



# Evaluation of Plasma Extracellular Vesicle MicroRNA Signatures for Lung Adenocarcinoma and Granuloma With Monte-Carlo Feature Selection Method

Xiangbo Chen<sup>1,2</sup>, Yunjie Jin<sup>3</sup> and Yu Feng<sup>4\*</sup>

<sup>1</sup> Key Laboratory of Molecular Epigenetics of the Ministry of Education, Northeast Normal University, Changchun, China, <sup>2</sup> Hangzhou Baocheng Biotechnology Co., Ltd., Hangzhou, China, <sup>3</sup> Department of Oncology, Shanghai Putuo People's Hospital, Shanghai, China, <sup>4</sup> Shuguang Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China

## OPEN ACCESS

### Edited by:

Tao Huang,  
Shanghai Institutes for Biological  
Sciences (CAS), China

### Reviewed by:

Shijia Zhu,  
The University of Texas Southwestern  
Medical Center, United States  
Cheng Guo,  
Columbia University, United States

### \*Correspondence:

Yu Feng  
Fengyu@shutcm.edu.cn;  
dryufeng021@163.com

### Specialty section:

This article was submitted to  
Bioinformatics and Computational  
Biology,  
a section of the journal  
Frontiers in Genetics

Received: 18 December 2018

Accepted: 05 April 2019

Published: 26 April 2019

### Citation:

Chen X, Jin Y and Feng Y (2019)  
Evaluation of Plasma Extracellular  
Vesicle MicroRNA Signatures for Lung  
Adenocarcinoma and Granuloma  
With Monte-Carlo Feature Selection  
Method. *Front. Genet.* 10:367.  
doi: 10.3389/fgene.2019.00367

Extracellular Vesicle (EV) is a compilation of secreted vesicles, including micro vesicles, large oncosomes, and exosomes. It can be used in non-invasive diagnosis. MicroRNAs (miRNAs) processed by exosomes can be detected by liquid biopsy. To objectively evaluate the discriminative ability of miRNAs from whole plasma, EV and EV-free plasma, we analyzed the miRNA expression profiles in whole plasma, EV and EV-free plasma of 10 lung adenocarcinoma and 9 granuloma patients. With Monte-Carlo feature selection method, the top discriminative miRNAs in whole plasma, EV and EV-free plasma were identified, and they were quite different. Using the Repeated Incremental Pruning to Produce Error Reduction (RIPPER) method, we learned the classification rules: in whole plasma, granuloma patients did not express hsa-miR-223-3p while the lung adenocarcinoma patients expressed hsa-miR-223-3p; in EV, the hsa-miR-23b-3p was highly expressed in granuloma patients but not lung adenocarcinoma patients; in EV-free plasma, hsa-miR-376a-3p was expressed in granuloma patients but barely expressed in lung adenocarcinoma patients. For prediction performance, whole plasma had the highest weighted accuracy and EV outperformed EV-free plasma. Our results suggested that EV can be used as lung cancer biomarker. However, since it is less stable and not easy to detect, there are still technological difficulties to overcome.

**Keywords:** microRNA signatures, biomarker, classification, lung adenocarcinoma, granuloma

## INTRODUCTION

Blood is a mixture of plasma, blood platelet and various blood cells, such as erythrocytes, leukocytes, neutrophilic granulocytes, eosinophilic granulocytes, basophilic granulocytes, monocytes, and lymphocytes (Basu and Kulkarni, 2014). It can reflect the body health and wellness. Extracellular Vesicle (EV) is a compilation of secreted vesicles, including micro vesicles, large oncosomes, and exosomes (Lawson et al., 2018). Exosomes, with a diameter of 30–100 nm, are a kind of membrane-bound EVs and originate from endosome (Raposo and Stoorvogel, 2013). Nearly all kinds of cells can secrete exosomes whether under normal or stressful conditions (Srivastava et al., 2015). When compared with normal cells, tumor cells of the specific organs have been

proven to secrete more exosomes. Besides, the membrane of exosomes richly contains plenty of functional proteins, including tetraspanin, endosome-related membrane transport and fusion proteins and multivesicular bodies-genesis proteins, and thus exosomes could be applied as biomarkers (O'Driscoll, 2015). Exosomes can be extracted from diverse body fluids, which contain numerous biological molecules (DNAs, RNAs, and proteins). Recently, liquid biopsy has been developed as a novel, non-invasive diagnosis method to explore tumor development (Sheridan, 2016).

MicroRNAs (miRNAs) processed by exosomes could be detected by liquid biopsy (Iranifar et al., 2019). miRNAs are a group of non-coding RNAs, which regulate gene expression at the post-transcriptional and translational levels (Inamura, 2017). Dysregulation of miRNA expression is related to the progression of lung adenocarcinoma. Besides, Nadal et al. (2014) have demonstrated that different morphological subtypes of lung adenocarcinoma have specific miRNA expression profiles, for instance, miR-212-3p, miR-132-5p, and miR-27a-3p are found significantly upregulated in adenocarcinomas with solid subtype. A mass of miRNAs play important roles in the process of lung cancer pathogenesis and are recognized as potential diagnostic biomarkers and tumor targeted therapeutic molecules (Inamura, 2017).

