

A Novel Three-MicroRNA Signature for Predicting Survival in Patients with Nasopharyngeal Carcinoma

Shan-Qiang Zhang (✉ zhangshanqiang@stu.edu.cn)

Yue Bei People's Hospital, Shantou University Medical College <https://orcid.org/0000-0003-3502-9877>

Jun Liu

Yue Bei People's Hospital

Hai-Bin Chen

Shantou University Medical College

Wen-Jie Dai

Yue Bei People's Hospital

Li-Qing Zhou

Yue Bei People's Hospital

Chong-Wei Xie

Yue Bei People's Hospital

Ji-Cheng LI

Yue Bei People's Hospital <https://orcid.org/0000-0002-9631-8931>

Research Article

Keywords: Nasopharyngeal carcinoma, prognosis, microRNA, bioinformatics analysis

Posted Date: June 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-622292/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Journal of Dental Sciences on September 1st, 2021. See the published version at <https://doi.org/10.1016/j.jds.2021.08.017>.

Abstract

This study aims to use integrated bioinformatics technology to dig a predictive miRNA-signature correlated with the prognosis of Nasopharyngeal carcinoma (NPC). Initially, 94 up-regulated and 91 down-regulated differentially expressed microRNAs (DEMs) of NPC were identified in the GEO dataset, and univariate COX regression analysis showed their abnormal expression were significantly associated with poor prognosis of NPC, respectively ($P=0.002$, $P<0.001$ and $P<0.001$). Subsequently, hsa-miR-29c, hsa-miR-30e and hsa-miR-93 identified by random forest algorithm were used to construct a predictive signature through multivariate COX regression analysis. Moreover, PCA, Kaplan-Meier analysis, time-dependent ROC analysis, and univariate and multivariate COX regression analysis demonstrated that there were significant differences in risk score, survival time and the expression of 3 DEMs between the high-risk group and the low-risk group ($P< 0.001$), and the high-risk group had worse OS ($P<0.001$). Furthermore, the average AUC values of 1- to 5-year OS, DFS and DMFS predicted by the signature were all above 0.7, and showed better clinical independence than other commonly used indexes. Eventually, 295 differentially expressed mRNAs were obtained by the intersection of the predicted results of TargetScan and the differentially expressed mRNAs in the datasets. Among them, 56 differentially expressed mRNAs were related to PFS. GO and KEGG enrichment analysis indicated that the poor prognosis of NPC was related to the abnormality of chromosomes, cytokines, and chemokines. Collectively, we constructed a 3-miRNA signature with better independent performance in predicting the prognosis for NPC, which may provide a basis for finding new therapeutic targets of NPC patients.

Introduction

Nasopharyngeal carcinoma (NPC), a type of head and neck malignant tumor derived from the columnar epithelium of nasopharyngeal mucosa, has caused substantial mortality and huge economic burden in Southeast Asian countries over the years due to its insidious onset and highly invasive characteristics [1]. According to the International Agency for Research on Cancer (IARC), there were an estimated 133,354 new cases of NPC and 80,008 NPC-related deaths in 2020 worldwide, showing an increasing trend compared with 2018 [2, 3]. As a country with the highest incidence rate of NPC in the world, in 2018, the corresponding new cases of NPC and NPC-related deaths in China were as high as 60, 558 and 31, 413 respectively, and most of which were confirmed to be locoregionally advanced (LA) NPC at the time of initial diagnosis [1]. One of the predominant reasons for the high mortality of NPC is tumor relapse and distant metastases, which is mainly related to the treatment failure of patients with LA-NPC [4]. Therefore, it is important to strengthen the prognostic evaluation of NPC in order to guide the individualized treatment for NPC patients.

At present, tumor-node-metastasis (TNM) staging system is the main determinant for NPC treatment, but the disadvantages of this method are that it is difficult to predict the prognosis of NPC patients and to identify the biological heterogeneity of the tumor [5]. Although the application of intensity-modulated radiation therapy has been shown to improve the overall survival rate in NPC patients with early stage disease, 30–40% of the patients eventually developed tumor relapse and distant metastasis [6, 7].

Furthermore, the prognosis in a considerable number of LA-NPC patients has not been improved due to the induction chemotherapy or adjuvant chemotherapy, suggesting that the differences in molecular biology or gene of NPC in different individuals may be the main reason for the differences in prognosis of NPC patients [8, 9]. Several studies have explored the influence of Epstein–Barr virus (EBV) DNA load and antibody titer changes on the prognosis of NPC, while others also have analyzed the prognostic evaluation capacity of tissue inhibitor metalloproteinase-2 (TIMP-2), L-selectin (SELL), Matrix metalloproteinase-3 (MMP-3), interleukin-8 (IL8), tumor-infiltrating lymphocytes, as well as cytokines such as cystic fibrosis transmembrane conduction regulator (CFTR) and platelet to lymphocyte ratio (PLR) in NPC [10–14]. However, despite the encouraging answers given by these studies, the current evidences are still not enough to start their clinical trials. Therefore, it is important to explore new molecular markers related to the survival outcome of NPC.

