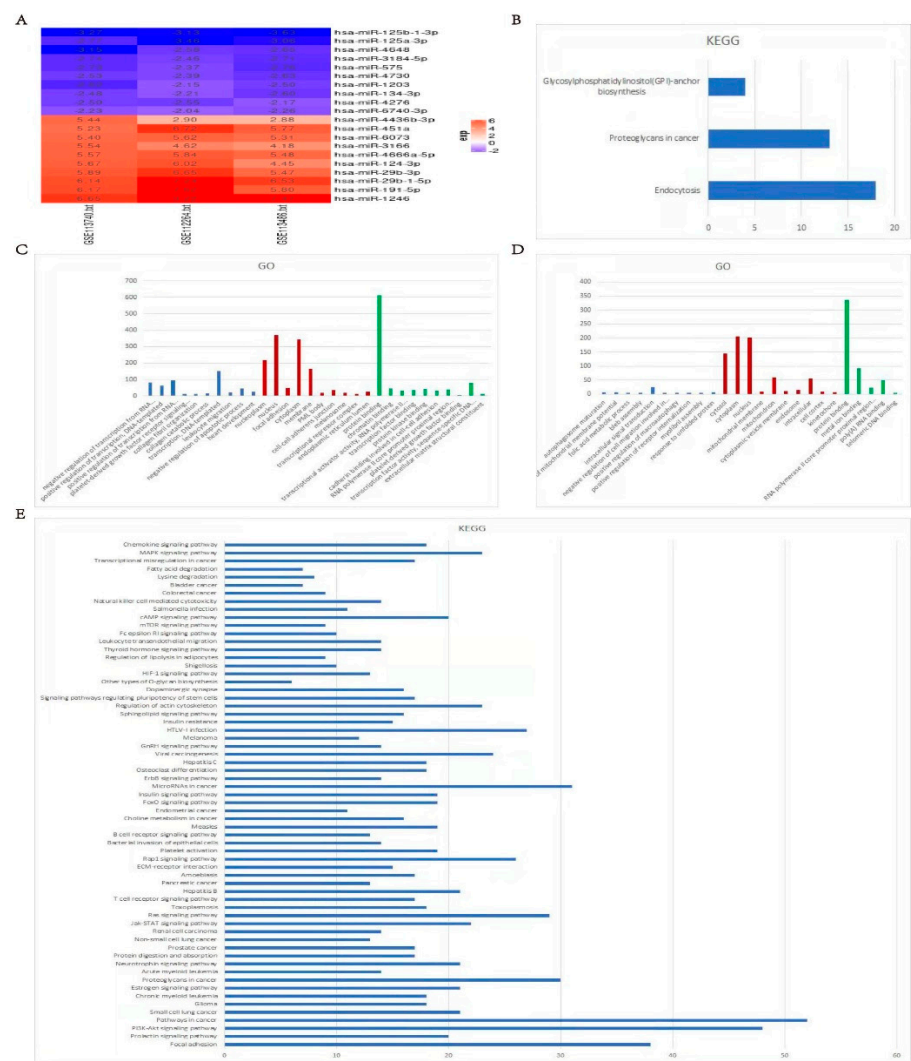
The SangerBox was used to analyze the GSE113486, GSE112264 and GSE113740 data to screen and obtain the abnormally expressed miRNAs in the blood of gastric cancer patients. The results detected that in the chip GSE113486, compared with the normal controls, there were 2086 miRNAs with statistically significant difference in the group of gastric cancer. Therein to raised miRNAs were 1784 and decreased miRNAs were 302 (Figure 2A). The most remarkably up-regulated 50 miRNAs and down-regulated 50 miRNAs were shown in the Figure 2B. In the chip GSE112264, there were a total of 2086 abnormal miRNAs in the blood of patients with gastric cancer that are statistically significant, with an increase of 1831 and a decrease of 255 (Figure 2C). The 50 miRNAs with the most significant up-regulation and down-regulation were shown in the Figure 2D. In the chip GSE113740, compared with the normal blood of the control, there were 1482 abnormal miRNAs in the blood of gastric cancer patients, which were statistically significant, with an increase of 1340 and a decrease of 142 (Figure 2E), of which the most significant up-regulation and down-regulation and the 50 miRNAs were shown in the Figure 2F.

### 2.2. The prediction of target genes directly silenced by the most differentially expressed miRNAs and the enrichment results of GO and KEGG

The RRA algorithm was used to find the top 10 up-regulated miRNAs and the top 10 down-regulated miRNAs from the three chip datasets (Figure 3A). With the combination and comparison of the predictions from Targetscan, Mirtarbase and miRDB, 691 over-expressed genes and 1074 down-regulated ones were obtained as the possible targets of the selected 20 miRNAs. GO and KEGG enrichment analysis on differentially expressed genes by DAVID received the molecular functions, cellular components, corresponding biological processes and signaling pathways of the target genes. KEGG enrichment results showed that Endocytosis, Proteoglycans in cancer and Glycosylphosphatidylinositol (GPI)-anchor biosynthesis were the main signaling pathways involved in overexpressed target genes-related mechanism (Figure 3B). Up-regulated genes (targets of down-regulated miRNAs) GO enrichment results suggested that the main biological processes included autophagosomes maturation, regulation of mitochondrial membrane potential and folic acid metabolic process. The corresponding main molecular functions included protein binding, metal ion binding and RNA polymerase II core promoter proximal region sequence-specific DNA binding. The cellular components included cytosol, cytoplasm and nucleus (Figure 3C). On the contrary, for the down-regulated genes (targets of up-regulated miRNAs), the main involved biological processes included regulation of transcription from RNA polymerase II promoter, platelet-derived growth factor receptor signaling pathway and DNA-templated positive regulation of transcription. The main molecular functions included protein binding, chromatin binding, transcriptional activator activity, and RNA polymerase II core promoter proximal region sequence-specific binding. The cellular components included nucleoplasm, nucleus and focal adhesion (Figure 3D). The signaling pathways involved in down-expressed genes regulation were Focal adhesion, Prolactin signaling pathway, PI3K-Akt signaling pathway and Focal adhesion (Figure 3E).