As a well-studied, common cancer, lung cancer maintains the leading cause of cancer-specific death around the world. Adenocarcinoma accounts for nearly half of all lung cancer types, remaining the most common histologic subtype (Travis et al., 2011; Rosell and Karachaliou, 2018). Although the development of new therapies has significantly improved the prognosis of patients with lung adenocarcinoma, the 5-year survival rate remains low (less than 16%) (Crino et al., 2010).

To evaluate the discriminative ability of miRNAs from whole plasma, EV, and EV-free plasma, we analyzed the miRNA expression profiles in whole plasma, EV, and EV-free plasma of lung adenocarcinoma and granuloma patients. The same feature selection method, Monte-Carlo feature selection and the same rule learner, Repeated Incremental Pruning to Produce Error Reduction (RIPPER), were applied in the three miRNA expression datasets for lung adenocarcinoma and granuloma patients. The prediction performances and classification rules of whole plasma, EV, and EV-free plasma were compared and analyzed. Our results suggested that the prediction performance of EV miRNAs was better than EV-free plasma miRNAs. What's more, we identified EV specific miRNA expression pattern in lung cancer. These results supported the usage of EV miRNAs as lung cancer biomarkers but the whole plasma achieved a better prediction performance. The utilization of EV biomarkers still has a long way to go.

## MATERIALS AND METHODS

### The MicroRNA Expression Profiles in Whole Plasma, EV, and EV-Free Plasma

We downloaded the processed miRNA expression profiles in whole plasma, EV, and EV-free plasma of 10 lung

adenocarcinoma patients and the miRNA expression profiles in whole plasma, EV and EV-free plasma of 9 granuloma patients from GEO (Gene Expression Omnibus) under accession number of GSE71661 on August 30, 2018. The expression levels of miRNAs were measured with next generation sequencing using Illumina HiSeq 2500. The reads were mapped onto known human miRNA in miRbase Release 21 using Blast and Bowtie. The mapped reads were normalized with the total number of reads. In each miRNA dataset of whole plasma, EV, and EV-free plasma, there were 10 lung adenocarcinoma and 9 granuloma patients; there were 1,509 miRNAs. The downloaded miRNA profiles were provided in **Supplementary Table S1**.

To systematically compare the miRNA expression difference between lung adenocarcinoma and granuloma patients, whole plasma, EV, and EV-free plasma were analyzed separately. Our goal was to compare their prediction performance and unique expression of miRNAs.

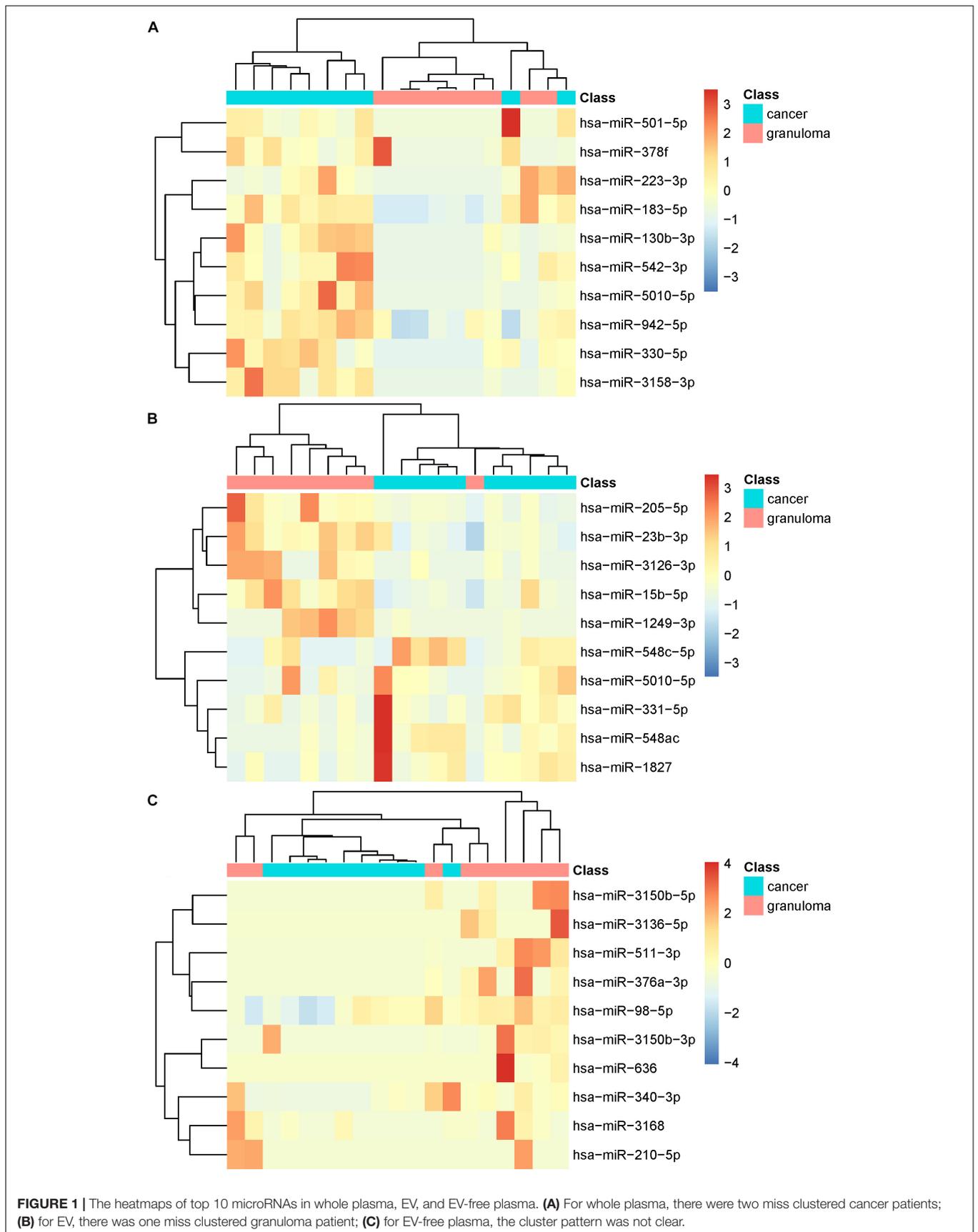
### Key MicroRNAs in Whole Plasma, EV, and EV-Free Plasma Identified With Monte-Carlo Feature Selection