As a class of short (≈ 20 – 22 nucleotides), non-coding RNA molecules, micro RNAs (miRNAs) have been proved to play important roles in the occurrence, development and outcome of tumor [15]. In recent years, with the rapid development of high-throughput technology and bioinformatics algorithm, investigators have found that miRNAs expression in tissue and bodily fluids of NPC patients can not only be used as potential biomarkers for early diagnosis and efficacy prediction, but can also participate in prognostic evaluation of the disease [5, 16]. For instance, Zhou et al [17] found that five differentially expressed miRNAs in NPC tissues were involved in distant metastasis. Liu et al [18] identified 5-miRNA signature in large-scale tissue samples, which were associated with the prognosis of NPC patients with high risk of progression. Lian et al [19] found that miRNA-183 and miRNA-141 were potential risk factors for poor prognosis in NPC patients. However, although the research on miRNAs as prognostic biomarkers of NPC is developing rapidly, the overlap of the results is very limited. So, it is necessary to verify and optimize the research data to provide more accurate basis for its feasibility of clinical transformation.

Gene Expression Omnibus (GEO) database stores a large number of published high-throughput data, which makes it more convenient for investigators to study the role of genes in tumorigenesis [20]. In this study, we constructed a prognostic signature composed of three miRNAs such as, hsa-mir-93, hsa-mir-29c and hsa-mir-30e, based on the GSE32960 dataset from the GEO database. Comprehensive analysis showed that the 3-miRNA signature has a satisfactory independent predictive performance for the overall survival (OS), disease-free survival (DFS), and distant metastasis-free survival (DMFS) in NPC patients.

Materials And Methods

Data acquisition and processing

All datasets used in this study were searched and downloaded from the Gene Expression Omnibus (GEO) database of national center for biotechnology information (NCBI) (<https://www.ncbi.nlm.nih.gov/geo>), including miRNA expression dataset GSE32960, mRNA expression dataset GSE12452, GSE53819 and GSE102349. Among them, the GSE32960 dataset contained two cohorts of healthy and NPC tissue samples, for a total of 330 cases (18 and 312, respectively). The number of normal nasopharyngeal

tissues and NPC tissues in GSE12452 and GSE53819 were 10 and 31, and 18 and 18, respectively, while GSE102349 contained 113 NPC tissue samples alone. Meanwhile, the clinical data of GSE32960 dataset and GSE 102349 dataset were also downloaded. All microarray data were processed for homogenization, and the data of GSE102349 was processed by fragments per kilobase of exon model per million mapped reads (FPKM) and *log2*-transformed for further analyses.

Screening of differentially expressed miRNAs and mRNAs

The limma bioconductor R package was used to screen differentially expressed miRNAs (DEMs) between NPC tissues and normal nasopharyngeal tissues. Considering the limited expression of miRNAs in tissues and analyzing DEMs as much as possible, the P-value [false discovery rate (FDR)] < 0.05 and |fold change (FC)| >1.5 were adjusted as the cut-off criteria to identify DEMs in GSE32960 dataset. Similarly, the P-value (FDR) < 0.05 and |FC| >2 were adjusted as the cut-off criteria to identify differentially expressed mRNAs in GSE12452 and GSE53819 datasets.

Functional enrichment analysis

FunRich (Version 3.1.3), an online tool for functional enrichment and interaction network analysis, was used to identify the pathways involved in the DEMs and the transcription factors that regulate their expression. In addition, the “clusterProfiler”, “enrichplot” and “org.Hs.eg.db” packages were adopted to conduct gene ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of the potential target genes of selected hub DEMs.

Construction and evaluation of 3-miRNA signature

In order to better screen and identify the DEMs related to the prognosis of NPC, 312 patients in GSE32960 dataset were randomly divided into the training set and the testing set by R package “caret”, and the whole set was used to represent the integration of the two sets. Then, the univariate Cox regression analysis was used to identify the DEMs related to OS of patients in the training set, and the random forest algorithm was used to reduce the dimension of the survival data to complete the importance ranking of DEMs. According to relative importance > 0.1, three hub DEMs (hsa-miR-29c, hsa-miR-30e, and hsa-miR-93) were selected for multivariate Cox regression analysis to construct the 3-miRNA prognostic signature. The index of signature was calculated as follows:

Prognostic index (Pi) = (β * expression level of hsa-miR-29c) + (β * expression level of hsa-miR-30e) + (β * expression level of hsa-miR-93).

Subsequently, Kaplan-Meier survival curve and time-dependent ROC curve analysis were used to evaluate and verify the predictive capacity of the 3-miRNA signature for risk identification, OS, DFS and DMFS of patients in the training set, testing set and whole set, respectively. Univariate and multivariate Cox regression analyses were used to demonstrate whether it can be used as a prognostic indicator for NPC independent of other clinical indexes.

Target gene analysis

Since miRNA can target mRNA and regulate gene expression at the post-transcriptional level, and the imbalance of miRNA-mRNA regulatory network is considered to be significantly associated with cancer treatment and prognosis, TargetScan (an online miRNA binding prediction database) was used to predict the target genes of 3 miRNAs. Furthermore, the predicted results were intersected with the differentially expressed mRNAs in GSE12452 and GSE53819 datasets to obtain the potential target genes of 3 miRNAs, and the overlapping results were displayed by drawing the Venn diagram on the bioinformatics & Evolutionary Genomics website (<http://bioinformatics.psb.ugent.be/webtools/Venn/>). Finally, the genes associated with NPC progression free survival (PFS) were screened in GSE102349 dataset by the univariate Cox regression analysis.

Statistical analysis

R software (version 3.6.3) was used for all the statistical analyses. The R statistical packages "survival" and "survminer" were used to plot the Kaplan-Meier survival curve with the median used as a cut-off value, while the package "survivalROC" was used to generate the receiver operating characteristic (ROC) curve of the 3-miRNA signature.