**Figure 2.** The differentially expressed blood miRNAs in gastric carcinoma. (A) The differentially expressed miRNAs in the bloods of gastric cancer patients compared with the normal controls from the dataset GSE113486. (B) The most significantly up-regulated 50 miRNAs and down-regulated 50 miRNAs in GSE113486. (C) The differentially expressed miRNAs in the bloods of gastric cancer patients compared with the normal controls from the dataset GSE112264. (D) The most significantly up-regulated 50 miRNAs and down-regulated 50 miRNAs in GSE112264. (E) The differentially expressed miRNAs in the bloods of gastric cancer patients compared with the normal controls from the dataset GSE113740. (F) The most significantly up-regulated 50 miRNAs and down-regulated 50 miRNAs in GSE113740.

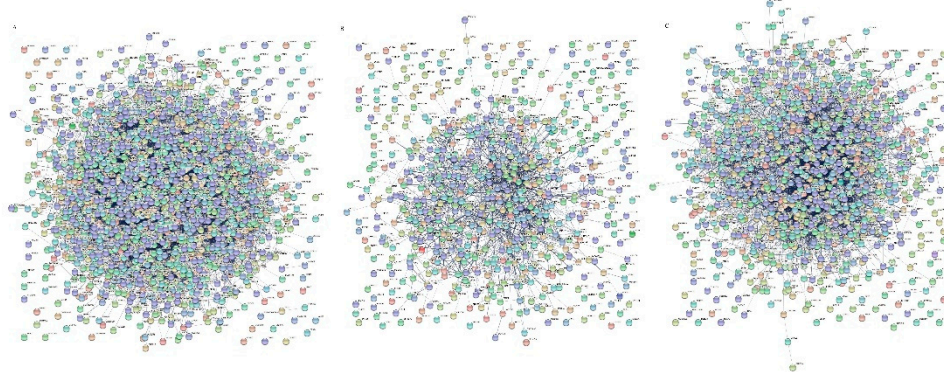


**Figure 3.** The GO and KEGG enrichment results of target genes that were directly regulated by the key blood miRNAs. (A) The highest-scoring up-regulated and down-regulated miRNAs obtained by the RRA algorithm. (B) The main biological processes, molecular functions and cellular components of down-regulated miRNAs and the corresponding up-regulated target genes. (C) KEGG- signaling pathway enrichment results of down-regulated miRNAs and the corresponding up-regulated target genes. (D) The main biological processes, molecular functions and cellular components of up-regulated miRNAs and the corresponding down-regulated target genes. (E) KEGG- signaling pathway enrichment results of up-regulated miRNAs and the corresponding down-regulated target genes.

2.3. The analysis of the proteins interaction network and selection of the key blood miRNAs with the correlant target genes

All the 20 most differently expressed blood miRNAs and the corresponding target genes were analyzed in the STRING database (Figure 4A–C). The results of the PPI network were imported into Cytoscape software for analysis. Then a total of 38 nodes were obtained which indicated that the 20 most differently expressed blood miRNAs were closely related to each other. And they were involved in multiple signaling pathways which might play important roles in the tumorigenesis. At the same time, the number of nodes in the up-regulated genes-encoded proteins network was 650. The number of edges was 1571. The average node degree was 4.83. The average local clustering coefficient was

0.337. And the enrichment P value was less than 0.0123. The number of nodes in the down-regulated genes-encoded proteins network was 1007. The number of edges was 5171. The average node degree was 10.3. The average local clustering coefficient was 0.308. And the enrichment P value is less than  $1.0 \times 10^{-16}$ . The results showed the data was reliable and can be used for further analysis.

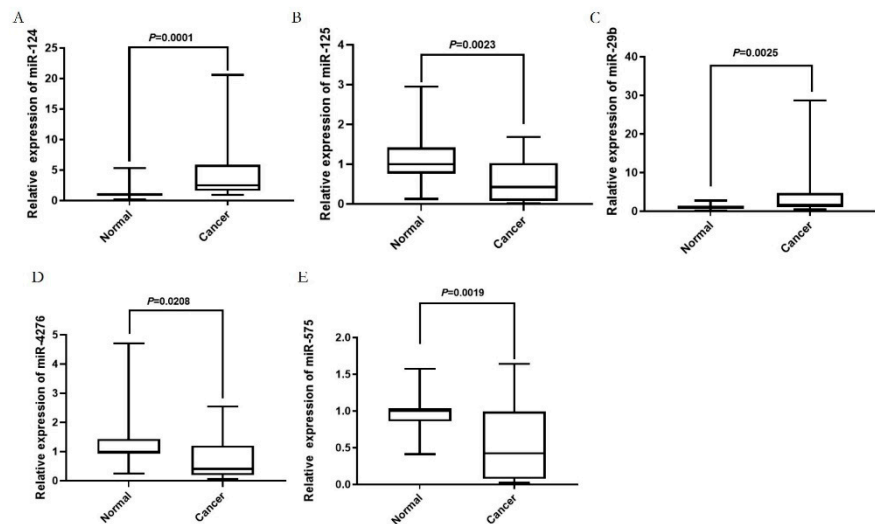


**Figure 4.** The analysis of the proteins interaction network by STRING database. (A) The network of all the 20 most differently expressed blood miRNAs and the corresponding target genes-coded proteins. (B) The PPI of the up-regulated genes-encoded proteins. (C) The PPI of the down-regulated genes-encoded proteins.