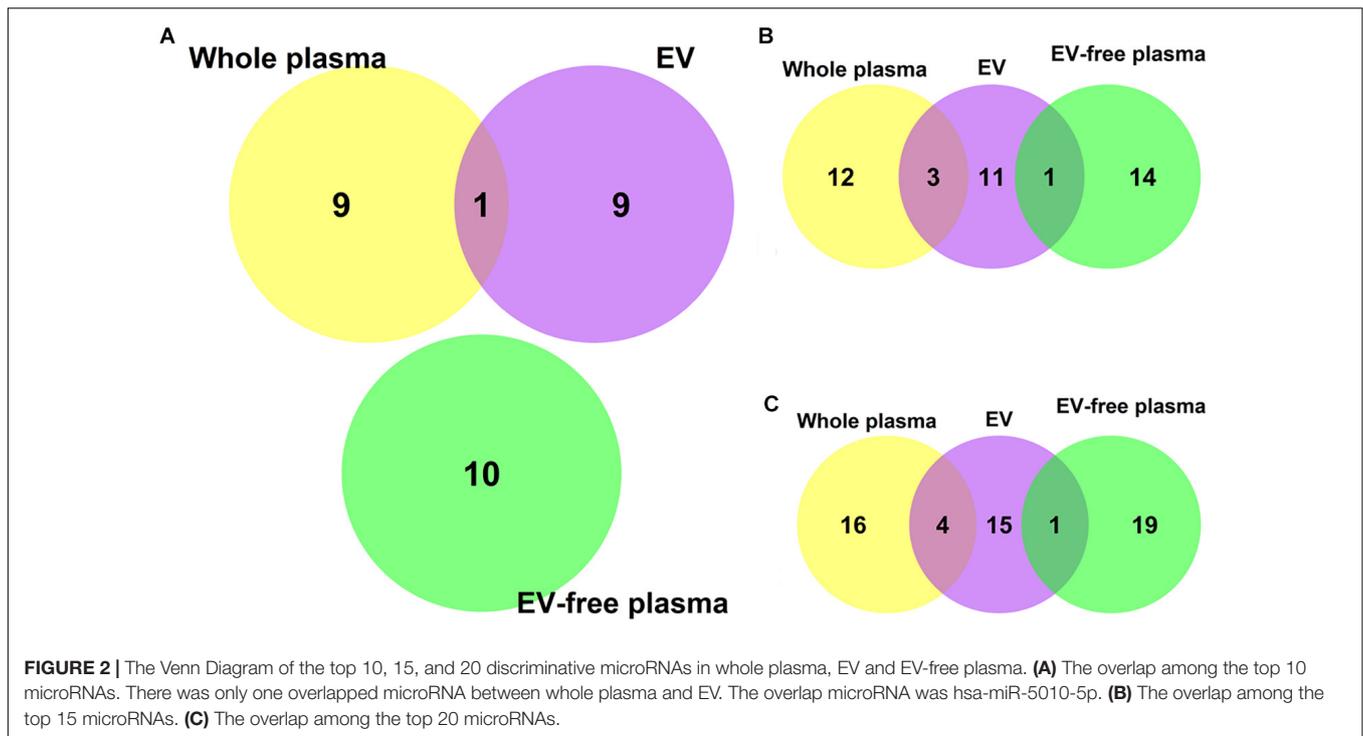
Since there were 19 samples and 1,509 miRNA features in whole plasma, EV, and EV-free plasma dataset, the number of features was much greater than the sample size. If we use all miRNAs to build the classification model, all samples will be perfectly classified. But it will be overfitting and will have no actual meanings. Therefore, we adopted the Monte-Carlo feature selection (Draminski et al., 2008) to identify the key miRNA features and then used these few key features to construct the classification model. The Monte-Carlo feature selection method has been widely used and has achieved great performance in many fields (Chen et al., 2018c,e; Pan et al., 2019).

