Current Genomics, 2022, *23*, 109-117

RESEARCH ARTICLE

Prospective Analysis of Proteins Carried in Extracellular Vesicles with Clinical Outcome in Hepatocellular Carcinoma

Wenbiao Chen[1,2,3,#], Feng Zhang[4,#], Huixuan Xu[3], Xianliang Hou[3] and Donge Tang[3,*]

[1]*Central Molecular Laboratory, People's Hospital of Longhua, The Affiliated Hospital of Southern Medical University, Shenzhen, 518109, China;* [2]*Department of Respiratory Medicine, People's Hospital of Longhua, The Affiliated Hospital of Southern Medical University, Shenzhen, 518109, China;* [3]*Clinical Medical Research Center, Guangdong Provincial Engineering Research Center of Autoimmune Disease Precision Medicine, The First Affiliated Hospital of Southern University of Science and Technology, The Second Clinical Medical College of Jinan University, Shenzhen People's Hospital, Shenzhen, 518020, China;* [4]*Intensive Care Unit, The First Affiliated Hospital of Jinan University, Guangzhou, 510632, China*

ARTICLE HISTORY

Received: October 26, 2021
Revised: December 26, 2021
Accepted: January 30, 2022

DOI:
10.2174/1389202923666220304125458

Abstract: ***Background*:** Extracellular vehicles (EVs) contain different proteins that relay information between tumor cells, thus promoting tumorigenesis. Therefore, EVs can serve as an ideal marker for tumor pathogenesis and clinical application.

***Objective*:** Here, we characterised EV-specific proteins in hepatocellular carcinoma (HCC) samples and established their potential protein-protein interaction (PPI) networks.

***Materials and Methods*:** We used multi-dimensional bioinformatics methods to mine a network module to use as a prognostic signature and validated the model's prediction using additional datasets. The relationship between the prognostic model and tumor immune cells or the tumor microenvironment status was also examined.

***Results*:** 1134 proteins from 316 HCC samples were mapped to the exoRBase database. HCC-specific EVs specifically expressed a total of 437 proteins. The PPI network revealed 321 proteins and 938 interaction pathways, which were mined to identify a three network module (3NM) with significant prognostic prediction ability. Validation of the 3NM in two more datasets demonstrated that the model outperformed the other signatures in prognostic prediction ability. Functional analysis revealed that the network proteins were involved in various tumor-related pathways. Additionally, these findings demonstrated a favorable association between the 3NM signature and macrophages, dendritic, and mast cells. Besides, the 3NM revealed the tumor microenvironment status, including hypoxia and inflammation.

***Conclusion*:** These findings demonstrate that the 3NM signature reliably predicts HCC pathogenesis. Therefore, the model may be used as an effective prognostic biomarker in managing patients with HCC.

Keywords: Proteins, exosome, hepatocellular carcinoma, prognostic signature, gene, PPI networks.

1. INTRODUCTION

HCC is a common malignancy of the digestive system, particularly in Asia. Due to hepatitis virus infections, HCC morbidity and mortality have remained persistently high [1]. The burden of HCC is complicated because the clinical manifestations of HCC are hidden in the early stages, which means that the majority of diagnoses occur in the later stages. Late-stage diagnoses results in a poor prognosis and a high rate of post-treatment recurrence, which puts patients' lives at risk [2]. Presently, early surgical intervention is the most effective treatment for HCC. Following early diagnosis and treatment, the 5-year survival rate is approximately 70%, compared to 16% in patients with advanced stages [3]. Therefore, early diagnosis and treatment of HCC patients, and accurate monitoring of disease progression, are critical in improving the overall survival rate. However, there are currently no reliable biomarkers for the early detection and management of HCC.

Extracellular vesicles (EVs) are tiny membranous vesicles, secreted by nearly all cell types and released into the

*Address correspondence to this author at the Clinical Medical Research Center, Guangdong Provincial Engineering Research Center of Autoimmune Disease Precision Medicine, The First Affiliated Hospital of Southern University of Science and Technology, The Second Clinical Medical College of Jinan University, Shenzhen People's Hospital, Shenzhen, 518020, China; Tel: +86 0755-25533018; Fax: +86 0755-25533018;
E-mail: donge66@126.com
 # These authors contributed equally to this work.

1875-5488/22 $65.00+.00

extracellular matrix [4]. Previously, EVs were believed to be responsible for cellular waste disposal. However, research indicates that vesicles function as messengers and are important mediators of cell signal transduction [5]. EVs play an important role in tumor development, metastasis, and progression [6]. EVs secreted by tumor cells contain oncogene virus-derived molecules, pathogenic genetic materials, protein components, and other substances [7]. Interactions between these EVs and other cells promote tumorigenesis in new cells. Furthermore, EVs promote tumor cell diffusion by creating a favorable microenvironment at local and distant metastatic sites [8]. Therefore, EVs contain the overall molecular information of the tumor cells, making them ideal biomarkers in clinical applications.

To date, no large population studies have been conducted to assess changes in EV biopsies across clinical stages and disease markers [9]. Technically, distinguishing between tumor-specific EVs from those released by normal tissues has remained a significant challenge due to a lack of specific biomarkers [10]. EVs carry a variety of small molecules, including proteins, non-coding RNA, or enzymes, that exhibit varying cellular expression levels in tumor cells [11]. It has been demonstrated that EVs containing proteins-*EGFR* extracted from tumor cells stimulate the formation of a liver-like microenvironment and promote liver-specific metastasis [12]. HCC-derived EVs-related protein-*ANGPT2*, on the other hand, was delivered to human umbilical vein endothelial cells by exosome endocytosis, resulting in enhanced angiogenesis [13]. Therefore, we hypothesised that proteins carried by EVs could serve as potential HCC-specific biomarkers.

In this study, we used a large sample size to examine the protein profiles carried by EVs in HCC. We established a reliable HCC prognostic signature using a PPI network. Validation experiments demonstrated that the 3NM signature was a highly predictive biomarker for disease progression in HCC patients. Additionally, the positive correlation between the 3NM and tumor immune cells and tumor microenvironment suggested that the 3NM signature was involved in HCC pathogenesis. Fig. (**1**) depicts the overall flow diagram of this study.

2. METHODS

2.1. Data Preparation and Preprocessing

The HCC samples were collected from three datasets; transcript sequencing data from TCGA-LIHC (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga), Japanese transcript sequencing data from ICGC-LIHC-JP (https://dcc.icgc.org/), and chip sequencing data of the Singapore Bioinformatics Centre from GEO-GSE76427 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE76427). The three distinct cohorts comprised two from different populations in Asia and Europe. The TCGA-LIHC data were used as a training dataset in this study, while the ICGC-LIHC-JP and GEO-GSE76427 data were used as validation datasets. We excluded samples with an overall survival (OS) of less than 30 days and insufficient clinical information. We analysed 317, 161, and 94 HCC samples from TCGA-LIHC, ICGC-LIHC-JP, and GSE76427, respectively (Supplementary Table **S1**). The clinical information of the three cohorts is summarised in Supplementary Table **S2**. Probe values (log_2 intensity) and probe annotations were extracted from the downloaded files. The Z-transform was used to normalise the Fragments Per Kilobase per Million (FPKM) data from the TCGA-LIHC, ICGC-LIHC-JP, and GEO-GSE76427 datasets. The TCGA-LIHC read counts were used to compare the differential expression of coding genes in tumor and normal tissues. Furthermore, the EV-related proteins associated with HCC were collected from exoRBase (http://www.exorbase.org/exoRBase/browse/tomRNAIndex). There were 17211 EV-specific HCC proteins in the exoRBase database.