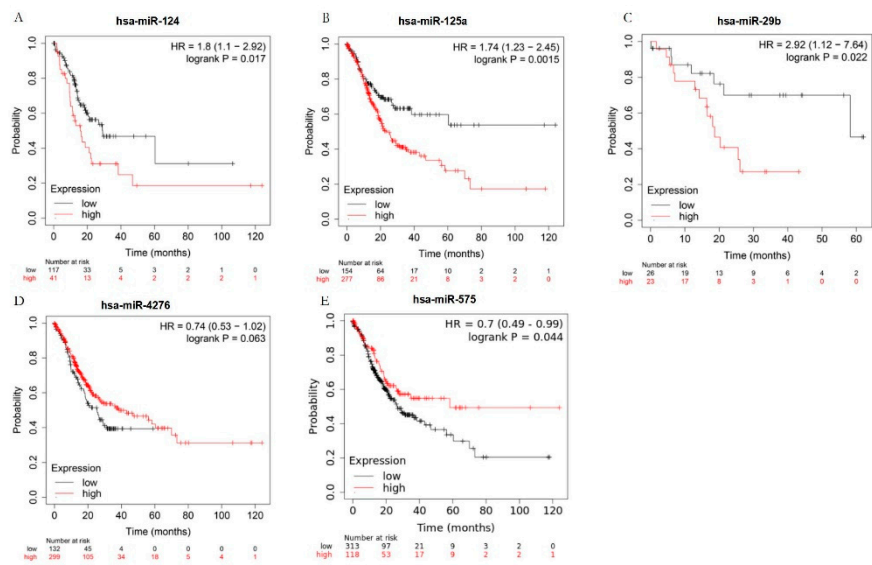
The number of occurrences of the blood miRNAs on the key nodes was obtained by Cytoscape software, which determined the importance of the corresponding miRNAs (Table S1). The results suggested the key blood miRNAs involved in the up-regulated ones were hsa-miR-124-3p and hsa-miR-29b-3p, whereas in down-regulated ones were hsa-miR-125a-3p, hsa-miR-4276 and hsa-miR-575. They could be the possible key miRNAs for the early diagnosis and treatment of gastric carcinoma. The sub-network genes obtained in the PPI network, which might be the targets of the key miRNAs, were involved in the survival analysis. The results showed there were 31 blood miRNAs-targeted genes which were associated with gastric cancer (Table S2).

#### *2.4. The key blood miRNAs with the correlant target genes were detected by the examination of the expression levels in human blood species and the survival curve analysis.*

The basic information of the patients who provided the blood samples and the correlation with the expression of key miRNAs was shown in Table S3. The results of qRT-PCR which determined the expression of selected key miRNAs in human blood species were shown in Figure 5. The survival curves which suggested the correlation of selected miRNAs expression with the patient prognosis were shown in Figure 6. The survival curves of the corresponding target genes with the patient prognosis were shown in Figures S1 and S2. From the results, it was preliminarily found that the significant up-regulation of hsa-miR-124-3p and hsa-miR-29b-3p as well as down-regulation of hsa-miR-575 was closely related to the prognosis of gastric cancer patients ( $P < 0.05$ ) in human blood species. Whereas of hsa-miR-125a-3p and hsa-miR-4276, the expression in human blood species was consistent with bioinformatics prediction results and inconsistent with the prognosis of gastric cancer patients. At the same time, although the prognosis analysis suggested the expression of hsa-miR-4276 was insignificant with the patient survival rate ( $P > 0.05$ ), the lower expression of hsa-miR-4276 had a tendency to reduce the survival rate of patients. The prognosis analysis also identified that the expression of most target genes was associated with the patient survival possibility. The expression of the key miRNAs had nothing to do with the patients' age and gender ( $P > 0.05$ ). All these results determined that the prediction of bioinformatics could be useful for diagnosis, which might not be completely consistent with the patients' conditions. It will be necessary to use more patients blood samples for testing and verification.



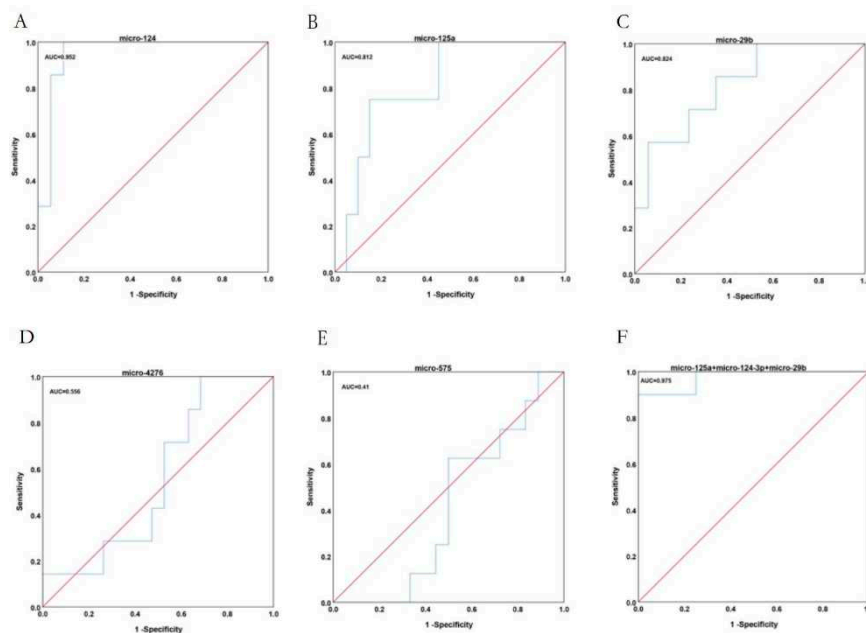
**Figure 5.** The expression of key miRNAs in human blood species by qRT-PCR detection. (A) The expression of hsa-miR-124-3p increased in the blood samples of gastric cancer patients ( $P < 0.01$  vs. normal). (B) The blood expression of hsa-miR-125a-3p decreased in the blood samples of gastric cancer patients ( $P < 0.01$  vs. normal). (C) The expression of hsa-miR-29b-3p increased in the blood samples of gastric cancer patients ( $P < 0.05$  vs. normal). (D) The expression of hsa-miR-4276 decreased in the blood samples of gastric cancer patients ( $P < 0.01$  vs. normal). (E) The expression of hsa-miR-575 decreased in the blood samples of gastric cancer patients ( $P < 0.01$  vs. normal).



**Figure 6.** Survival analysis results of the key miRNAs. (A) Kapler-Meier Plotter analysis showed that the high expression of hsa-miR-124-3p, which was predicted as an oncogene, significantly reduced the survival rate of patients ( $P < 0.05$ ). (B) The high expression of hsa-miR-125a-3p significantly reduced the survival rate of patients ( $P < 0.01$ ). (C) The high expression of hsa-miR-29b-3p significantly reduced the survival rate of patients ( $P < 0.05$ ). (D) The high expression of hsa-miR-4276 with tumor suppressive effect improved the survival rate of patients to a certain extent, but there was no significant difference ( $P=0.063$ ). (E) The high expression of hsa-miR-575 improved the survival rate of patients significantly ( $P < 0.05$ ).