The Monte-Carlo feature selection method will randomly choose several features multiple times and then construct a series of tree classifiers (Chen et al., 2018a; Pan et al., 2018b; Wang et al., 2018). Based on the frequency and classification accuracies of the feature nodes on these classification trees, each feature will be assigned with a relative importance. Intuitively speaking, if a feature has been selected many times to construct the classification tree, it is important as the classification tree will find the most discriminative features to be the nodes.

Let's denote the total number of miRNA features with  $d$ , i.e., 1,509 in this study.  $m$  miRNA features ( $m \ll d$ ) will be randomly selected and be used to construct  $t$  classification trees for  $s$  times. Each of the  $t$  trees was trained and tested based on the training and test patient samples randomly divided from the full dataset. Therefore,  $s \cdot t$  classification trees will be constructed. Based on how many times a miRNA feature  $g$  has been selected by these  $s \cdot t$  trees and how much this miRNA feature  $g$  has contributed to the classification of the  $s \cdot t$  trees, its relative importance (RI) can be calculated:

$$RI_g = \sum_{\tau=1}^{st} (wAcc)^\mu \sum_{n_g(\tau)} IG(n_g(\tau)) \left( \frac{\text{no} \cdot \text{in } n_g(\tau)}{\text{no} \cdot \text{in } \tau} \right)^\nu \quad (1)$$





where  $wAcc$  is the weighted classification accuracy of decision tree  $\tau$ ,  $IG(n_g(\tau))$  is the information gain of node  $n_g(\tau)$ , which is a decision rule using the expression levels of miRNA feature  $g$ , ( $no \cdot$  in  $n_g(\tau)$ ) is the number of samples under node  $n_g(\tau)$ , ( $no \cdot$  in  $\tau$ ) is the number of samples in decision tree  $\tau$ ,  $u$ , and  $v$  are adjust parameters.

By analyzing the  $s \cdot t$  classification trees, each miRNA feature will be assigned with a RI and will be ranked decreasingly.

The Monte-Carlo feature selection method was applied using the dmLab software (Draminski et al., 2008) downloaded from <http://www.ipipan.eu/staff/m.draminski/mcfs.html>.

### Classification Rules for Lung Adenocarcinoma and Granuloma in Whole Plasma, EV, and EV-Free Plasma Learned With RIPPER

Repeated Incremental Pruning to Produce Error Reduction is a widely used method to learn the classification rules (Cai et al., 2018; Chen et al., 2018a,c,e,f; Pan et al., 2018a). Since we want to evaluate the prediction performance objectively, we did the 10-fold cross-validation for three times and combined the three-time results. In each cross validation (Wang et al., 2017; Zhang et al., 2017; Chen et al., 2018b,d; Li et al., 2018), the samples were randomly divided into 10 parts and each part was used as test dataset for once. After 10 rounds, all samples have been tested. As the random splits of data may cause bias, we repeated the 10-fold cross-validation for three times. In this study, the lung adenocarcinoma patients and granuloma patients were treated as positive samples and negative samples, respectively. We used weighted accuracy to evaluate the RIPPER prediction

performance, i.e., the average of the accuracies of positive samples and negative samples.

## RESULTS

### The Discriminative MicroRNAs Between Lung Adenocarcinoma and Granuloma Patients in Whole Plasma, EV, and EV-Free Plasma

The miRNA expression profiles of lung adenocarcinoma and granuloma patients in whole plasma, EV and EV-free plasma were analyzed separately. In whole plasma, the top 10 discriminative miRNAs were hsa-miR-223-3p, hsa-miR-501-5p, hsa-miR-130b-3p, hsa-miR-5010-5p, hsa-miR-330-5p, hsa-miR-378f, hsa-miR-3158-3p, hsa-miR-542-3p, hsa-miR-183-5p and hsa-miR-942-5p. In EV, the top 10 discriminative miRNAs were hsa-miR-23b-3p, hsa-miR-548ac, hsa-miR-3126-3p, hsa-miR-15b-5p, hsa-miR-205-5p, hsa-miR-5010-5p, hsa-miR-331-5p, hsa-miR-1249-3p, hsa-miR-548c-5p, and hsa-miR-1827. In EV-free plasma, the top 10 discriminative miRNAs were hsa-miR-511-3p, hsa-miR-376a-3p, hsa-miR-3150b-3p, hsa-miR-3150b-5p, hsa-miR-3168, hsa-miR-98-5p, hsa-miR-3136-5p, hsa-miR-210-5p, hsa-miR-340-3p, and hsa-miR-636. **Figure 1** shows the heatmaps of the top 10 miRNAs in whole plasma, EV and EV-free plasma. The miRNAs and patients were clustered using ward D2 method (Murtagh and Legendre, 2014) based on Euclidean distance. The R package pheatmap<sup>1</sup> was applied to plot the heatmaps. It can be seen

<sup>1</sup><https://CRAN.R-project.org/package=pheatmap>

from **Figure 1** that for whole plasma, there were two miss clustered cancer patients; for EV, there was one miss clustered granuloma patient; for EV-free plasma, the cluster pattern was not clear. The miRNAs in EV-free plasma were not suitable as cancer biomarkers.

We plotted the Venn Diagram of the top 10 discriminative miRNAs in whole plasma, EV and EV-free plasma in **Figure 2A**. There was only one overlapped miRNA between whole plasma and EV. The overlap miRNA was hsa-miR-5010-5p. It can be seen that the miRNA expression pattern was different in whole plasma, EV and EV-free plasma. It was necessary to investigate which blood compartments should be used for biomarker discovery.