2.2. Establishment of the PPI Network

The EBSeq R package analysed the differentially-expressed proteins between tumor and normal tissues. We identified 1134 proteins having a posterior probability of differential expression (PPED) >0.99. 1134 differentially-expressed proteins and 7211 EVs-specific HCC proteins were then used to perform intersection operations. A total of 437 proteins highly expressed by extracellular vesicles in HCC were used to conduct PPI analysis and construct a human PPI network in STRING (https://string-db.org). STRING experiments were executed using the Bioconductor R package.

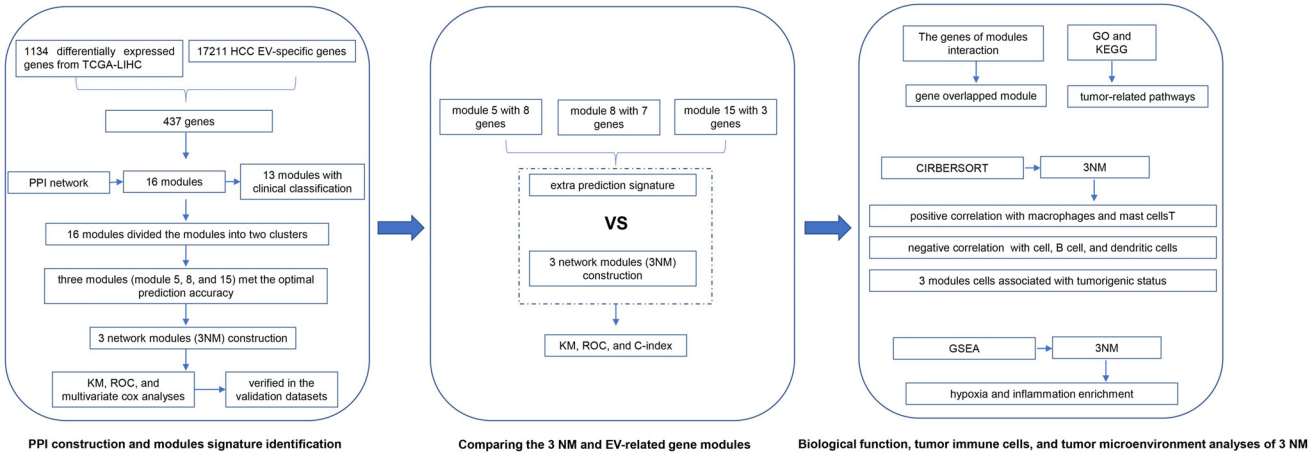

Fig. (1). The Flow chart of this study. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

2.3. Selection of a Network Module Signature

The Molecular Complex Detection (MCODE) algorithm was used to identify significant modules based on the PPI network. The algorithm was executed using the Cytoscape software, with a degree cutoff=2, node score cutoff=0.2, K-score=2, and Max-depth =100. To identify the modules with a significant clinical classification function, we integrated the proteins' profile of the training dataset (TCGA-LIHC) with the PPI network and then calculated the expression score (e) of each module using the following formula:

$$e_j = \sum_{i}^{m} \frac{Z_{ij}}{\sqrt{m}},$$

where m is the number of proteins in a given module M, j is the sample of expression score of module M, and Z_{ij} is the protein expression value of proteins i through Z in the transformation conversion. Next, the mutual information (MI) between e and clinical classification (c) was determined by discriminative score $\mathrm{S(M)}$ as shown:

$$\mathrm{S(M)} = \mathrm{MI}(e',\ c) = \sum_{X}^{e'} \sum_{Y}^{c} p(X,Y)\, log \frac{p(X,Y)}{p(X)p(Y)},$$

Where e' is the discretised form of e, calculated by discretising the expression score=9 ($\log_2$(n)+1) (n is the number of the sample) [14]. X and Y represent the enumerative values of e and c, respectively. $p(X,Y)$ is the joint probability density function (PDF) of e' and c. $p(X)$ and $p(Y)$ are the marginal pdf's of e and c. The greedy searching program was used to identify the modules with a locally maximal stage score [14].

We randomly selected proteins with the same number of M modules from the PPI network to calculate the MI value. Each module was randomly repeated 1000 times. The results from the random sampling and the actual module outcomes were statistically analyzed, and the modules with a significance level of (P<0.0001) were further processed. To construct the prognosis prediction signature, we used the random forest (RF) algorithm in the randomForest R package to score the candidate module protein expression, then used for feature selection and model establishment. Briefly, the predictive value of each candidate module was estimated using an initial RF of 5000 trees. The stepwise adverse selection was used to determine the optimal combination of the candidate modules for prognostic prediction. Each iteration excluded 10% of the features, and the remaining features were used to construct an RF with 3000 trees. The program was terminated when only two functions remained. The RF model with the smallest amount of features was selected among all iteration outcomes. Finally, the last RF model included an out of bag (OOB) probability to describe patient prognosis.

2.4. Identification of Tumor Immune Cells and Tumor Microenvironment Status

CIBERSORT method [15] was used to calculate the abundance of immune and non-immune cells in the tumor microenvironment. The calculation was based on the protein profiles on tissue infiltrating cell populations. The standard-annotated proteins were uploaded to the CIBERSORT website, and the calculation was performed using LM22 signatures and 1000 permutations. Previous studies revealed the signature of tumor microenvironment status, including hypoxia [16], angiogenesis [17], and inflammation [18]. These tumor cell signature scores were derived using the corresponding proteins' average Z- normalised expression values. Pheatmap R package was used to create cluster maps. Additionally, using the GSVA R package, the tumor microenvironment status between subgroups classified by module signatures was analysed using the Gene Set Enrichment Analysis (GSEA).

2.5. Statistical Analysis

The clusterProfiler R package was used to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses the module's signature identified from the EVs-related proteins. The $p < 0.05$ was set as the threshold for statistical significance. The OS was estimated using the Kaplan–Meier (KM) method, while the sensitivity and specificity of the survival curve were assessed using the receiver operating characteristic (ROC) curve. The ROC curve was obtained from calculations of the area under the curve (AUC) using the pROC R package. A hypergeometric algorithm was used to determine the overlapping proteins across the significant signature modules. Multivariate cox analysis was performed to determine the correlation between signature modules and clinical characteristics. On the other hand, variables were compared between groups using the independent Chi-square test. Pearson's correlation was used to investigate the association between immune cells and the tumor microenvironment status signature modules. For OS prediction, Harrell's C-index was used to compare the signatures of modules and proteins. All statistical analyses were performed using the SPSS Version 25.0 software and the R software version 3.4.0, $P < 0.05$.