Results

Identification and functional enrichment analysis of DEMs in NPC

The roadmap of the present study is demonstrated in Fig. S1. The clinical characteristics of 312 NPC patients in GSE32960 dataset are listed in Table 1. According to the above screening criteria ($|FC| > 1.5$ and $FDR < 0.05$), a total of 185 miRNAs in NPC tissues reached statistical significance compared with the normal tissues. Among them, 94 significantly up-regulated miRNAs and 91 significantly down-regulated miRNAs are illustrated in the volcanic map (Fig. 1a), while the heat map of these DEMs is shown in Fig. 1b. The enrichment analysis results by FunRich showed that these DEMs were mainly involved in ErbB receptor signaling network, sphingosine 1-phosphate pathway, glypican pathway, plasma Membrane Estrogen Receptor, proteoglycan syndecan-mediated signaling, vascular endothelial growth factor (VEGF) and VEGF receptor signaling network, and other biological pathways, as shown in Fig. 1c. In addition, the top 10 upstream transcription factors predicted by FunRich software, namely SP1, EGR1, SP4, POU2F1, RREB1, NFIC, SOX1, TFAP4, E2F1 and RORA, may have regulatory effects on these 185 DEMs, as shown in Fig. 1d.

Construction of 3-miRNA signature and risk identification of NPC patients

The results of univariate Cox regression analysis in the training set revealed that 17 DEMs were associated with OS in NPC patients (Fig. S2). By using random forest algorithm to evaluate the importance of these DEMs, three miRNAs with relative importance > 0.1 , such as hsa-miR-29c, hsa-miR-30e and hsa-miR-93, were selected for future analysis. The relationship between the error rate and the

number of classification trees and the out of bag importance of the 3 miRNAs is shown in Fig. 2a. Multivariate Cox regression analysis was then performed to construct a 3-miRNA signature, the hazard ratios (HRs) and confidence interval of 3 miRNAs were shown in Fig. 2b. The signature risk score was calculated as follows: Risk Score = 0.843* exphsa-miR-93-0.543* exphsa-miR-29c-0.749* exphsa-miR-30e. Next, the 3-miRNA signature was used to calculate the risk score of each patient in the training set, and the patients were divided into the high-risk group and the low-risk group according to the median value. As shown in Fig. 2c, the results showed that there were significant differences in risk score, survival time, expression of three miRNAs and principal component analysis (PCA) between the two groups, indicating that the 3-miRNA signature constructed in this study has a good power to identify the prognostic risk of NPC. As expected, similar results were also observed in the testing set and the whole set, which further increased the credibility of analysis results in the present study (Fig. 2d, e).

Evaluation of the predictive performance for OS and independent prognostic value of 3-miRNA signature in NPC

To evaluate the predictive performance of 3-miRNA signature for OS in NPC patients, Kaplan-Meier survival curves were generated in the training set, the testing set, and the whole set, respectively. The results revealed that the high-risk patients identified by 3-miRNA signature had worse survival probability ($P < 0.001$) (Fig. 3a-c). Subsequently, the time-dependent ROC curve analysis was performed to further evaluate the predictive capacity of 3-miRNA signature in each group. As shown in Fig. 3d, e and f, the AUC values of the signature for predicting 1- to 5-year OS of NPC patients were 0.803, 0.806, 0.793, 0.806 and 0.806 in the training set, 0.981, 0.685, 0.699, 0.685 and 0.685 in the testing set, and 0.838, 0.750, 0.750, 0.752 and 0.756 in the whole set, respectively. These results indicated that the 3-miRNA signature constructed in this study is accurate and stable in predicting OS of NPC patients. Furthermore, the results of univariate and multivariate Cox regression analyses showed that whether in the training set, testing set, or whole set, 3-miRNA signature was superior to any clinical index such as gender, age, and TNM stage in predicting prognosis, suggesting that it is suitable as an independent prognostic indicator for NPC patients (Fig. 3g-i). In addition, these results also implied that hsa-miR-93, hsa-miR-29c, and hsa-miR-30e may play important roles in NPC progression and can be used as potential targets for NPC treatment.

Evaluation of DFS and DMFS prediction performance of 3-miRNA signature in NPC

In order to comprehensively evaluate the prognostic value of 3-miRNA signature in NPC, the Kaplan-Meier survival curves of DFS and DMFS of NPC patients in GSE32960 dataset were further generated. The results revealed that the high-risk patients differentiated by the 3-miRNA signature constructed in this study had worse survival probability ($P < 0.0001$) (Fig. 4a, b). Furthermore, time-dependent ROC curve analysis was performed to evaluate the predictive capacity of 3-miRNA signature for DFS and DMFS in NPC patients. The results indicated that the AUC values of the signature for predicting 1- to 5-year DFS and DMFS of NPC patients were 0.718, 0.723, 0.702, 0.723 and 0.723, and 0.753, 0.774, 0.749, 0.774 and 0.774, respectively (Fig. 4c, d). Moreover, the results of univariate and multivariate Cox regression

analyses also demonstrated that the 3-miRNA signature had better independent prognostic performance in predicting NPC DFS and DMFS than gender, age, TNM stage, radiotherapy interrupt, radiotherapy boosting and concurrent chemoradiotherapy (Fig. 5a, b). These results further corroborated the robustness of prognostic value of the 3-miRNA signature for NPC patients, which is worthy of further study.