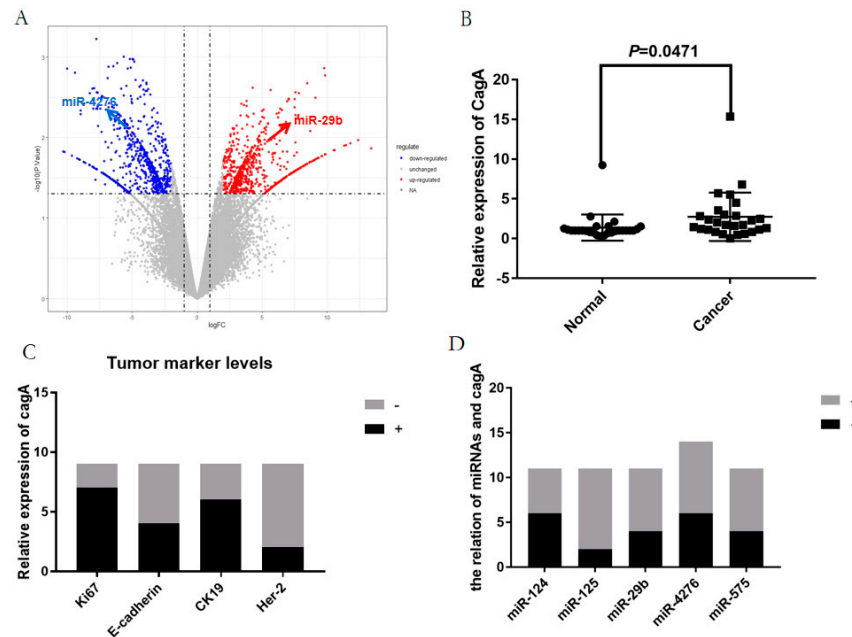
### 2.5. The diagnostic value of key blood miRNAs in patients with GC.

In order to further clarify the value of key blood miRNAs in the diagnosis of gastric carcinoma, ROC analysis was conducted in an expression-dependent manner. The results showed that miR-124-3p, miR-125-3p and miR-29b-3p could have the potential application in diagnosis, with AUC values of 0.952, 0.812 and 0.824, respectively. The ROC value of the group to combine miR-124-3p, miR-125-3p and miR-29b-3p could get to 0.975. Although the ROC value of hsa-miR-4276 and hsa-miR-575 were relatively low to 0.556 and 0.41, the results showed the model composed of the key miRNAs was excellent in distinguishing patients with and without GC (Figure 7).



**Figure 7.** The key miRNAs could have diagnostic value in gastric cancer determined by ROC analysis. (A) The value of hsa-miR-124-3p area under the ROC curve (AUC) was 0.952,  $P < 0.001$ . (B) The value of hsa-miR-125a-3p AUC was 0.812,  $P < 0.001$ . (C) The value of miR-29b-3p AUC was 0.824,  $P < 0.001$ . (D) The value of hsa-miR-4276 AUC was 0.556,  $P > 0.05$ . (E) The value of hsa-miR-575 AUC was 0.41,  $P > 0.05$ . (F) The value of (hsa-miR-124-3p + hsa-miR-125a-3p + miR-29b-3p) AUC was 0.975,  $P < 0.001$ .

The model was further verified in GSE108307 which involved the miRNAs induced by *H. pylori* infection. The cross verification showed the up-regulated hsa-miR-29b-3p and down-regulated hsa-miR-4276 could be found in the dataset (Figure 8A). CagA expression was examined in human blood species using qRT-PCR. The result suggested that CagA was overexpressed in gastric cancer patients group compared with the normal (Figure 8B). The clinically used markers for pathological diagnosis of gastric cancer and the key miRNAs were detected in the samples of gastric cancer patients which was divided into CagA-positive and CagA-negative groups. Compared with the clinically used markers, the model composed of the key miRNAs gave accurate diagnosis which demonstrated the correlation between CagA-positive group and the expression of the key miRNAs (Figure 8C,D). In conclusion, the five key miRNAs totally showed the diagnostic value in patients with GC by blood samples check.



**Figure 8.** The expression of key miRNAs in CagA-positive human blood species by qRT-PCR detection. (A) The differentially expressed miRNAs in *H. pylori*-infected gastric cancer patients compared with the normal controls from the dataset GSE108307, including the key hsa-miR-29b-3p and hsa-miR-4276. (B) The expression of CagA, the key virulence factor of *H. pylori*, was remarkably overexpressed in human gastric cancer blood species compared with the normal controls ( $P < 0.05$ ). (C) The correlation between CagA expression and clinically used markers expression levels including Ki67, E-Cadherin, CK19 and Her-2. (D) The correlation between CagA expression and key miRNAs expression levels.

### 3. Discussion

In the research, GSE113486, GSE112264 and GSE113740 selected from the GEO database were analyzed with the bioinformatics tools to get the key blood miRNAs in gastric carcinoma. RRA analysis were performed to get the ranking of the corresponding miRNAs. Then the 20 most differentially expressed blood miRNAs (10 up-regulated and the other 10 down-regulated) became the aims to predict the target genes and the relative biological activities. There were 691 up-regulated differentially expressed genes and 1074 down-regulated genes which might be directly silenced by the 20 blood miRNAs. GO and KEGG enrichment analysis were proceeded on the differentially expressed genes by DAVID. The results suggested the biological processes and the signaling pathways in which the differentially expressed genes were mainly involved in this system. On this basis, analysis of the target genes – encoded proteins with PPI network suggested five crucial blood miRNAs which were hsa-miR-124-3p, hsa-miR-125a-3p, hsa-miR-29b-3p, hsa-miR-4276 and hsa-miR-575. The results in human blood detection and survival curve analysis confirmed the important roles of the key blood miRNAs in gastric carcinoma. ROC analysis and the cross validation with *H. pylori*-induced miRNAs as well as the comparative study by the clinical used markers all suggested the key miRNAs could be useful in gastric cancer diagnosis.