3. RESULTS

3.1. Construction of the PPI and Identification of Modules Signature

The intersection of 1134 differentially expressed proteins with 17211 HCC EVs-specific proteins resulted in the identification of 437 proteins. The proteins were used to construct the PPI network, which contained 321 proteins and 938 interactions (Supplementary Fig. **1**). The MCODE algorithm was used to identify 16 modules in the PPI network (Supplementary Fig. **2**). Next, the greedy searching program and random sampling identified 13 modules that significantly distinguished clinical classification (P<0.001) (Supplementary Table **S3**). The clustering map for the expression score profile of the 16 modules divided them into two clusters (cluster 1, 2) (Fig. **2A**). Cluster 1 had 3 modules (modules 1, 4, and 6) that contained upregulated proteins in tumor stages I and II, while cluster 2 had 13 modules that contained upregulated proteins in tumor stage III+ IV (P=0.033). Additionally, the 16 modules were associated with the clinic-pathologic classification of HCC, such as T (P=0.043), or grade (P=0.046).

Because the modules were associated with clinical characteristics of HCC, we considered using the 13 modules

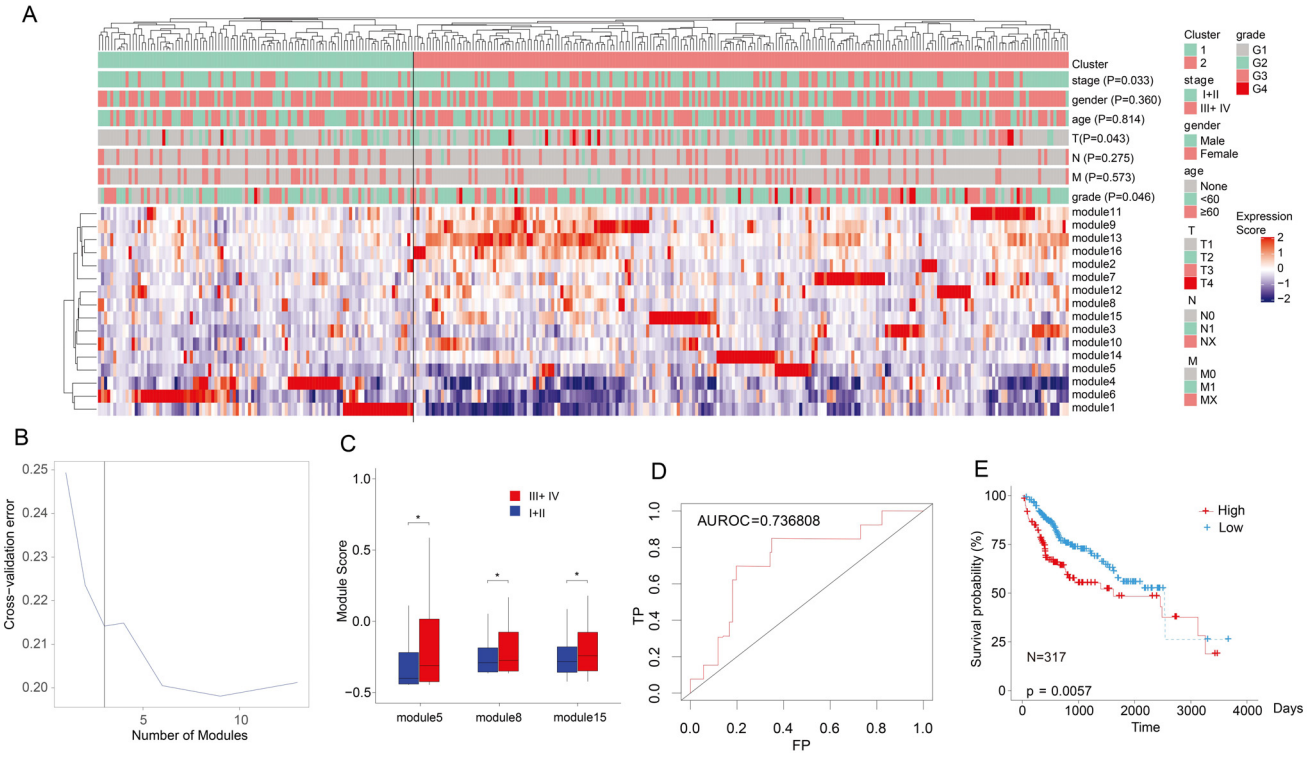

Fig. (2). Modules signature identification based on the PPI network. (**A**) Clustering map of the 16 modules based on the expression score of their profiles. The clustering map also showed the association between the clusters and Clinicopathologic staging. T-N-M: the form of tumor classification represented Tumor- Node- Metastasis. (**B**) Identification of optimal modules using the RF algorithm. The X-axis represented the number of modules, while the Y-axis indicated the cross−validation of each prediction model. (**C**) Comparison of expression scores based on 3 modules between high (III+ IV) and low (I+II) clinical tumor stage. (**D**) The ROC curve was used to validate the 3NM's prognostic prediction (**E**) KM analysis was used to investigate the survival of HCC patients in the TCGA-LIHC database using the 3NM signature. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

with a significant discriminative score to establish an optimal signature for predicting prognosis in HCC patients. We found that combining three modules (modules 5, 8, and 15) resulted in optimal prediction accuracy when using the RF algorithm (Fig. **2B**). In advanced clinical tumor stages (III+ IV), all 3 modules showed higher expression levels compared to early tumor stages (I+II) (Fig. **2C**). The 3 modules were then used to construct a prognostic prediction signature, referred as the 3 network modules (3NM). The ROC curve verified the prognostic prediction potential of 3NM with an AUC of 0.736 (Fig. **2D**).

317 HCC patients from TCGA-LIHC were divided into high or low-risk groups based on the predicted risk coefficient using the 3NM signature. The signature accurately predicted a lower survival rate in the high-risk group compared to the low-risk groups (p = 0.0057) (Fig. **2E**). Additionally, multivariate cox analysis found that the 3NM was a strong independent risk factor for survival (P=0.045) (Supplementary Table **S4**). The prognostic prediction of the 3NM was also validated using validation datasets. The KM analysis revealed that the high-risk group had a significantly shorter OS than the low-risk groups in both the GEO-GSE76427 (p = 0.0039) and ICGC-LIHC-JP (p < 0.0001) datasets (Fig. **3**). These findings demonstrated that the 3NM signature based on EVs-related proteins from the PPI network was a potential biomarker for predicting prognosis in HCC patients.

3.2. Functional Analysis of 3NM and their Corresponding EVs-related Proteins

There were 8, 7, and 3 EV-related proteins, corresponding to modules 5, 8, and 15, respectively. Protein-ISL1 overlapped modules 5 and 8, whereas 4 proteins, (PVALB, NPY, PF4, and SPAG17) were located at the intersection of modules 8 and 15 (Fig. **4A**). GO, and KEGG analyses revealed that all 3 modules were associated with tumor-related pathways, including the toll-like receptor signaling pathway, Wnt signaling pathway, and inflammatory response (Fig. **4B**). These findings inferred that the 3 modules were connected by EVs-related proteins and played a role in HCC pathogenesis.

3.3. Comparing the 3 NM and Exosome-related Gene Modules

As with 3 NM, the remaining 18 EV-related proteins were used to construct an additional prediction signature. In both the training and validation datasets, the developed signature could distinguish survival time between high- and low-risk groups. The high-risk group had a significantly shorter survival time than the low-risk group (Figs. **5A-C**), which corresponded to the prediction performance of the 3NM signature. Unlike the new signature, however, the 3 NM ROC curve showed a higher AUC and C-index (Figs. **5D-G**). These findings demonstrated a superior prediction performance of the 3NM signature.