Advantages of the 3-miRNA signature

Currently, gender, age and TNM stage are the main references to evaluate the prognosis of NPC patients. Therefore, in order to further confirm the advantages of the 3-miRNA signature constructed in this study in predicting prognosis, time-dependent ROC curves of the comparison between 3-miRNA signature, gender, age and TNM stage in predicting 1-year, 3-year and 5-year OS of NPC patients in the GSE32960 dataset were conducted. As shown in Fig. 6a, b and c, the AUC values of 3-miRNA signature in predicting 1-year, 3-year, and 5-year OS of NPC patients were 0.838, 0.750, and 0.745, respectively. Compared with the clinical indexes mentioned above, the 3-miRNA signature showed better predictive power. In addition, Kaplan-Meier survival curve analyses of the 3 miRNAs revealed that the patients with high has-miR-93 expression had poor prognosis (OS, $P < 0.001$), and low expressions of has-miR-30e and has-miR-29c were associated with poor prognosis (OS, $P < 0.001$ and $P = 0.002$) (Fig. 6d-f). These results demonstrated that the 3-miRNA signature is a better potential prognostic biomarker for NPC compared with the commonly used clinical indexes, and can be used to guide the individualized treatment of NPC.

Analysis of 3-miRNA target genes

Using TargetScan, a total of 9642 target genes for the 3 miRNAs were predicted. Meanwhile, 1322 and 2439 differentially expressed mRNAs were screened in GSE12452 and GSE53819 datasets, respectively. The intersection of overlapping genes between the three sets of data is shown in the Venn diagram, which indicated that 295 differentially expressed mRNAs could be the potential target genes of the 3 miRNAs (Fig. S3a). The results of GO enrichment analysis revealed that the 295 differentially expressed mRNAs were mainly involved in biological processes such as, nuclear chromosome segregation, and endodermal cell differentiation. In terms of cellular components, these potential target genes were mainly enriched the apical part of the cell, endoplasmic reticulum lumen, and apical plasma membrane, while their molecular function were mainly associated with the extracellular matrix structural constituent, cytokine activity, chemokine receptor binding, and extracellular matrix structural constituent conferring tensile strength (Fig. S3b). Furthermore, KEGG analysis indicated that the potential target genes of 3 miRNAs were significantly enriched in the cytokine-cytokine receptor interaction, chemokine signaling pathway, PI3K-Akt signaling pathway, and other pathways associated with tumorigenesis (Fig. S3c). Finally, the univariate Cox regression analysis revealed that there were 56 differentially expressed genes in GSE102349 dataset, which were related to progression-free survival (PFS), and the top 40 genes were illustrated in Fig. S3d.

Discussion

Currently, miRNAs have been demonstrated to play an important role in the radioresistance and chemoresistance of NPC, and can affect the survival outcome of patients [21]. Several studies have explored the potential role of miRNAs as prognostic biomarkers for NPC based on different human biological samples [18, 22, 23]. In spite of the emerging biological importance of miRNAs, their characteristics such as small molecular weight, low expression and high sequence homology, as well as the expensive, time-consuming and complex operation of microarray and quantitative real time polymerase chain reaction (qRT-PCR) can restrict the further clinical transformation research of miRNAs [24, 25]. Therefore, integrative bioinformatics analysis leveraging a number of public datasets has become one of the effective methods to excavate miRNAs with better clinical application potential [26].

In the present study, miRNA expression data from GSE32960 dataset were adopted for differential expression analysis, and 185 DEMs were identified. The miRNA-enrichment analysis showed that these DEMs were involved in various pathophysiological pathways such as tumor proliferation, invasion, metastasis and angiogenesis, indicating that they can play an indispensable role in the prognosis of NPC. Considering that transcription factors and miRNAs are crucial trans-regulatory factors that can play important roles in controlling gene regulation at the transcriptional and post-transcriptional level, the transcription factors that regulate these DEMs were also predicted. As a sequence-specific DNA binding protein, transcription factor specificity protein-1 (SP1) has been proved to be involved in regulating the expression of genes that play a key role in the proliferation and metastasis of a variety of tumor cells, and the high expression level of SP1 protein is often considered as a negative prognostic factor in tumor patients [27]. A recent study showed that exogenous SP1 could inhibit the expression of miR7-5p and could reverse the growth of NPC cells inhibited by solamargine [28]. Ho et al. [29] found that melatonin can regulate the expression of matrix metalloprotein 9 (MMP-9) gene induced by tumor promoter 12-o-tetradecanoylphosphate-3-acetate (TPA) by inhibiting the binding ability of SP1 to DNA, thus inhibiting the metastasis activity of NPC. Our prediction showed that SP1 may regulate nearly 58% of DEMs, which is helpful to further clarify the invasion and metastasis mechanism of NPC, and more experiments are needed to verify it in the future.