MiRNAs have been identified as a key regulator in complex biological processes, including the pathogenesis of cancer [10]. Human cancers express the characteristics such as sustained proliferation signals, activation of invasion and metastasis, angiogenesis, escape from immune destruction and so on [11]. MiRNAs can be involved in the every feature and affect multiple signal pathways [12]. Blood miRNAs are becoming one of the most useful biomarkers in the diagnosis and treatment of cancers [13-16]. MiRNAs have the specific expression profile in different cancer cells and tissues which can enter the circulation. Circulating free miRNA can be detected in blood, plasma and other body fluids. Mature miRNAs are very stable in body fluids and have high specificity in different cancer states, which determines miRNAs a potential non-invasive tumor marker. There have been some research

on the blood miRNAs in gastric carcinoma, such as miR-210, miR-1, miR-20a, miR-34a, miR-423-5p, and so on [17, 18]. In this research, with the bioinformatics tools, five new blood miRNAs, which were hsa-miR-124-3p, hsa-miR-125a-3p, hsa-miR-29b-3p, hsa-miR-4276 and hsa-miR-575, might be the potential targets in gastric carcinoma diagnosis and treatment. The key miRNAs and the relative target genes were involved in the biological activities which are closely correlant to tumorigenesis. For example, platelet-derived growth factor receptor signaling pathway and transcription RNA polymerase II both play important roles in gastric carcinoma and development [19, 20]. Focal adhesion and PI3K-Akt signaling pathway are closely related to the carcinogenesis [21, 22]. These results detected the importance of the selected blood miRNAs in gastric carcinoma.

By tracing the upstream miRNAs of the target genes, the results of Cytoscape which screened out key node genes from the PPI network showed the appearance frequency of hsa-miR-124-3p was the highest in the key miRNAs. However, there was no clear conclusion about the relationship between hsa-miR-124-3p and tumorigenesis in the miRbase database. A few research are beginning to hint the potential role of miR-124-3p in tumorigenesis [23, 24]. Both the examination in human blood species and the analysis of the survival curve suggested the correlation between hsa-miR-124-3p and the gastric carcinogenesis. The further ROC analysis also suggested the importance of hsa-miR-124-3p in gastric cancer diagnosis. In the predicted target genes of hsa-miR-124-3p, a number of genes had a higher correlation with the prognosis of gastric cancer, such as ANXA5 and CAV1. ANXA5 is a member of the Annexin family and is involved in the tumorigenesis and development of a variety of cancers [25], which plays the role on gastric cancer by regulating the ERK signal pathway [26]. The low expression of CAV1 means decreased expression of E-cadherin, cell morphology changes and an increase in the migration ability of gastric cancer cells [27]. Other miRNAs, such as miR-6792-3p, can also be involved in gastric carcinogenesis by their inhibitory effect on the target CAV1 [28]. So, it can be speculated that hsa-miR-124-3p may facilitate gastric carcinoma by negatively regulating multiple downstream target genes in one network. At present, there are no studies on the correlation between *H. pylori* infection and hsa-miR-124-3p. This study suggested the up-regulation of hsa-miR-124-3p in CagA-positive gastric cancer patients' blood samples. All the results detected the value of hsa-miR-124-3p in gastric carcinogenesis and diagnosis.

It has been explored that hsa-miR-125a-3p could influence immunity and carcinogenesis by regulating tumor-associated signal transduction pathway, such as the Hippo pathway [29]. In addition, hsa-miR-125a-3p is related to the tumorigenesis of prostate cancer, colon cancer, and so on [30]. However, the research of hsa-miR-125a-3p on gastric carcinoma is still few [31]. In this research, the abnormal expression of hsa-miR-125a-3p was identified in the blood species of gastric cancer patients. At the same time, the downstream target of hsa-miR-125a-3p was predicted as PR domain zinc finger protein 1 (PRDM1) gene. PRDM1 is the protein involved in regulating the differentiation of B cells and T cells, which plays an important role in immunosuppression [32]. Previous reports have shown that PRDM1 is associated with many kinds of cancers, such as adrenocortical cancer, colon cancer, acute myeloid leukemia, brain cancer and lung cancer [33]. The expression of PRDM1 in non-germinal center B cell-like (non-GCB) patients was also associated with a worse prognosis [34]. In this research, PRDM1 also showed its correlation with the survival possibility of gastric cancer patients. There are no studies on the correlation between *H. pylori* infection and hsa-miR-125a-3p / PRDM1 until now. In this study, CagA induced the inhibition of hsa-miR-125a-3p which was consistent with the down-regulation of hsa-miR-125a-3p in blood samples of GC patients. Therefore, it may have high probability that hsa-miR-125a-3p is participating in the gastric carcinogenesis by inhibiting the expression of PRDM1, which need in-depth research later.

Hsa-miR-29b-3p plays a key role in the tumorigenesis and metastatic of glioblastoma, colorectal cancer, breast cancer and so on [35-37]. Hsa-miR-29b-3p can also regulate the expression of VEGFA in pancreatic ductal adenocarcinoma [38]. The research also showed hsa-miR-29b-3p might act as one of the important circulating microRNAs and have clinical value in endometrial cancer Regulation [39]. In addition, as one of the important target genes of hsa-miR-29b-3p, activation of PER1 can inhibit the progression of pancreatic cancer [40]. The GC blood check displayed the correlation between circulating hsa-miR-29b-3p expression and gastric cancer prognosis and diagnostic.

Although there is few studies about the role of hsa-miR-29b-3p in *H. pylori* infection, the bioinformatics prediction suggested the importance of the miRNA and the negatively regulated target genes in this process.

Hsa-miR-4276 is involved in lung epithelial cells resistance on influenza A infection by inducing the expression of cytochrome c oxidase VIc [41]. There is few relevant research to prove that hsa-miR-4276 is related to tumorigenesis, also with its target genes RNF217 and IP6K1. On the other hand, among the key nodes obtained by Cytoscape, the number of downstream genes of hsa-miR-4276 was 24 times. This suggested that hsa-miR-4276 and its target genes might have a strong correlation with gastric cancer. The result from patient blood check and the prediction from *H. pylori* - induced miRNAs suggested this possibility. However, it need more in-depth exploration.