To investigate whether the overlap pattern would change when more miRNAs were analyzed, we plotted Venn Diagrams of the top 15 and top 20 miRNAs as **Figures 2B,C**, respectively. There was still no overlap among the whole plasma, EV and EV-free plasma. The overlap between whole plasma and EV became larger when more top miRNAs were included but the overlap between EV and EV-free plasma remained to be one no matter whether the top 15 or 20 miRNAs were analyzed. The EV miRNAs were more similar with the whole plasma miRNAs than the EV-free plasma miRNAs.

### The Prediction Accuracies of MicroRNA Signatures for Lung Adenocarcinoma and Granuloma Patients in Whole Plasma, EV, and EV-Free Plasma

We evaluated the prediction accuracies of miRNA signatures for lung adenocarcinoma and granuloma patients in whole plasma, EV and EV-free plasma with 10-fold cross validations. To avoid the bias of random splits of samples, we repeated the 10-fold cross validation for three times. Therefore, the samples size in the confusion matrix will be the original sample size 19 multiplied by 3 which was 57.

The confusion matrices of miRNA signatures in whole plasma, EV and EV-free plasma were given in **Table 1**. The weighted accuracies using whole plasma, EV and EV-free plasma miRNA data were 77.22, 65.19, and 64.82%, respectively. The EV miRNAs performed better than the EV-free plasma miRNAs. The accuracy of granuloma in EV-free plasma, 29.63%, was extremely low.

### The Classification Rules in Whole Plasma, EV, and EV-Free Plasma

With the RIPPER method, we learned the classifications of miRNA expression levels in whole plasma, EV and EV-free plasma. These rules were given in **Table 2**. In whole plasma, granuloma patients did not express hsa-miR-223-3p while the lung adenocarcinoma patients expressed hsa-miR-223-3p. In EV, the hsa-miR-23b-3p was highly expressed in granuloma patients but not lung adenocarcinoma patients. In EV-free plasma, hsa-miR-376a-3p was expressed in granuloma patients but barely expressed in lung adenocarcinoma patients. We compared the mean expression levels of hsa-miR-23b-3p in whole plasma cancer, whole plasma granuloma, EV cancer

**TABLE 1** | The confusion matrices of whole plasma, EV, and EV-free plasma microRNAs.

	Whole plasma		EV		EV-free plasma	
	Predicted granuloma	Predicted adenocarcinoma	Predicted granuloma	Predicted adenocarcinoma	Predicted granuloma	Predicted adenocarcinoma
Actual granuloma	21	6	19	8	8	19
Actual adenocarcinoma	7	23	12	18	0	30
Granuloma accuracy	21/(21+6) = 77.77%		19/(19+8) = 70.37%		8/(8+19) = 29.63%	
Adenocarcinoma accuracy	23/(7+23) = 76.67%		18/(12+18) = 60.00%		30/(0+30) = 100.00%	
Weighted accuracy	(77.77%+76.67%)/2 = 77.22%		(70.37%+60.00%)/2 = 65.19%		(29.63%+100.00%)/2 = 64.82%	

**TABLE 2** | The RIPPER rules in whole plasma, EV and EV-free plasma.

Whole plasma		EV		EV-free plasma	
Granuloma	hsa-miR-223-3p $\leq 0$	Granuloma	hsa-miR-23b-3p $\geq 210.43$	Granuloma	hsa-miR-376a-3p $\geq 4.50$
Adenocarcinoma	Others	Adenocarcinoma	Others	Adenocarcinoma	Others

and EV granuloma. We found that in EV, hsa-miR-23b-3p was more highly expressed in granuloma than cancer with a fold change of 1.82, while in whole plasma, hsa-miR-23b-3p was more lowly expressed in granuloma than cancer with fold change of 0.84. What's more, we compared the mean expression levels of hsa-miR-376a-3p in EV-free plasma as well. We found that in EV-free plasma, the mean expression levels of hsa-miR-376a-3p in cancer and granuloma were 0 and 10.30, respectively, while in whole plasma, the mean expression levels of hsa-miR-376a-3p in cancer and granuloma were 1.45 and 0, respectively. The expression pattern between EV or EV-free plasma and whole plasma were different. These results suggested it was necessary to measure the EV, EV-free plasma and whole plasma, separately.

hsa-miR-223-3p was reported to have an increased expression in *H. pylori*-infected gastric cancer patients, which was related to progressive proliferation and migration of cancer cells (Ma et al., 2014; Wang et al., 2015). Thus, in plasma, the expression of hsa-miR-223-3p in granuloma patients would not be as high as in cancer patients.

Zhou et al. (2015) found that cancer patients with higher expression of has-miR-23b had better outcomes than those with lower expression. In our study, we found that has-miR-23b-3p had higher expression in granuloma patients compared to in lung adenocarcinoma patients.

Joerger et al. (2014) reported that hsa-miR-376a was insensitive to perturbations in advanced non-small cell lung cancer patients. We found has-miR-376a-3p had a higher expression in granuloma patients, while its expression was very low in lung adenocarcinoma patients.