For clinicians, the current difficulty in NPC treatment is that some patients often do not respond to radiotherapy or chemotherapy, resulting in tumor progression or distant metastasis due to not being treated in time, leading to poor prognosis [8, 30]. Therefore, we constructed a 3-miRNA signature to evaluate the prognosis of patients with NPC. Different from other studies, we employed random forest algorithm to select hsa-miR-93, hsa-miR-29c and hsa-miR-30e in the training set to form the prognostic signature, and confirmed the accuracy and stability of the signature in the testing set and the whole set. It is known that hsa-miR-93 is overexpressed in many types of cancer, and can target transforming growth factor beta-receptor type II (TGF β R2) and disabled-2 (DAB2) through a variety of signaling pathways to promote the growth and metastasis of NPC cells [31, 32]. Recent studies have shown that upregulated miR-30e-5p can target metastasis-associated protein 1 (MTA1) and ubiquitin-specific protease 22 (USP22), respectively, to reduce the ability of 5-8F cells migration, invasion and metastasis [33, 34]. In addition, hsa-miR-29c can directly target T lymphoma invasion and metastasis inducing factor 1 (TIAM1), inhibit the migration and invasion of NPC cells, and downregulate Integrin Beta 1 (ITGB1) to

enhance the sensitivity of NPC cells to paclitaxel [35, 36]. In this study, we found high expression of hsa-miR-93 and low expression of has-miR-30e and hsa-miR-29c in NPC tissues through data analysis, and Kaplan-Meier curves indicated that patients with this expression pattern tend to have shorter OS. These results are consistent with the findings of other studies mentioned above, which support the feasibility of 3-miRNA signature as the prognostic risk factor for patients with NPC.

By analyzing the risk score, survival time, the expression of three miRNAs, and other data of patients, we found that the signature composed of hsa-miR-93, hsa-miR-29c and hsa-miR-30e could classify patients into the high-risk group and the low-risk group in GSE32960 dataset, and the Kaplan-Meier analysis showed that these identified high-risk patients have worse survival probability, which indicated that the 3-miRNA signature constructed in this study has a better prognosis risk identification ability for NPC patients. Time dependent ROC analysis showed that AUC values of 3-miRNA signature in predicting 5-year OS of patients within training set, testing set and whole set were 0.806, 0.685 and 0.756, respectively, which were better than the 3-miRNA signature constructed by Wang et al [37], 4-miRNA signature constructed by Zhang et al [38], 4-miRNA signature constructed by Zhao et al [39] and 4-miRNA signature constructed by Liu et al [40]. Furthermore, compared with the published signatures mentioned above, the AUC values of the 3-miRNA signature constructed in this study for predicting 1- to 5-year DFS and DMFS of NPC patients were all above 0.7, and showed better power in predicting OS than gender, age and TNM stage, indicating that the 3-miRNA signature constructed in this study has the potential to be an independent prognostic factor for NPC patients.

In the following analysis, the results of target gene prediction showed that 295 differentially expressed mRNAs may be the potential target genes for the three miRNAs, indicating that the dysregulation of these miRNA-mRNA pairs may predict the poor survival outcome of NPC patients. However, since mRNA can also be targeted by a variety of miRNAs, their complete interaction network remains to be further revealed [41]. In the subsequent GO enrichment analysis, the results showed that the biological process and cellular components of these target genes were related to the nuclear chromosome segregation and centromeric region of chromosome, respectively, suggesting that chromosomal abnormalities are important factors for poor prognosis of NPC. At present, studies have shown that chromosomal aberration, loss of heterozygosity, chromosome rearrangements and chromosomal imbalances can play important role in the occurrence and development of NPC [42]. We speculated that during NPC resistance to radiotherapy or chemotherapy, these post-transcriptional target genes were regulated by the 3 miRNAs, resulting in chromosomal rearrangements and causing distant metastasis or recurrence of tumor. KEGG analysis showed that the pathways involved in these target genes, such as PI3K-Akt signaling pathway, small cell lung cancer and osteoclast differentiation, were related to tumor invasion and metastasis, which further confirmed our conjecture [31, 43, 44]. In addition, the molecular function annotation and KEGG analysis showed that these target genes were mainly enriched in cytokine activity, chemokine receptor binding, and cytokine-cytokine receptor interaction and cytokine signaling pathway, which may be due to the reason that these potential target genes promoted the secretion of cytokines and chemokines by NPC cells under the regulation of the 3 miRNAs, or interacted with stromal cells by releasing tumor exosomes, resulting in immune escape and enhanced tumor metastasis [45]. On the

other hand, the 3-miRNA signature constructed in this study may be adopted as a target for NPC immunotherapy. Of course, further research is needed to clarify the mechanism of their interaction with EBV and tumor microenvironment [46].

Although the above analysis has fully demonstrated that our constructed 3-miRNA signature may be a good predictive tool for the survival outcome of patients with NPC, this study still has the following limitations. Initially, because it is difficult to obtain large-scale tissue microarray datasets with complete clinical information, we only adopted the GSE32960 dataset to screen and identify DEMs, and more external validations of the predictive performance of the developed signature is needed as next step. Moreover, the specific mechanism of the 3 miRNAs in NPC needs to be further studied by molecular experiments, and the related lncRNA-miRNA-mRNA competitive endogenous RNA (CeRNA) networks need to be established. Eventually, the performance of 3-miRNA signature constructed in this study needs to be further modified and optimized in combination with clinical research.

Conclusions

Taken together, in this study, three DEMs (hsa-miR-93, hsa-miR-29c and hsa-miR-30e) were screened from NPC tissue microarray dataset by using novel random forest algorithm. The results of transcription factor prediction, prognosis evaluation and target gene prediction demonstrated that the signature, which is composed of the three DEMs, has the potential to be used as an independent prognostic factor for NPC, and its performance is better than the published signatures. The results of this study may lay the foundation for further elucidating the underlying mechanism of miRNAs influencing the prognosis of NPC and exploring new therapeutic targets of NPC to improve the survival outcome of patients.