MSRB3 has been identified to be the key protein that can regulate the proliferation and migration of gastric cancer cells which might be an effective marker to predict gastric cancer peritoneal metastasis and poor prognosis [42, 43]. MSRB3 was predicted the target gene of hsa-miR-575 in this research. It was also detected in blood check that hsa-miR-575 was differentially expressed in gastric cancer and could be regulated by CagA. Hsa-miR-575 might has the trend correlated with the prognosis of gastric cancer patients, which need more research to determine. The role of hsa-miR-575 which can negatively regulate MSRB3 in gastric carcinogenesis can be one of the targets of future research.

In summary, all the research implied the accuracy and feasibility with the abnormal expression level of the key miRNAs in blood for gastric cancer prevention, diagnosis and treatment. Through bioinformatics analysis and check in human blood species, the model composed of hsa-miR-124-3p, hsa-miR-125a-3p, hsa-miR-29b-3p, hsa-miR-4276 and hsa-miR-575 could be involved in gastric carcinoma and used as circulating markers. The blood miRNAs may become new biomarkers for the early diagnosis of gastric cancer. More research is needed in the future to further confirm the mechanism and biological activities of the five key miRNAs in gastric carcinoma for better utilization.

## 4. Materials and Methods

### 4.1. Data selection

GEO (<https://www.ncbi.nlm.nih.gov/geo/>) is built by the National Center for Biotechnology Information (NCBI) which contains high-throughput gene expression data submitted by the scientific researchers around the world. The database was scanned with the keywords "gastric cancer", "blood", "human" and "miRNA". GSE113486, GSE112264 and GSE113740 were confirmed as the most suitable datasets. The datasets contain 115 cases of gastric cancer samples and 151 cases of normal samples. GSE108307 was used to show the miRNAs affected by *H. pylori* infection.

### 4.2. Preliminary screening of differentially expressed miRNAs and predicted target genes

SangerBox is one of the biometric analysis platforms based on R language (<http://sangerbox.com>). Using the GEO convenience converter in SangerBox, the sample files were carried on differential analysis by DECenter. Then its network tool library was used to execute RRA analysis. Thus 20 miRNAs which showed the most significant changes in the gastric cancer group compared with the normal were screened out as the research targets (10 up-regulated and 10 down-regulated, respectively). Targetscan (<http://www.targetscan.org/>), miRDB (<http://mirdb.org/>) and mirtarbase (<http://mirtarbase.mbc.nctu.edu.tw>) are the three websites for online prediction of the miRNA's possible target genes.

### 4.3. Functional enrichment of differentially expressed genes

DAVID6.8 (<https://david.ncifcrf.gov/>) is one of widely used bioinformatics databases that can provide systematic and comprehensive information for large-scale genes or proteins lists. In this research, it was used to proceed GO and KEGG enrichment analysis on differentially expressed genes and relative signal pathways that might play key roles in gastric carcinoma, which provided a rough

understanding of the relevant molecular mechanisms. The  $P < 0.05$  and gene counts  $\geq 2$  were rated to be statistically significant.

#### 4.4. Protein-protein interaction network (PPI network)

PPI network is composed of individual proteins through their interactions. The study of PPI network can help to understand the functional links among proteins and to obtain the core regulators. String (<https://string-db.org>) is one of the proteins interaction databases with the most exhaustive species and the most comprehensive interaction information. After obtaining the PPI network of 20 differentially expressed genes, the important sub-network and the corresponding genes were analyzed with the MCODE plug-in clustering function module of Cytoscape (version 3.6.1). The minimum score required for interaction is medium confidence (0.400).

#### 4.5. Selection and survival curve analysis of the key miRNAs and relative target genes

It was to cluster the sub-network nodes of the PPI network, and sort the most important genes which were negatively regulated by the corresponding key miRNAs. The occurrence number of a miRNA was used to decide whether it should be regarded as the key one in the research. ONCOLNC and GEPIA are the two tools that can interactively explore survival dependencies. In this study, the correlation between the key nodal genes and the survival rate of gastric cancer patients was analyzed with ONCOLNC and GEPIA. The survival curves of the key miRNAs and the target genes were examined by the Kaplan-Meier Plotter online tools (<https://kmplot.com/analysis/>) to determine the influence of their expression level on the prognosis of gastric cancer patients. The significance was defined as a  $P$  value of  $< 0.05$ .

#### 4.6. The human blood species and qRT-PCR determination

The human blood species were collected from Xiajin County People's Hospital, Dezhou, Shandong Province. They were divided into two groups: one control group ( $n=17$ ) which came from randomly selected normal volunteers, and the other experimental group ( $n=29$ ) which came from patients who were newly diagnosed with gastric cancer without any treatment. The gender, age, lymph node metastasis and gastric cancer marker expression were all derived from the pathological examination reports of patients. This research has been approved by the Ethics Committee of Shandong University School of Basic Medical Sciences.

Total RNA was extracted using Invitrogen TRIzol (Thermo Fisher SCIENTIFIC, Code: 15596018). For miRNA quantification, the synthesis of cDNA was performed using Evo M-MLV RT Kit for qPCR (Accurate Biology, Code: AG11707). The qPCR primers of miR-575, miR-125a-3p, miR-124-3p, miR-4276 and miR-29b-3p were designed and synthesized with the Bulge-Loop miRNA qRT-PCR Primer sets (RIBOBIO, Code: MQPSCM001). Real-time PCR was carried on with SYBR Green Premix Pro Taq HS qPCR Kit (Accurate Biology, Code: AG11701) on a Bio-Rad CFX-96 real-time system. U6 was detected as the internal control. CagA expression was determined with the primers forward: 5'-CAAGTCCGTGGGCATCATGT-3' and reverse: 5'-GAGGACACTCGGTCTCTAGC-3'. All qRT-PCR reactions were conducted in triplicates, and relative quantification was calculated by the  $2^{-\Delta\Delta Ct}$  method (95% confidence interval) with calibration to the corresponding control. The expression data were analyzed with a Student's  $t$ -test. The  $P < 0.05$  was considered significant difference in statistics.