## DISCUSSION

We identified the discriminative miRNAs in different blood compartments, such as hsa-miR-501-5p and hsa-miR-130b-3p in plasma; hsa-miR-548ac in EV and hsa-miR-511-3p in EV-free plasma.

hsa-miR-501 has been proven to have an association with clear cell renal cell carcinoma (Liu et al., 2018), pancreatic ductal adenocarcinoma (Liao et al., 2018), cervical cancer (Guo et al., 2018) and so on. Besides, they all found upregulation of has-miR-501 enhances tumor cell proliferation, migration and invasion.

hsa-miR-130b-3p is a novel miRNA in lung cancer, we found hsa-miR-130b-3p are upregulated in the plasma of lung cancer patients, which would be applied as a new biomarker to distinguish cancer and granuloma, and further guide therapeutic decisions clinically.

As for hsa-miR-548, Liu et al. (2015) investigated hsa-miR-548 expression in fresh tumor tissues from 22 patients with primary

non-small cell lung cancer via RT-PCR and they found that the hsa-miR-548 expression level was significantly higher ( $p < 0.01$ ) in adjacent non-tumor tissues than that in the tumor. That is, non-small cell lung cancer would down-regulate the expression of hsa-miR-548. Furthermore, they also observed that hsa-miR-548 was involved in the migration and invasion of non-small cell lung cancer cells by targeting the AKT1 signaling pathway.

For hsa-miR-511-3p, it has been reported to be related to lung adenocarcinoma by triggering BAX (Zhang et al., 2014) and TRIB2 (Zhang et al., 2012).

As for the diagnostic value, plasma is the most valuable, followed by EV and EV-free plasma. Previous studies have demonstrated that exosomes can be used as a type of novel biomarker for tumors and some benign diseases (Principe et al., 2013; Vella et al., 2016). Considering the diagnostic value of testing plasma is better than testing exosomes in plasma, many useful information may be missed when only exosomes in plasma were tested. The reasons are as follows: (1) Methods like OptiPrep<sup>TM</sup> density-based separation (DG-Exos), ultracentrifugation (UC-Exos), and immunoaffinity capture using anti-EpCAM-coated magnetic beads (IAC-Exos) are not effective enough to isolate exosomes and may destroy exosomes during the isolation process (Greening et al., 2015); (2) exosomes are not stable and are easily degraded, which could cause a bias (Kumar et al., 2018).

Since the sample size of this study was limited, the results should be validated in an independent large cohort. Another factor that may have affected the results was the disease type. For lung adenocarcinoma, the results were like this. But for other diseases, which release a large amount of RNAs and proteins into the circulatory system directly, the importance of exosome may decrease.

## CONCLUSION

Extracellular Vesicle is a promising technology for non-invasive diagnosis. miRNAs processed by exosomes can be detected by liquid biopsy and used as biomarkers. To evaluate the discriminative ability of miRNAs from whole plasma, EV and EV-free plasma, we analyzed the miRNA expression profiles in whole plasma, EV and EV-free plasma of lung adenocarcinoma and granuloma patients. We found that the top discriminative miRNAs in whole plasma, EV and EV-free plasma were quite different, and the classification rules also varied. The prediction performance of whole plasma was the best but the EV outperformed EV-free plasma. Our results suggested that EV can be used as a lung cancer biomarker but EV may be less stable or difficult to detect than whole plasma, therefore, the whole plasma was still a good choice as lung cancer signatures.

## AUTHOR CONTRIBUTIONS

XC and YJ did the conception and design, performed the sample collection, analyzed and interpreted the data, and wrote, reviewed, and/or revised the manuscript. XC oversaw the developmental methodology. All authors read and approved the final manuscript.