Declarations

Funding

This work was supported by the Shaoguan Science and Technology Plan Projects in 2020 (Grant no.200812094530421)

Conflicts of interest

The authors declare that there is no conflict of interest.

Availability of data and material

The GSE32960, GSE12452, GSE53819 and GSE102349 datasets used in this study can be queried and downloaded from GEO database. GEO belong to public databases and users can download relevant data for free for research and publish relevant articles.

Authors' contributions

All authors contributed to the study conception and design. Data collection and analysis were performed by Shan-Qiang Zhang and Jun Liu. The first draft of the manuscript was written by Shan-Qiang Zhang. Funding was obtained by Shan-Qiang Zhang. The statistical analysis was performed by Jun Liu. Figure modification was implemented by Wen-Jie Dai, Chong-Wei Xie and Li-Qing Zhou. The manuscript was revised by Ji-Cheng Li and Hai-Bin Chen. All authors read and approved the final version of this manuscript.

Ethics approval

The patients involved in the database have obtained ethical approval. Our study is based on open source data, so there are no ethical issues and other conflicts of interest. The Research Ethics Committee of Yue Bei People's Hospital has confirmed that no ethical approval and informed consent are required.

Acknowledgements

We acknowledge GEO database for providing their platforms and contributors for uploading their meaningful datasets. This work was supported by the Shaoguan Science and Technology Plan Projects in 2020 (Grant no.200812094530421)

References

1. Chen YP, Chan ATC, Le QT, Blanchard P, Sun Y, Ma J. Nasopharyngeal carcinoma. *Lancet*. 2019;394(10192):64–80. [https://doi.org/10.1016/S0140-6736\(19\)30956-0](https://doi.org/10.1016/S0140-6736(19)30956-0).
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021; 71(3): 209–49. <https://doi.org/10.3322/caac.21660>.
3. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394–424. <https://doi.org/10.3322/caac.21492>.
4. Zhang Y, Chen L, Hu GQ, Zhang N, Zhu XD, Yang KY, et al. Gemcitabine and Cisplatin Induction Chemotherapy in Nasopharyngeal Carcinoma. *N Engl J Med*. 2019;381:1124–35. <https://doi.org/10.1056/NEJMoa1905287>.
5. Wen W, Mai SJ, Lin HX, Zhang MY, Huang JL, Hua X, et al. Identification of two microRNA signatures in whole blood as novel biomarkers for diagnosis of nasopharyngeal carcinoma. *J Trans Med*. 2019;17(1):186. <https://doi.org/10.1186/s12967-019-1923-2>.
6. Jin YB, Zhang GY, Lin KR, Chen XP, Cui JH, Wang YJ, et al. Changes of plasma cytokines and chemokines expression level in nasopharyngeal carcinoma patients after treatment with definitive intensity-modulated radiotherapy (IMRT). *PLoS one*. 2017; 2 (2): e0172264. <https://doi.org/10.1371/journal.pone.0172264>.

7. Liu B, Tan Z, Jiang Y, Chen Y, Chen Y, Ling K. Correlation Between the Expression of miR150 and FOXO4 and the Local Recurrence and Metastasis of Nasopharyngeal Carcinoma After Intensive Radiotherapy. *J BUON*. 2018; (6): 1671–8.
8. Liang Y, Li J, Li Q, Tang L, Chen L, Mao Y, et al. Plasma protein-based signature predicts distant metastasis and induction chemotherapy benefit in Nasopharyngeal Carcinoma. *Theranostics*. 2020;10(21):9767–78. <https://doi.org/10.7150/thno.47882>.
9. Janvilisri T. Omics-Based Identification of Biomarkers for Nasopharyngeal Carcinoma. *Dis Markers*. 2015; 2015: 762128. <https://doi.org/10.1155/2015/762128>.
10. Lee HM, Okuda KS, González FE, Patel V. Current Perspectives on Nasopharyngeal Carcinoma. *Adv Exp Med Biol*. 2019;1164:11–34. https://doi.org/10.1007/978-3-030-22254-3_2.
11. Gong D, Li Z, Ding R, Cheng M, Huang H, Liu A, et al. Extensive serum biomarker analysis in patients with nasopharyngeal carcinoma. *Cytokine*. 2019;118:107–14. <https://doi.org/10.1016/j.cyto.2018.04.031>.
12. Tu Z, Chen Q, Zhang JT, Jiang X, Xia Y, Chan HC. CFTR is a potential marker for nasopharyngeal carcinoma prognosis and metastasis. *Oncotarget*. 2016;7(47):76955–65. <https://doi.org/10.18632/oncotarget.12762>.
13. Zhang J, Feng W, Ye Z, Wei Y, Li L, Yang Y. Prognostic significance of platelet-to-lymphocyte ratio in patients with nasopharyngeal carcinoma: a meta-analysis. *Future Oncol*. 2020;16(5):117–27. <https://doi.org/10.2217/fon-2019-0520>.
14. Liu W, Chen G, Zhang C, Liao X, Xie J, Liang T, et al. Prognostic significance of tumor-infiltrating lymphocytes and macrophages in nasopharyngeal carcinoma: a systematic review and meta-analysis. *Eur Arch Otorhinolaryngol*. 2021. doi:10.1007/s00405-021-06879-2.
15. Spence T, Bruce J, Yip KW, Liu FF. MicroRNAs in nasopharyngeal carcinoma. *Chin clin oncol*. 2016;5(2):17. <https://doi.org/10.21037/cco.2016.03.09>.
16. Wang S, Claret FX, Wu W. MicroRNAs as Therapeutic Targets in Nasopharyngeal Carcinoma. *Front Oncol*. 2019;9:756. <https://doi.org/10.3389/fonc.2019.00756>.
17. Zhou W, Chang A, Zhao H, Ye H, Li D, Zhuo X. Identification of a novel microRNA profile including miR-106b, miR-17, miR-20b, miR-18a and miR-93 in the metastasis of nasopharyngeal carcinoma. *Cancer Biomark*. 2020;27(4):533–9. <https://doi.org/10.3233/CBM-190601>.
18. Liu N, Chen NY, Cui RX, Li WF, Li Y, Wei RR, et al. Prognostic value of a microRNA signature in nasopharyngeal carcinoma: a microRNA expression analysis. *Lancet Oncol*. 2012; 13(6): 633–41. [https://doi.org/10.1016/S1470-2045\(12\)70102-X](https://doi.org/10.1016/S1470-2045(12)70102-X).
19. Lian J, Li Y, Yu M. MicroRNA-183 and microRNA-141 are potential risk factors for poor prognosis in patients with nasopharyngeal carcinoma. *Oncol Lett*. 2019; 17(1): 1172–6. <https://doi.org/10.3892/ol.2018.9650>.
20. Zhou X, Du YL, Jin P, Ma F. Bioinformatic analysis of cancer-related microRNAs and their target genes. *Yi Chuan*. 2015;37(9):855–64. <https://doi.org/10.16288/j.ycz.14-439>.