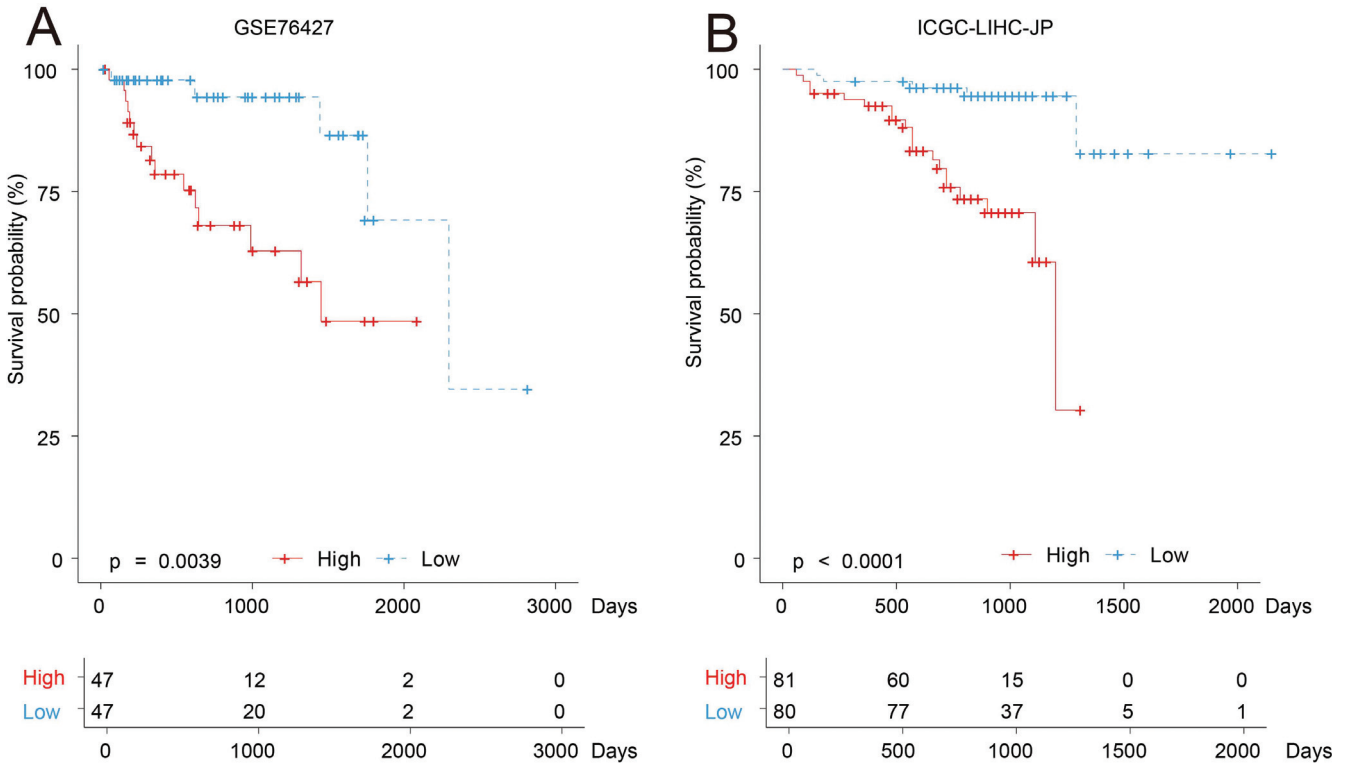

Fig. (3). KM analysis was used to verify the prognostic prediction of HCC patients in (**A**) GEO-GSE76427 and (**B**) ICGC-LIHC-JP datasets based on the 3NM signature. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

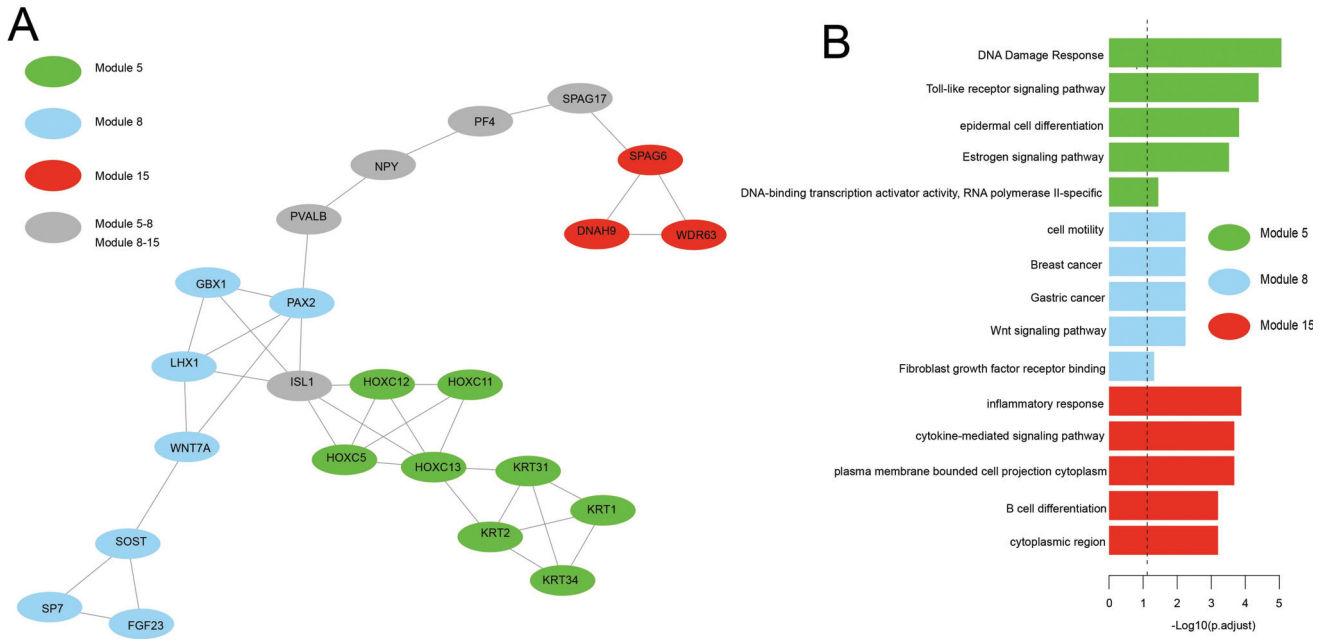

Fig. (4). Functional analysis of the 3 modules. (**A**) The extracellular vesicle-related genes from 3 modules interacted by overlapping with proteins. (**B**) The GO and KEGG analyses of 3 modules. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

3.4. 3 NM is Associated with Tumor Immune Cells and Tumor Microenvironment Status

CIRBERSORT was used to identify the proportion of immune-related cells in the TCGA-LIHC dataset. There was a direct positive correlation between the 3 NM, macrophages, and mast cells. In contrast, various other immune cells, including the T and cell B cell, and dendritic cells, correlated negatively with the 3 modules (Fig. **6A**). Additionally, the 3 modules and their corresponding EVs-related proteins were positively associated with tumorigenic status (hypoxia,