## REFERENCES

- Basu, D., and Kulkarni, R. (2014). Overview of blood components and their preparation. *Indian J. Anaesth.* 58, 529–537. doi: 10.4103/0019-5049.144647
- Cai, Y. D., Zhang, S., Zhang, Y. H., Pan, X., Feng, K., Chen, L., et al. (2018). Identification of the gene expression rules that define the subtypes in Glioma. *J. Clin. Med.* 7:E350. doi: 10.3390/jcm7100350
- Chen, L., Li, J., Zhang, Y. H., Feng, K., Wang, S., Zhang, Y., et al. (2018a). Identification of gene expression signatures across different types of neural stem cells with the Monte-Carlo feature selection method. *J. Cell. Biochem.* 119, 3394–3403. doi: 10.1002/jcb.26507
- Chen, L., Pan, X., Hu, X., Zhang, Y. H., Wang, S., Huang, T., et al. (2018b). Gene expression differences among different MSI statuses in colorectal cancer. *Int. J. Cancer* doi: 10.1002/ijc.31554 [Epub ahead of print].
- Chen, L., Pan, X., Zhang, Y. H., Kong, X., Huang, T., and Cai, Y. D. (2018c). Tissue differences revealed by gene expression profiles of various cell lines. *J. Cell. Biochem.* doi: 10.1002/jcb.27977 [Epub ahead of print].
- Chen, L., Wang, S., Zhang, Y. H., Wei, L., Xu, X., Huang, T., et al. (2018d). Prediction of nitrated tyrosine residues in protein sequences by extreme learning machine and feature selection methods. *Comb. Chem. High Throughput Screen.* 21, 393–402. doi: 10.2174/1386207321666180531091619
- Chen, L., Zhang, S., Pan, X., Hu, X., Zhang, Y. H., Yuan, F., et al. (2018e). HIV infection alters the human epigenetic landscape. *Gene Ther.* 26, 29–39. doi: 10.1038/s41434-018-0051-6
- Chen, L., Zhang, Y. H., Pan, X., Liu, M., Wang, S., Huang, T., et al. (2018f). Tissue expression difference between mRNAs and lncRNAs. *Int. J. Mol. Sci.* 19:E3416.
- Crino, L., Weder, W., van Meerbeek, J., Felip, E., and Group, E. G. W. (2010). Early stage and locally advanced (non-metastatic) non-small-cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 21(Suppl. 5), v103–v115.
- Draminski, M., Rada-Iglesias, A., Enroth, S., Wadelius, C., Koronacki, J., and Komorowski, J. (2008). Monte Carlo feature selection for supervised classification. *Bioinformatics* 24, 110–117.
- Greening, D. W., Xu, R., Ji, H., Tauro, B. J., and Simpson, R. J. (2015). A protocol for exosome isolation and characterization: evaluation of ultracentrifugation, density-gradient separation, and immunoaffinity capture methods. *Methods Mol. Biol.* 1295, 179–209. doi: 10.1007/978-1-4939-2550-6\_15
- Guo, J., Jin, D., Wu, Y., Yang, L., Du, J., Gong, K., et al. (2018). The miR 495-UBE2C-ABCG2/ERCC1 axis reverses cisplatin resistance by downregulating drug resistance genes in cisplatin-resistant non-small cell lung cancer cells. *EBioMedicine* 35, 204–221. doi: 10.1016/j.ebiom.2018.08.001
- Inamura, K. (2017). Diagnostic and therapeutic potential of MicroRNAs in lung cancer. *Cancers* 9:E49. doi: 10.3390/cancers9050049
- Iranifar, E., Seresht, B. M., Momeni, F., Fadaei, E., Mehr, M. H., Ebrahimi, Z., et al. (2019). Exosomes and microRNAs: new potential therapeutic candidates in Alzheimer disease therapy. *J. Cell. Physiol.* 234, 2296–2305. doi: 10.1002/jcp.27214
- Joerger, M., Baty, F., Fruh, M., Droegge, C., Stahel, R. A., Betticher, D. C., et al. (2014). Circulating microRNA profiling in patients with advanced non-squamous NSCLC receiving bevacizumab/erlotinib followed by platinum-based chemotherapy at progression (SAKK 19/05). *Lung Cancer* 85, 306–313. doi: 10.1016/j.lungcan.2014.04.014
- Kumar, S., Michael, I. J., Park, J., Granick, S., and Cho, Y. K. (2018). Cloaked exosomes: biocompatible, durable, and degradable encapsulation. *Small* 14:e1802052. doi: 10.1002/smll.201802052
- Lawson, J., Dickman, C., Towle, R., Jabalee, J., Rani, A., and Garnis, C. (2018). Extracellular vesicle secretion of miR-142-3p from lung adenocarcinoma

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2019.00367/full#supplementary-material>