21. Mohammadi C, Gholamzadeh Khoei S, Fayazi N, Mohammadi Y, Najafi R. miRNA as promising theragnostic biomarkers for predicting radioresistance in cancer: A systematic review and meta-analysis. *Crit Rev Oncol Hematol*. 2021; 157: 103183. <https://doi.org/10.1016/j.critrevonc.2020.103183>.
22. Mu JW, Zhou XY, Wang QJ, Han LH, Jiao JB. MicroRNA-141-3p promoted the progression of nasopharyngeal carcinoma through targeting DLC1. *Eur Rev Med Pharmacol Sci*. 2020;24(21):11105–13. https://doi.org/10.26355/eurrev_202011_23597.
23. Zhang Z, Huang J, Wang G, Jin F, Zheng J, Xiao H. Serum miRNAs, a potential prognosis marker of loco-regionally advanced nasopharyngeal carcinoma patients treated with CCRT. *BMC Cancer*. 2020;20(1):183. <https://doi.org/10.1186/s12885-020-6689-7>.
24. Simonson B, Das S. MicroRNA Therapeutics: the Next Magic Bullet? *Mini Rev Med Chem*. 2015; 15(6): 467–74. <https://doi.org/10.2174/1389557515666150324123208>.
25. Ganju A, Khan S, Hafeez BB, Behrman SW, Yallapu MM, Chauhan SC, et al. miRNA nanotherapeutics for cancer. *Drug Discov Today*. 2017;22(2):424–32. <https://doi.org/10.1016/j.drudis.2016.10.014>.
26. Chen L, Heikkinen L, Wang C, Yang Y, Sun H, Wong G. Trends in the development of miRNA bioinformatics tools. *Brief Bioinform*. 2019;20(5):1836–52. <https://doi.org/10.1093/bib/bby054>.
27. Vizcaíno C, Mansilla S, Portugal J. Sp1 transcription factor: A long-standing target in cancer chemotherapy. *Pharmacol Ther*. 2015;152:111–24. <https://doi.org/10.1016/j.pharmthera.2015.05.008>.
28. Wu J, Tang X, Ma C, Shi Y, Wu W, Hann SS. The regulation and interaction of colon cancer-associated transcript-1 and miR7-5p contribute to the inhibition of SP1 expression by solamargine in human nasopharyngeal carcinoma cells. *Phytother Res*. 2020; 34(1): 201–13. <https://doi.org/10.1002/ptr.6555>.
29. Ho HY, Lin CW, Chien MH, Reiter RJ, Su SC, Hsieh YH, et al. Melatonin suppresses TPA-induced metastasis by downregulating matrix metalloproteinase-9 expression through JNK/SP-1 signaling in nasopharyngeal carcinoma. *J Pineal Res*. 2016;61(4):479–92. <https://doi.org/10.1111/jpi.12365>.
30. Zhang SQ, Pan SM, Liang SX, Han YS, Chen HB, Li JC. Research status and prospects of biomarkers for nasopharyngeal carcinoma in the era of highthroughput omics. *Int J Oncol*. 2021;58(4):1. <https://doi.org/10.3892/ijo.2021.5188>.
31. Lyu X, Fang W, Cai L, Zheng H, Ye Y, Zhang L, et al. TGFβR2 is a major target of miR-93 in nasopharyngeal carcinoma aggressiveness. *Mol Cancer*. 2014;13:51. <https://doi.org/10.1186/1476-4598-13-51>.
32. Xu YF, Mao YP, Li YQ, Ren XY, He QM, Tang XR, et al. MicroRNA-93 promotes cell growth and invasion in nasopharyngeal carcinoma by targeting disabled homolog-2. *Cancer Lett*. 2015;363(2):146–55. <https://doi.org/10.1016/j.canlet.2015.04.006>.
33. Hu W, Yao W, Li H, Chen L. MiR-30e-5p inhibits the migration and invasion of nasopharyngeal carcinoma via regulating the expression of MTA1. *Biosci Rep*. 2020;40(5):BSR20194309. <https://doi.org/10.1042/BSR20194309>.