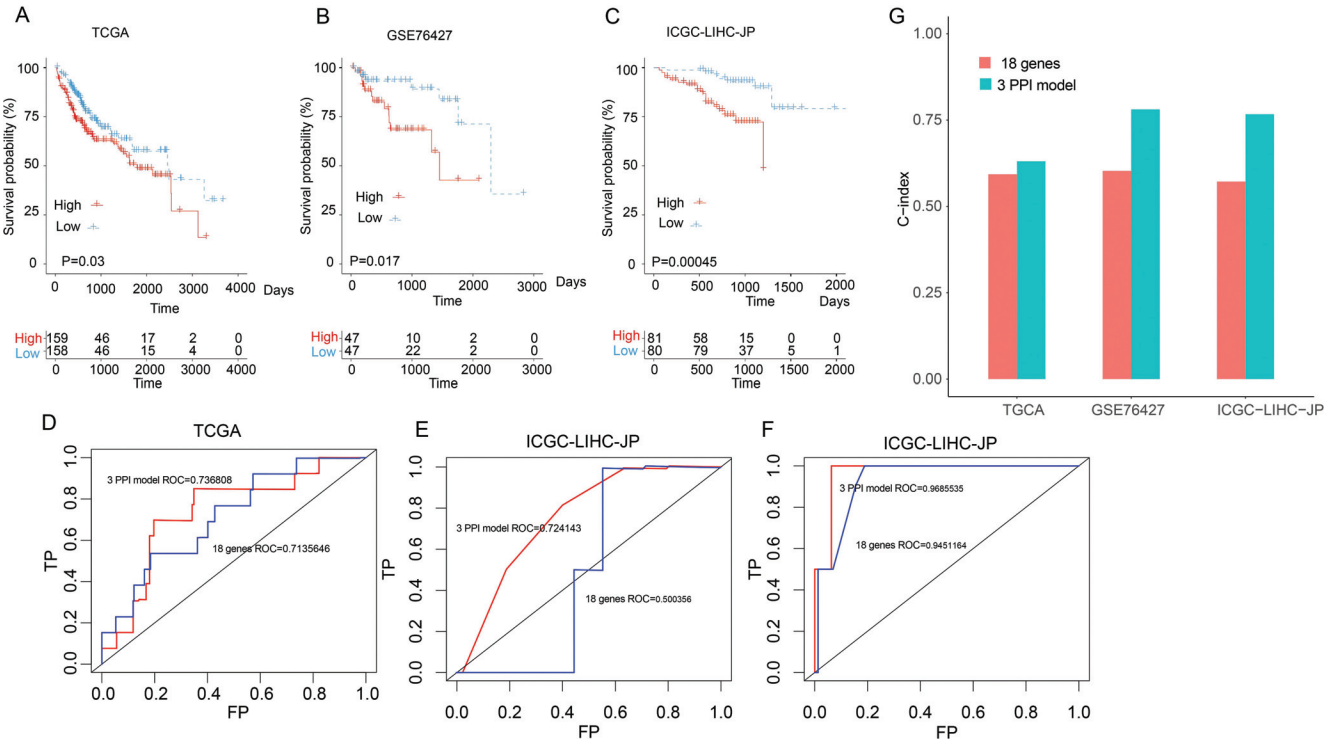

Fig. (5). A comparison of the prognostic prediction ability between 3NM and their corresponding EVs-related proteins. KM analysis was used to investigate the survival of HCC patients in (**A**) TCGA-LIHC, (**B**) ICGC-LIHC-JP, and (**C**) GEO-GSE76427 datasets based on 3NM corresponding EVs-related proteins. ROC curve analysis was used to compare the accuracy of 3NM and their corresponding EVs-related proteins in predicting (**D**)TCGA-LIHC, (**E**)ICGC-LIHC-JP, and (**F**) GEO-GSE76427 datasets HCC patient outcomes. (**G**) The C-index was used to assess the prognostic prediction abilities of 3NM and its corresponding EVs-related proteins. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

HIF1A, angiogenesis, and inflammation) (Fig. **6A**). Hypoxia, HIF1A, angiogenesis, and inflammatory cytokines were significantly more prevalent in the high-risk group than in the low-risk group (Fig. **6B**). HIF1A regulates various pathways in HCC cells that control cellular inflammation, angiogenesis, proliferation, and migration, allowing HCC cells to adapt to low-oxygen environments [19]. Therefore, hypoxia was favorably associated with HIF1A, both of which were upregulated in high-risk groups (Fig. **6C**). GSEA analysis revealed hypoxia and inflammation enrichment in the high-risk group classified by the 3 NM signature (Fig. **6D**). The data established an association between the 3 NM signature and HCC pathogenesis, characterised by immune-related factors and tumor microenvironment status.

4. DISCUSSION

EVs containing nucleic acids and other genetic material are widely distributed in body fluids. They are an ideal candidate for use as noninvasive clinical biomarkers.

It is crucial to identify specific biomarkers from tumor-specific EVs for the early diagnosis and prognosis of cancerous cells [20]. Here, we systematically investigated the EVs-related proteins in HCC patients and developed a prognostic signature for use in the clinical management of HCC. The 3NM signature was found to be a highly predictive biomarker for prognostic prediction in HCC patients. Mechanistically, the 3NM was associated with several tumor-related pathways, tumor immune cells, and tumor microenvironment status. Our research analysed the prognostic value of EVs-related proteins in HCC patients using a large and diverse sample size.

Unlike most research, which used biochemical methods to obtain human EVs, this study used high-throughput sequencing to identify tumor-specific proteins in EVs [21]. The EV-related proteins were obtained from HCC databases in this study [22]. Additionally, the differentially expressed proteins in HCC and normal tissue intersected with the EVs-related proteins from the database. This approach ensured that the selected proteins were accurate and HCC-specific. A previous study reported that EVs-related proteins released into body fluid were proportional to the tumor tissue suggesting that the examination of EVs-related proteins may be used to assess tumor status [23]. We mined EV-related proteins for prognostic prediction of HCC using a multi-dimensional bioinformatics approach. The 3NM signature showed prognostic predictive ability not only in the training cohort but also in the two validation cohorts. This study provided a novel strategy to determine the prognostic significance of specific EV-related proteins in HCC.

The establishment of the 3NM signature was based on constructing a PPI network. The network, coupled with the proteins expression profile, aided in studying the molecular mechanisms of HCC and systematic identification of a reliable biomarker [24]. The 3 modules were inter-connected by