**TABLE S1** | The processed microRNA expression profiles.

- cells induces tumor promoting changes in the stroma through cell-cell communication. *Mol. Carcinog.* 58, 376–387. doi: 10.1002/mc.22935
- Li, J., Lan, C. N., Kong, Y., Feng, S. S., and Huang, T. (2018). Identification and analysis of blood gene expression signature for osteoarthritis with advanced feature selection methods. *Front. Genet.* 9:246. doi: 10.3389/fgene.2018.00246
- Liao, X., Wang, X., Huang, K., Yang, C., Yu, T., Han, C., et al. (2018). Genome-scale analysis to identify prognostic microRNA biomarkers in patients with early stage pancreatic ductal adenocarcinoma after pancreaticoduodenectomy. *Cancer Manag. Res.* 10, 2537–2551. doi: 10.2147/CMAR.S168351
- Liu, C., Yang, H., Xu, Z., Li, D., Zhou, M., Xiao, K., et al. (2015). microRNA-548l is involved in the migration and invasion of non-small cell lung cancer by targeting the AKT1 signaling pathway. *J. Cancer Res. Clin. Oncol.* 141, 431–441. doi: 10.1007/s00432-014-1836-7
- Liu, J., Liu, B., Guo, Y., Chen, Z., Sun, W., Gao, W., et al. (2018). Key miRNAs and target genes played roles in the development of clear cell renal cell carcinoma. *Cancer Biomark.* 23, 279–290. doi: 10.3233/CBM-181558
- Ma, L., Chen, Y., Zhang, B., and Liu, G. (2014). Increased microRNA-223 in *Helicobacter pylori*-associated gastric cancer contributed to cancer cell proliferation and migration. *Biosci. Biotechnol. Biochem.* 78, 602–608. doi: 10.1080/09168451.2014.895661
- Murtagg, F., and Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? *J. Classif.* 31, 274–295.
- Nadal, E., Zhong, J., Lin, J., Reddy, R. M., Ramnath, N., Orringer, M. B., et al. (2014). A MicroRNA cluster at 14q32 drives aggressive lung adenocarcinoma. *Clin. Cancer Res.* 20, 3107–3117. doi: 10.1158/1078-0432.CCR-13-3348
- O'Driscoll, L. (2015). Expanding on exosomes and ectosomes in cancer. *N. Engl. J. Med.* 372, 2359–2362.
- Pan, X., Hu, X., Zhang, Y. H., Chen, L., Zhu, L., Wan, S., et al. (2018a). Identification of the copy number variant biomarkers for breast cancer subtypes. *Mol. Genet. Genomics* 294, 95–110.
- Pan, X., Hu, X., Zhang, Y. H., Feng, K., Wang, S. P., Chen, L., et al. (2018b). Identifying patients with atrioventricular septal defect in down syndrome populations by using self-normalizing neural networks and feature selection. *Genes* 9:E208. doi: 10.3390/genes9040208
- Pan, X., Hu, X., Zhang, Y. H., Chen, L., Zhu, L., Wan, S., et al. (2019). Identification of the copy number variant biomarkers for breast cancer subtypes. *Mol. Genet. Genomics* 294, 95–110. doi: 10.1007/s00438-018-1488-4
- Principe, S., Hui, A. B., Bruce, J., Sinha, A., Liu, F. F., and Kislinger, T. (2013). Tumor-derived exosomes and microvesicles in head and neck cancer: implications for tumor biology and biomarker discovery. *Proteomics* 13, 1608–1623. doi: 10.1002/pmic.201200533
- Raposo, G., and Stoorvogel, W. (2013). Extracellular vesicles: exosomes, microvesicles, and friends. *J. Cell Biol.* 200, 373–383. doi: 10.1083/jcb.201211138
- Rosell, R., and Karachaliou, N. (2018). Gene expression signatures predicting survival and chemotherapy benefit in patients with resected non-small-cell lung cancer. *EBioMedicine* 33, 16–17.
- Sheridan, C. (2016). Exosome cancer diagnostic reaches market. *Nat. Biotechnol.* 34, 359–360.
- Srivastava, A., Filant, J., Moxley, K. M., Sood, A., McMeekin, S., and Ramesh, R. (2015). Exosomes: a role for naturally occurring nanovesicles in cancer growth, diagnosis and treatment. *Curr. Gene Ther.* 15, 182–192.
- Travis, W. D., Brambilla, E., Noguchi, M., Nicholson, A. G., Geisinger, K., Yatabe, Y., et al. (2011). International association for the study of lung cancer/American thoracic society/European respiratory society: international multidisciplinary classification of lung adenocarcinoma: executive summary. *Proc. Am. Thorac. Soc.* 8, 381–385. doi: 10.1513/pats.201107-042ST
- Vella, L. J., Hill, A. F., and Cheng, L. (2016). Focus on extracellular vesicles: exosomes and their role in protein trafficking and biomarker potential in

- Alzheimer's and Parkinson's disease. *Int. J. Mol. Sci.* 17:173. doi: 10.3390/ijms17020173
- Wang, C., Ding, M., Xia, M., Chen, S., Van Le, A., Soto-Gil, R., et al. (2015). A five-miRNA panel identified from a multicentric case-control study serves as a novel diagnostic tool for ethnically diverse non-small-cell lung cancer patients. *EBioMedicine* 2, 1377–1385. doi: 10.1016/j.ebiom.2015.07.034
- Wang, D., Li, J. R., Zhang, Y. H., Chen, L., Huang, T., and Cai, Y. D. (2018). Identification of differentially expressed genes between original breast cancer and xenograft using machine learning algorithms. *Genes* 9:E155. doi: 10.3390/genes9030155
- Wang, S., Wang, D., Li, J., Huang, T., and Cai, Y. D. (2017). Identification and analysis of the cleavage site in signal peptide by using SMOTE, Dagging, and feature selection methods. *Mol. Omics* 14, 64–73. doi: 10.1039/c7mo00030h
- Zhang, C., Chi, Y. L., Wang, P. Y., Wang, Y. Q., Zhang, Y. X., Deng, J., et al. (2012). miR-511 and miR-1297 inhibit human lung adenocarcinoma cell proliferation by targeting oncogene TRIB2. *PLoS One* 7:e46090. doi: 10.1371/journal.pone.0046090
- Zhang, H. H., Pang, M., Dong, W., Xin, J. X., Li, Y. J., Zhang, Z. C., et al. (2014). miR-511 induces the apoptosis of radioresistant lung adenocarcinoma cells by triggering BAX. *Oncol. Rep.* 31, 1473–1479. doi: 10.3892/or.2014.2973
- Zhang, Y. H., Huang, T., Chen, L., Xu, Y., Hu, Y., Hu, L. D., et al. (2017). Identifying and analyzing different cancer subtypes using RNA-seq data of blood platelets. *Oncotarget* 8, 87494–87511. doi: 10.18632/oncotarget.20903
- Zhou, X., Xu, X., Wang, J., Lin, J., and Chen, W. (2015). Identifying miRNA/mRNA negative regulation pairs in colorectal cancer. *Sci. Rep.* 5:12995. doi: 10.1038/srep12995

**Conflict of Interest Statement:** XC was employed by company Rongze Biotechnology Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Chen, Jin and Feng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.