#### 4.7. ROC analysis

The expression data of key miRNAs were used to construct the ROC curves by software SPSS 26.0 with area under the curve (AUC) values as the diagnosis value indicator. The  $P < 0.05$  was considered to be statistically significant.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Figure S1: Survival analysis results of the target genes regulated by the key hsa-miR-124-3p; Figure S2. Survival analysis results of the target genes regulated by the key miRNAs; Table S1. The most differentially expressed blood miRNAs in gastric carcinoma; Table S2. The key blood miRNAs-targeted

genes in gastric cancer; Table S3. The correlation between key miRNAs expression and the age and gender of gastric cancer patients providing blood species.

**Author Contributions:** Data curation, Qingyu Wang, Zhouhan Xu, Xuesi Yang, Ruiqi Zhao, Shujie Li and Xize Li; Funding acquisition, Jiping Zeng; Project administration, Jiping Zeng; Resources, Xueling Li; Supervision, Jiping Zeng; Writing - original draft, Qingyu Wang, Xize Li and Jiping Zeng; Writing - review and editing, Jiping Zeng. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Natural Science Foundation of China (No. 81871627) and Shandong University Student Innovation and Entrepreneurship Training Program (No. 2019325, S202110422208).

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of Shandong University School of Basic Medical Sciences (No.ECSBMSSDU2021-1-097 on March 3, 2021).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Acknowledgments:** Thanks to Fen Liu from Department of Pathogenic Microbiology, School of Basic Medical Sciences, Cheeloo College of Medicine, Shandong University, Yize Liang from Shandong University Qilu Hospital and Haocheng Wang from the 8-year Clinical Medicine program of Shanghai Jiao Tong University for their help in the clinical data analysis.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Xia C, Dong X, Li H, Cao M, Sun D, He S, Yang F, Yan X, Zhang S, Li N, Chen W. Cancer statistics in China and United States, 2022: profiles, trends, and determinants. *Chin Med J (Engl)*. 2022 Feb 9; 135(5):584-590.
2. Wang Q, Liu G, Hu C. Molecular Classification of Gastric Adenocarcinoma. *Gastroenterology Res*. 2019; 12(6):275-282.
3. Hrovatin K, Kunej T. Classification of miRNA-related sequence variations. *Epigenomics*. 2018; 10(4):463-481.
4. Fabian MR, Sonenberg N. The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. *Nat Struct Mol Biol*. 2012; 19(6):586-593.
5. Saliminejad K, Khorram Khorshid HR, Soleymani Fard S, Ghaffari SH. An overview of microRNAs: Biology, functions, therapeutics, and analysis methods. *J Cell Physiol*. 2019; 234(5):5451-5465.
6. Su F, Gao Z, Liu Y, Zhou G, Cui Y, Deng C, Liu Y, Zhang Y, Ma X, Wang Y, Guan L, Zhang Y, Liu B. Integrated Tissue and Blood miRNA Expression Profiles Identify Novel Biomarkers for Accurate Non-Invasive Diagnosis of Breast Cancer: Preliminary Results and Future Clinical Implications. *Genes (Basel)*. 2022; 13(11):1931.
7. Scionti F, Arbitrio M, Caracciolo D, Pensabene L, Tassone P, Tagliaferri P, Di Martino MT. Integration of DNA Microarray with Clinical and Genomic Data. *Methods Mol Biol*. 2022; 2401:239-248.
8. Berger MF, Mardis ER. The emerging clinical relevance of genomics in cancer medicine. *Nat Rev Clin Oncol*. 2018; 15(6):353-365.
9. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J*. 2014; 13:8-17.
10. Rupaimoole R, Slack FJ. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nat Rev Drug Discov*. 2017; 16(3):203-222.
11. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discov*. 2022; 12(1):31-46.
12. Van Roosbroeck K, Calin GA. Cancer Hallmarks and MicroRNAs: The Therapeutic Connection. *Adv Cancer Res*. 2017; 135:119-149.
13. Turchinovich A, Weiz L, Langheinz A, Burwinkel B. Characterization of extracellular circulating microRNA. *Nucleic Acids Res*. 2011; 39(16):7223-7233.
14. Schwarzenbach H, Nishida N, Calin GA, Pantel K. Clinical relevance of circulating cell-free microRNAs in cancer. *Nat Rev Clin Oncol*. 2014; 11(3):145-156.
15. Giannopoulou N, Constantinou C. Recent developments in diagnostic and prognostic biomarkers for colorectal cancer: a narrative review. *Oncology*. 2023 Jun 26.
16. Van den Broek D, Groen HJM. Screening approaches for lung cancer by blood-based biomarkers: Challenges and opportunities. *Tumour Biol*. 2023 Jun 30.
17. Aalami AH, Aalami F, Sahebkar A. Gastric Cancer and Circulating microRNAs: An Updated Systematic Review and Diagnostic Meta-Analysis. *Curr Med Chem*. 2023; 30(33):3798-3814.
18. Jelski W, Mroczko B. Molecular and Circulating Biomarkers of Gastric Cancer. *Int J Mol Sci*. 2022; 23(14):7588.