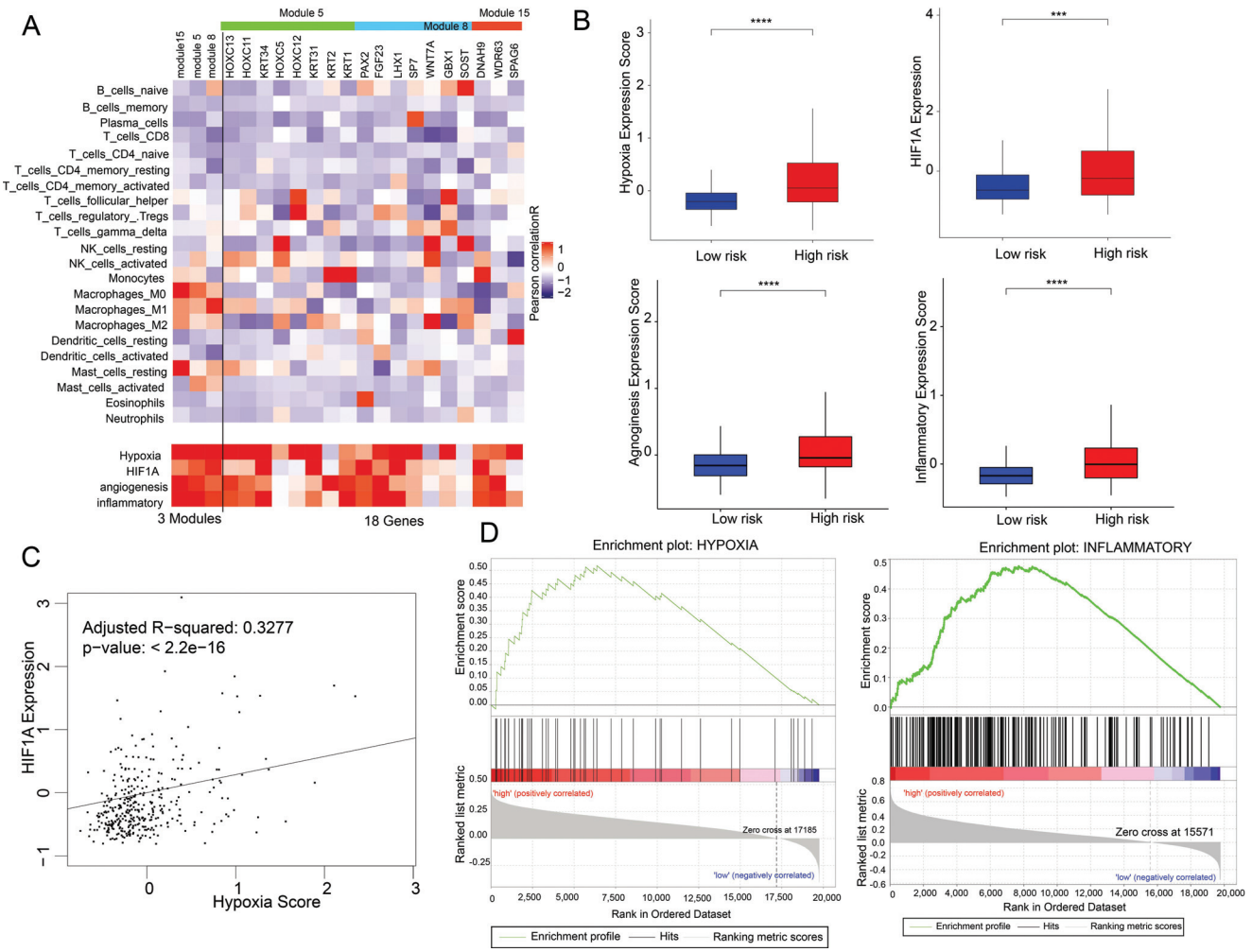

Fig. (6). The association between 3NM, tumor immune cells, and tumor microenvironment status. (**A**) The association between 3NM, tumor immune cells, and tumor microenvironment status. (**B**) Tumor microenvironment status score was used to compare between high and low-risk groups. (**C**) Pearson's correlation algorithm was used to determine the association between HIF1A and hypoxia. (**D**) GSEA was used to determine the hypoxia and inflammatory status of the tumor microenvironment based on the 3NM categorisation. (*A higher resolution / colour version of this figure is available in the electronic copy of the article*).

overlapping proteins, some of which were associated with tumorigenesis. Indeed, several important proteins such as FGF23 [25] in module 8, HOXC5 [26] in module 5, and WDR63 [27] in module 15 were identified as HCC oncogenes. However, using a single protein as a biomarker lacked specificity and stability and was susceptible to external factors. The shortage could be alleviated by classifying protein expression and function and constructing models based on bioinformatics algorithms [28]. Our study constructed the 3NM signature to assess its prognostic predictive value in HCC patients. Additionally, the 3NM was associated with the clinicopathologic staging of HCC. This study uncovered a complex protein expression and PPI interaction network, which provided a strategy for developing a potential HCC prognostic prediction signature.

EVs are a critical component of the tumor microenvironment and contribute significantly to tumorigenesis. Thus, EVs were identified as a novel biomarker for communication mediation in the tumor microenvironment [29]. The 3NM was found to be associated with hypoxia, angiogenesis, and inflammatory reactions. Hypoxia, an important feature in solid tumors, is associated with invasion and a poor prognosis. Previous research has established that hypoxia promotes the secretion of EVs by cancer cells. On the other hand, hypoxic conditions significantly alter EVs, indicating a communication interface between tumor cells and the local or distant microenvironment [30]. Similarly, this study found a direct correlation between hypoxia and all of the 3 modules. Additionally, HIF1A, a key protein involved in inducing hypoxia in extracellular vesicles [31], was also isolated. Angiogenesis and inflammatory reactions were associated with the 3NM and had higher expression levels in high-risk groups. There is growing evidence that EVs promote immunosuppression by altering the composition of different immune cell types, thereby contributing to tumor progression [32]. This study showed that the 3 modules had abnormal immune cell composition and high levels of macrophages (M0, M1, and M2) and mast cells. Overall, these findings suggested that the 3NM signature was associated with HCC pathogenesis, thus might be used as a potential biomarker for HCC progression prediction.

With the rapid advancement of bio-information technology and the rise of precision medicine enabled by big data,

many cancer databases with different functions have been established [33]. The use of modern biological information technology for in-depth mining and analysis of a large number of clinical medical data,provides a new research method for tumor mechanism research and also supports rich data for clinical diagnosis and treatment [33, 34]. The purpose of this study was to determine the value of clinical information about proteins carried by extracellular vesicles in HCC based on big data and bioinformatics methods. Bioinformatics research using big data has identified several biomarkers, such as Zhang *et al.* finding that a three-miRNA signature can serve as an independent prognostic indicator for HCC patients [35], and Du *et al.* revealing that a specific seven-biomarker signature may be clinically useful in predicting HCC prognoses in addition to conventional clinicopathological factors [36]. While the bioinformatics analysis methods for big data are generally the same, the research perspectives vary. Because HCC is associated with mechanisms involving the genome, proteome, metabolism, immunity, apoptosis, and exosome conduction, bioinformatics analysis revealed the etiology of HCC and identified biological targets *via* different mechanisms [37, 38]. Therefore, this study focused on the proteins carried by EVs, which are critical in the signaling of HCC tumor cells. Compared to other studies involving bioinformatics research using large datasets on HCC, the data provides the 3 network modules signature composed of specific proteins, rather than individual proteins. Moreover, comparing 3 NM and exosome-related proteins showed a superior prediction performance of the 3NM signature.

While this study established that 3 NM was a reliable HCC prognostic signature for HCC, the limitation of this study is the lack of comprehensive verification. Due to data constraints, the 3 NM were validated in only 2 additional databases. Additionally, the 3 NM was not validated in actual clinical practice. Thus, the 3 NM is preliminary and will require further research and clinical validation on a large scale in the future. We suggested that the 3 NM identified in this study be combined with clinical indicators of HCC to complement each other and improve the clinical diagnosis and treatment specificity of HCC.

CONCLUSION

In conclusion, this study constructed a 3NM signature for prognostic prediction of HCC patients using used specific proteins carried by EVs. The 3NM was associated with the clinicopathologic stage, tumor immune cells, and tumor microenvironment status. We hypothesised that the 3NM was a sufficiently robust biomarker for clinical application and pathogenesis research in HCC.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

FUNDING

The research was supported by the China Postdoctoral Science Foundation (NO. 2021M703372); the Basic and Applied Basic Research Fund of Guangdong Provincial and Municipal Joint Fund (NO. 21201910240002230); Medical Science and Technology Research Fund of Guangdong Province (B2022103).

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

SUPPLEMENTARY MATERIAL

Supplementary material is available on the publisher's website along with the published article.

REFERENCES

[1] Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.,* **2018**, *68*(6), 394-424. http://dx.doi.org/10.3322/caac.21492 PMID: 30207593

[2] Yang, J.D.; Hainaut, P.; Gores, G.J.; Amadou, A.; Plymoth, A.; Roberts, L.R. A global view of hepatocellular carcinoma: Trends, risk, prevention and management. *Nat. Rev. Gastroenterol. Hepatol.,* **2019**, *16*(10), 589-604. http://dx.doi.org/10.1038/s41575-019-0186-y PMID: 31439937

[3] Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2015. *CA Cancer J. Clin.,* **2015**, *65*(1), 5-29. http://dx.doi.org/10.3322/caac.21254 PMID: 25559415

[4] Kalluri, R.; LeBleu, V.S. The biology**,** function**,** and biomedical applications of exosomes. *Science,* **2020**, *367*(6478), eaau6977. http://dx.doi.org/10.1126/science.aau6977 PMID: 32029601

[5] Hessvik, N.P.; Llorente, A. Current knowledge on exosome biogenesis and release. *Cell. Mol. Life Sci.,* **2018**, *75*(2), 193-208. http://dx.doi.org/10.1007/s00018-017-2595-9 PMID: 28733901

[6] Whiteside, T.L. Tumor-derived exosomes and their role in cancer progression. *Adv. Clin. Chem.,* **2016**, *74*, 103-141. http://dx.doi.org/10.1016/bs.acc.2015.12.005 PMID: 27117662

[7] Kalluri, R. The biology and function of exosomes in cancer. *J. Clin. Invest.,* **2016**, *126*(4), 1208-1215. http://dx.doi.org/10.1172/JCI81135 PMID: 27035812

[8] McAndrews, K.M.; Kalluri, R. Mechanisms associated with biogenesis of exosomes in cancer. *Mol. Cancer,* **2019**, *18*(1), 52. http://dx.doi.org/10.1186/s12943-019-0963-9 PMID: 30925917

[9] Tang, Z.; Li, D.; Hou, S.; Zhu, X. The cancer exosomes: Clinical implications, applications and challenges. *Int. J. Cancer,* **2020**, *146*(11), 2946-2959. PMID: 31671207

[10] Wu, A.Y.; Ueda, K.; Lai, C.P. Proteomic analysis of extracellular vesicles for cancer diagnostics. *Proteomics,* **2019**, *19*(1-2), e1800162. http://dx.doi.org/10.1002/pmic.201800162 PMID: 30334355

[11] Wee, I.; Syn, N.; Sethi, G.; Goh, B.C.; Wang, L. Role of tumor-derived exosomes in cancer metastasis. *Biochim. Biophys. Acta Rev. Cancer,* **2019**, *1871*(1), 12-19. http://dx.doi.org/10.1016/j.bbcan.2018.10.004 PMID: 30419312

[12] Zhang, H.; Deng, T.; Liu, R.; Bai, M.; Zhou, L.; Wang, X.; Li, S.; Wang, X.; Yang, H.; Li, J.; Ning, T.; Huang, D.; Li, H.; Zhang, L.; Ying, G.; Ba, Y. Exosome-delivered EGFR regulates liver microenvironment to promote gastric cancer liver metastasis. *Nat. Commun.,* **2017**, *8*, 15016. http://dx.doi.org/10.1038/ncomms15016 PMID: 28393839

[13] Xie, J.Y.; Wei, J.X.; Lv, L.H.; Han, Q.F.; Yang, W.B.; Li, G.L.; Wang, P.X.; Wu, S.B.; Duan, J.X.; Zhuo, W.F.; Liu, P.Q.; Min, J. Angiopoietin-2 induces angiogenesis *via* exosomes in human hepatocellular carcinoma. *Cell Commun. Signal.,* **2020**, *18*(1), 46. http://dx.doi.org/10.1186/s12964-020-00535-8 PMID: 32183816

[14] Chuang, H.Y.; Lee, E.; Liu, Y.T.; Lee, D.; Ideker, T. Network-based classification of breast cancer metastasis. *Mol. Syst. Biol.,* **2007**, *3*, 140. http://dx.doi.org/10.1038/msb4100180 PMID: 17940530

[15] Newman, A.M.; Liu, C.L.; Green, M.R.; Gentles, A.J.; Feng, W.; Xu, Y.; Hoang, C.D.; Diehn, M.; Alizadeh, A.A. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods,* **2015**, *12*(5), 453-457. http://dx.doi.org/10.1038/nmeth.3337 PMID: 25822800

[16] Buffa, F.M.; Harris, A.L.; West, C.M.; Miller, C.J. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br. J. Cancer,* **2010**, *102*(2), 428-435. http://dx.doi.org/10.1038/sj.bjc.6605450 PMID: 20087356

[17] Masiero, M.; Simões, F.C.; Han, H.D.; Snell, C.; Peterkin, T.; Bridges, E.; Mangala, L.S.; Wu, S.Y.; Pradeep, S.; Li, D.; Han, C.; Dalton, H.; Lopez-Berestein, G.; Tuynman, J.B.; Mortensen, N.; Li, J.L.; Patient, R.; Sood, A.K.; Banham, A.H.; Harris, A.L.; Buffa, F.M. A core human primary tumor angiogenesis signature identifies the endothelial orphan receptor ELTD1 as a key regulator of angiogenesis. *Cancer Cell,* **2013**, *24*(2), 229-241. http://dx.doi.org/10.1016/j.ccr.2013.06.004 PMID: 23871637

[18] Saloura, V.; Zuo, Z.; Koeppen, H.; Keck, M.K.; Khattri, A.; Boe, M.; Hegde, P.S.; Xiao, Y.; Nakamura, Y.; Vokes, E.E. Correlation of T-cell inflamed phenotype with mesenchymal subtype, expression of PD-L1, and other immune checkpoints in head and neck cancer. *J. Clin. Oncol.,* **2014**, *32*(Suppl 15), 6009. http://dx.doi.org/10.1200/jco.2014.32.15_suppl.6009

[19] Wilson, G.K.; Tennant, D.A.; McKeating, J.A. Hypoxia inducible factors in liver disease and hepatocellular carcinoma: Current understanding and future directions. *J. Hepatol.,* **2014**, *61*(6), 1397-1406. http://dx.doi.org/10.1016/j.jhep.2014.08.025 PMID: 25157983

[20] Wang, Z.; Chen, J.Q.; Liu, J.L.; Tian, L. Exosomes in tumor microenvironment: Novel transporters and biomarkers. *J. Transl. Med.,* **2016**, *14*(1), 297. http://dx.doi.org/10.1186/s12967-016-1056-9 PMID: 27756426