19. Huang F, Wang M, Yang T, Cai J, Zhang Q, Sun Z, Wu X, Zhang X, Zhu W, Qian H, Xu W. Gastric cancer-derived MSC-secreted PDGF-DD promotes gastric cancer progression. *J Cancer Res Clin Oncol*. 2014; 140(11):1835-48.
20. Chen FX, Smith ER, Shilatifard A. Born to run: control of transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol*. 2018; 19(7):464-478.
21. Zhang H, Schaefer A, Wang Y, Hodge RG, Blake DR, Diehl JN, Papageorge AG, Stachler MD, Liao J, Zhou J, Wu Z, Akarca FG, de Klerk LK, Derks S, Pierobon M, Hoadley KA, Wang TC, Church G, Wong KK, Petricoin EF, Cox AD, Lowy DR, Der CJ, Bass AJ. Gain-of-Function RHOA Mutations Promote Focal Adhesion Kinase Activation and Dependency in Diffuse Gastric Cancer. *Cancer Discov*. 2020; 10(2):288-30.
22. Hu M, Zhu S, Xiong S, Xue X, Zhou X. MicroRNAs and the PTEN/PI3K/Akt pathway in gastric cancer (Review). *Oncol Rep*. 2019; 41(3):1439-1454.
23. Li Q, Liu S, Yan J, Sun MZ, Greenaway FT. The potential role of miR-124-3p in tumorigenesis and other related diseases. *Mol Biol Rep*. 2021; 48(4):3579-3591.
24. Romero-López MJ, Jiménez-Wences H, Cruz-De la Rosa MI, Román-Fernández IV, Fernández-Tilapa G. miR-23b-3p, miR-124-3p and miR-218-5p Synergistic or Additive Effects on Cellular Processes That Modulate Cervical Cancer Progression? A Molecular Balance That Needs Attention. *Int J Mol Sci*. 2022; 23(21):13551.
25. Peng B, Guo C, Guan H, Liu S, Sun MZ. Annexin A5 as a potential marker in tumors. *Clin Chim Acta*. 2014; 427:42-8.
26. Wang X, Dai Y, Zhao Y, Li M, Zhang J, Ci Y, Wang H, Li X. AnnexinA5 Might Suppress the Phenotype of Human Gastric Cancer Cells via ERK Pathway. *Front Oncol*. 2021; 11:665105.
27. Zhang K, Yang G, Wu W, Zhang J, Xia X, Jiang T, Cao J, Huang K, Qiu Z, Huang C. Decreased Expression of Caveolin-1 and E-Cadherin Correlates with the Clinicopathologic Features of Gastric Cancer and the EMT Process. *Recent Pat Anticancer Drug Discov*. 2016; 11(2):236-44.
28. Luo Z, Rong Z, Zhang J, Zhu Z, Yu Z, Li T, Fu Z, Qiu Z, Huang C. Circular RNA circCCDC9 acts as a miR-6792-3p sponge to suppress the progression of gastric cancer through regulating CAV1 expression. *Mol Cancer*. 2020; 19(1):86.
29. Wang JK, Wang Z, Li G. MicroRNA-125 in immunity and cancer. *Cancer Lett*. 2019; 454:134-145.
30. Russo A, Potenza N. Antiproliferative Activity of microRNA-125a and its Molecular Targets. *Microna*. 2019; 8(3):173-179.
31. Amiri-Dashatan N, Koushki M, Naghi-Zadeh M, Razzaghi MR, Mohaghegh Shalmani H. Prognostic value of microRNA-125a/b family in patients with gastric cancer: a meta-analysis. *Gastroenterol Hepatol Bed Bench*. 2021; 14(Suppl1):S1-S9.
32. Boi M, Zucca E, Inghirami G, Bertoni F. PRDM1/BLIMP1: a tumor suppressor gene in B and T cell lymphomas. *Leuk Lymphoma*. 2015; 56(5):1223-8.
33. Shen L, Chen Q, Yang C, Wu Y, Yuan H, Chen S, Ou S, Jiang Y, Huang T, Ke L, Mo J, Feng Z, Zhou P, Fan W. Role of PRDM1 in Tumor Immunity and Drug Response: A Pan-Cancer Analysis. *Front Pharmacol*. 2020; 11:593195.
34. Martin-Arruti M, Vaquero M, Díaz de Otazu R, Zabalza I, Ballesteros J, Roncador G, García-Orad A. Bcl-2 and BLIMP-1 expression predict worse prognosis in gastric diffuse large B cell lymphoma (DLCL) while other markers for nodal DLCL are not useful. *Histopathology*. 2012; 60(5):785-92.
35. Shin J, Shim HG, Hwang T, Kim H, Kang SH, Dho YS, Park SH, Kim SJ, Park CK. Restoration of miR-29b-3p exerts anti-cancer effects on glioblastoma. *Cancer Cell Int*. 2017; 17:104.
36. Ulivi P, Canale M, Passardi A, Marisi G, Valgiusti M, Frassinetti GL, Calistri D, Amadori D, Scarpi E. Circulating Plasma Levels of miR-20b, miR-29b-3p and miR-155 as Predictors of Bevacizumab Efficacy in Patients with Metastatic Colorectal Cancer. *Int J Mol Sci*. 2018; 19(1):307.
37. Amirian M, Jafari-Nozad AM, Darroudi M, Farkhondeh T, Samarghandian S. Overview of the miR-29 family members' function in breast cancer. *Int J Biol Macromol*. 2023; 230:123280.
38. Zhao X, Liu Y, Li Z, Zheng S, Wang Z, Li W, Bi Z, Li L, Jiang Y, Luo Y, Lin Q, Fu Z, Rufu C. Linc00511 acts as a competing endogenous RNA to regulate VEGFA expression through sponging hsa-miR-29b-3p in pancreatic ductal adenocarcinoma. *J Cell Mol Med*. 2018; 22(1):655-667.
39. Bloomfield J, Sabbah M, Castela M, Mehats C, Uzan C, Canlorbe G. Clinical Value and Molecular Function of Circulating MicroRNAs in Endometrial Cancer Regulation: A Systematic Review. *Cells*. 2022; 11(11):1836.
40. Guo X, Li K, Jiang W, Hu Y, Xiao W, Huang Y, Feng Y, Pan Q, Wan R. RNA demethylase ALKBH5 prevents pancreatic cancer progression by posttranscriptional activation of PER1 in an m6A-YTHDF2-dependent manner. *Mol Cancer*. 2020; 19(1):91.
41. Othumpangat S, Noti JD, Beezhold DH. Lung epithelial cells resist influenza A infection by inducing the expression of cytochrome c oxidase VIc which is modulated by miRNA 4276. *Virology*. 2014; 468-470:256-264.

42. Zhang S, Zang D, Cheng Y, Li Z, Yang B, Guo T, Liu Y, Qu X, Che X. Identification of Key Gene and Pathways for the Prediction of Peritoneal Metastasis of Gastric Cancer by Co-expression Analysis. *J Cancer*. 2020; 11(10):3041-3051.
43. Ma X, Wang J, Zhao M, Huang H, Wu J. Increased expression of methionine sulfoxide reductases B3 is associated with poor prognosis in gastric cancer. *Oncol Lett*. 2019; 18(1):465-471.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.