[21] Yang, X.X.; Sun, C.; Wang, L.; Guo, X.L. New insight into isolation, identification techniques and medical applications of exosomes. *J. Control Release.,* **2019**, *308*, 119-129. http://dx.doi.org/10.1016/j.jconrel.2019.07.021 PMID: 31325471

[22] Li, S.; Li, Y.; Chen, B.; Zhao, J.; Yu, S.; Tang, Y.; Zheng, Q.; Li, Y.; Wang, P.; He, X.; Huang, S. exoRBase: A database of circRNA, lncRNA and mRNA in human blood exosomes. *Nucleic Acids Res.,* **2018**, *46*(D1), D106-D112. http://dx.doi.org/10.1093/nar/gkx891 PMID: 30053265

[23] Jalalian, S.H.; Ramezani, M.; Jalalian, S.A.; Abnous, K.; Taghdisi, S.M. Exosomes, new biomarkers in early cancer detection. *Anal. Biochem.,* **2019**, *571*, 1-13. http://dx.doi.org/10.1016/j.ab.2019.02.013 PMID: 30776327

[24] Li, T.; Wernersson, R.; Hansen, R.B.; Horn, H.; Mercer, J.; Slodkowicz, G.; Workman, C.T.; Rigina, O.; Rapacki, K.; Stærfeldt, H.H.; Brunak, S.; Jensen, T.S.; Lage, K. A scored human protein-protein interaction network to catalyze genomic interpretation. *Nat. Methods,* **2017**, *14*(1), 61-64. http://dx.doi.org/10.1038/nmeth.4083 PMID: 27892958

[25] Degirolamo, C.; Sabbà, C.; Moschetta, A. Therapeutic potential of the endocrine fibroblast growth factors FGF19, FGF21 and FGF23. *Nat. Rev. Drug Discov.,* **2016**, *15*(1), 51-69. http://dx.doi.org/10.1038/nrd.2015.9 PMID: 26567701

[26] Yan, T.; Ooi, W.F.; Qamra, A.; Cheung, A.; Ma, D.; Sundaram, G.M.; Xu, C.; Xing, M.; Poon, L.; Wang, J.; Loh, Y.P.; Ho, J.H.J.; Ng, J.J.Q.; Ramlee, M.K.; Aswad, L.; Rozen, S.G.; Ghosh, S.; Bard, F.A.; Sampath, P.; Tergaonkar, V.; Davies, J.O.J.; Hughes, J.R.; Goh, E.; Bi, X.; Fullwood, M.J.; Tan, P.; Li, S. HoxC5 and miR-615-3p target newly evolved genomic regions to repress hTERT and inhibit tumorigenesis. *Nat. Commun.,* **2018**, *9*(1), 100. http://dx.doi.org/10.1038/s41467-017-02601-1 PMID: 29311615

[27] Koufaris, C. Human and primate-specific microRNAs in cancer: Evolution, and significance in comparison with more distantly-related research models: The great potential of evolutionary young microRNA in cancer research. *BioEssays,* **2016**, *38*(3), 286-294. http://dx.doi.org/10.1002/bies.201500135 PMID: 26800466

[28] Qi, L.; Chen, L.; Li, Y.; Qin, Y.; Pan, R.; Zhao, W.; Gu, Y.; Wang, H.; Wang, R.; Chen, X.; Guo, Z. Critical limitations of prognostic signatures based on risk scores summarized from gene expression levels: A case study for resected stage I non-small-cell lung cancer. *Brief. Bioinform.,* **2016**, *17*(2), 233-242. http://dx.doi.org/10.1093/bib/bbv064 PMID: 26254430

[29] Milane, L.; Singh, A.; Mattheolabakis, G.; Suresh, M.; Amiji, M.M. Exosome mediated communication within the tumor microenvironment. *J. Control Release.,* **2015**, *219*, 278-294. http://dx.doi.org/10.1016/j.jconrel.2015.06.029 PMID: 26143224

[30] Meng, W.; Hao, Y.; He, C.; Li, L.; Zhu, G. Exosome-orchestrated hypoxic tumor microenvironment. *Mol. Cancer,* **2019**, *18*(1), 57. http://dx.doi.org/10.1186/s12943-019-0982-6 PMID: 30925935

[31] Hu, C.; Chen, M.; Jiang, R.; Guo, Y.; Wu, M.; Zhang, X. Exosome-related tumor microenvironment. *J. Cancer,* **2018**, *9*(17), 3084-3092. http://dx.doi.org/10.7150/jca.26422 PMID: 30210631

[32] Filipazzi, P.; Bürdek, M.; Villa, A.; Rivoltini, L.; Huber, V. Recent advances on the role of tumor exosomes in immunosuppression and disease progression. *Semin. Cancer Biol.,* **2012**, *22*(4), 342-349. http://dx.doi.org/10.1016/j.semcancer.2012.02.005 PMID: 22369922

[33] Dlamini, Z.; Francies, F.Z.; Hull, R.; Marima, R. Artificial Intelligence (AI) and big data in cancer and precision oncology. *Comput. Struct. Biotechnol. J.,* **2020**, *18*, 2300-2311. http://dx.doi.org/10.1016/j.csbj.2020.08.019 PMID: 32994889

[34] Zheng, H.; Zhang, G.; Zhang, L.; Wang, Q.; Li, H.; Han, Y.; Xie, L.; Yan, Z.; Li, Y.; An, Y.; Dong, H.; Zhu, W.; Guo, X. Comprehensive review of web servers and bioinformatics tools for cancer prognosis analysis. *Front. Oncol.,* **2020**, *10*, 68. http://dx.doi.org/10.3389/fonc.2020.00068 PMID: 32117725

[35] Zhang, X.; Ma, L.; Zhai, L.; Chen, D.; Li, Y.; Shang, Z.; Zhang, Z.; Gao, Y.; Yang, W.; Li, Y.; Pan, Y. Construction and validation of a three-microRNA signature as prognostic biomarker in patients with hepatocellular carcinoma. *Int. J. Med. Sci.,* **2021**, *18*(4), 984-999. http://dx.doi.org/10.7150/ijms.49126 PMID: 33456356

[36] Du, X.; Zhang, Y. Integrated analysis of immunity- and ferroptosis-related biomarker signatures to improve the prognosis prediction of hepatocellular carcinoma. *Front. Genet.,* **2020**, *11*, 614888. http://dx.doi.org/10.3389/fgene.2020.614888 PMID: 33391356

[37] Craig, A.J.; von Felden, J.; Garcia-Lezana, T.; Sarcognato, S.; Villanueva, A. Tumour evolution in hepatocellular carcinoma. *Nat. Rev. Gastroenterol. Hepatol.,* **2020**, *17*(3), 139-152. http://dx.doi.org/10.1038/s41575-019-0229-4 PMID: 31792430

[38] Zhang, S.; Zhou, Y.; Wang, Y.; Wang, Z.; Xiao, Q.; Zhang, Y.; Lou, Y.; Qiu, Y.; Zhu, F. The mechanistic, diagnostic and therapeutic novel nucleic acids for hepatocellular carcinoma emerging in past score years. *Brief. Bioinform.,* **2021**, *22*(2), 1860-1883. http://dx.doi.org/10.1093/bib/bbaa023 PMID: 32249290