34. Ma YX, Zhang H, Li XH, Liu YH. MiR-30e-5p inhibits proliferation and metastasis of nasopharyngeal carcinoma cells by targeting USP22. *Eur Rev Med Pharmacol Sci.* 2018;22(10):6342–9. https://doi.org/10.26355/eurrev_201810_16045.
35. Liu N, Tang LL, Sun Y, Cui RX, Wang HY, Huang BJ, et al. MiR-29c suppresses invasion and metastasis by targeting TIAM1 in nasopharyngeal carcinoma. *Cancer Lett.* 2013; 329(2): 181–8. <https://doi.org/10.1016/j.canlet.2012.10.032>.
36. Huang L, Hu C, Chao H, Wang R, Lu H, Li H, et al. miR-29c regulates resistance to paclitaxel in nasopharyngeal cancer by targeting ITGB1. *Exp Cell Res.* 2019; 378(1): 1–10. <https://doi.org/10.1016/j.yexcr.2019.02.012>.
37. Wang T, Wu J, Wu Y, Cheng Y, Deng Y, Liao J. A novel microRNA-based signature predicts prognosis among nasopharyngeal cancer patients. *Exp Biol Med (Maywood).* 2021;246(1):72–83. <https://doi.org/10.1177/1535370220958680>.
38. Zhang S, Yue W, Xie Y, Liu L, Li S, Dang W, et al. The four-microRNA signature identified by bioinformatics analysis predicts the prognosis of nasopharyngeal carcinoma patients. *Oncol Rep.* 2019;42(5):1767–80. <https://doi.org/10.3892/or.2019.7316>.
39. Zhao L, Fong AHW, Liu N, Cho WCS. Molecular subtyping of nasopharyngeal carcinoma (NPC) and a microRNA-based prognostic model for distant metastasis. *J Biomed Sci.* 2018;25(1):16. <https://doi.org/10.1186/s12929-018-0417-5>.
40. Liu N, Cui RX, Sun Y, Guo R, Mao YP, Tang LL, et al. A four-miRNA signature identified from genome-wide serum miRNA profiling predicts survival in patients with nasopharyngeal carcinoma. *Int J Cancer.* 2014;134(6):1359–68. <https://doi.org/10.1002/ijc.28468>.
41. Zhao H, Chang A, Ling J, Zhou W, Ye H, Zhuo X. Construction and analysis of miRNA-mRNA regulatory networks in the radioresistance of nasopharyngeal carcinoma. *3 Biotech.* 2020;10(12):511. <https://doi.org/10.1007/s13205-020-02504-x>.
42. Xu Y, Huang X, Ye W, Zhang Y, Li C, Bai P, et al. Comprehensive analysis of key genes associated with ceRNA networks in nasopharyngeal carcinoma based on bioinformatics analysis. *Cancer Cell Int.* 2020;20:408. <https://doi.org/10.1186/s12935-020-01507-1>.
43. Jiang X, Feng L, Dai B, Li L, Lu W. Identification of key genes involved in nasopharyngeal carcinoma. *Braz J Otorhinolaryngol.* 2017;83(6):670–6. <https://doi.org/10.1016/j.bjorl.2016.09.003>.
44. Jiao S, Subudhi SK, Aparicio A, Ge Z, Guan B, Miura Y, et al. Differences in Tumor Microenvironment Dictate T Helper Lineage Polarization and Response to Immune Checkpoint Therapy. *Cell.* 2019;179(5):1177–90. <https://doi.org/10.1016/j.cell.2019.10.029>.
45. Huang SCM, Tsao SW, Tsang CM. Interplay of Viral Infection, Host Cell Factors and Tumor Microenvironment in the Pathogenesis of Nasopharyngeal Carcinoma. *Cancers (Basel).* 2018;10(4):106. <https://doi.org/10.3390/cancers10040106>.
46. Cortez MA, Anfossi S, Ramapriyan R, Menon H, Atalar SC, Aliru M, et al. Role of miRNAs in immune responses and immunotherapy in cancer. *Genes Chromosomes Cancer.* 2019; 58(4):244–53. <https://doi.org/10.1002/gcc.22725>.

Tables

Table 1
Baseline information of patients in GSE32960

| Clinical characteristics | | Total | Percentage (%) |
|--------------------------|------------|-------|----------------|
| GSE32960 in GEO | | 312 | 100 |
| Survival status | Survival | 246 | 78.85 |
| | Death | 66 | 21.15 |
| Age | ≤ 65 years | 291 | 93.27 |
| | >65 years | 21 | 6.73 |
| Gender | Female | 79 | 25.32 |
| | Male | 233 | 74.68 |
| Stage | I | 12 | 3.85 |
| | II | 86 | 27.56 |
| | III | 91 | 29.17 |
| | IV | 123 | 39.42 |
| RT period interruptions | 0 days | 178 | 57.05 |
| | ≥ 1 days | 134 | 42.95 |
| RT boosting | Yes | 149 | 47.76 |
| | No | 163 | 52.24 |
| CCRT | Yes | 268 | 85.90 |
| | No | 44 | 14.10 |

Figures

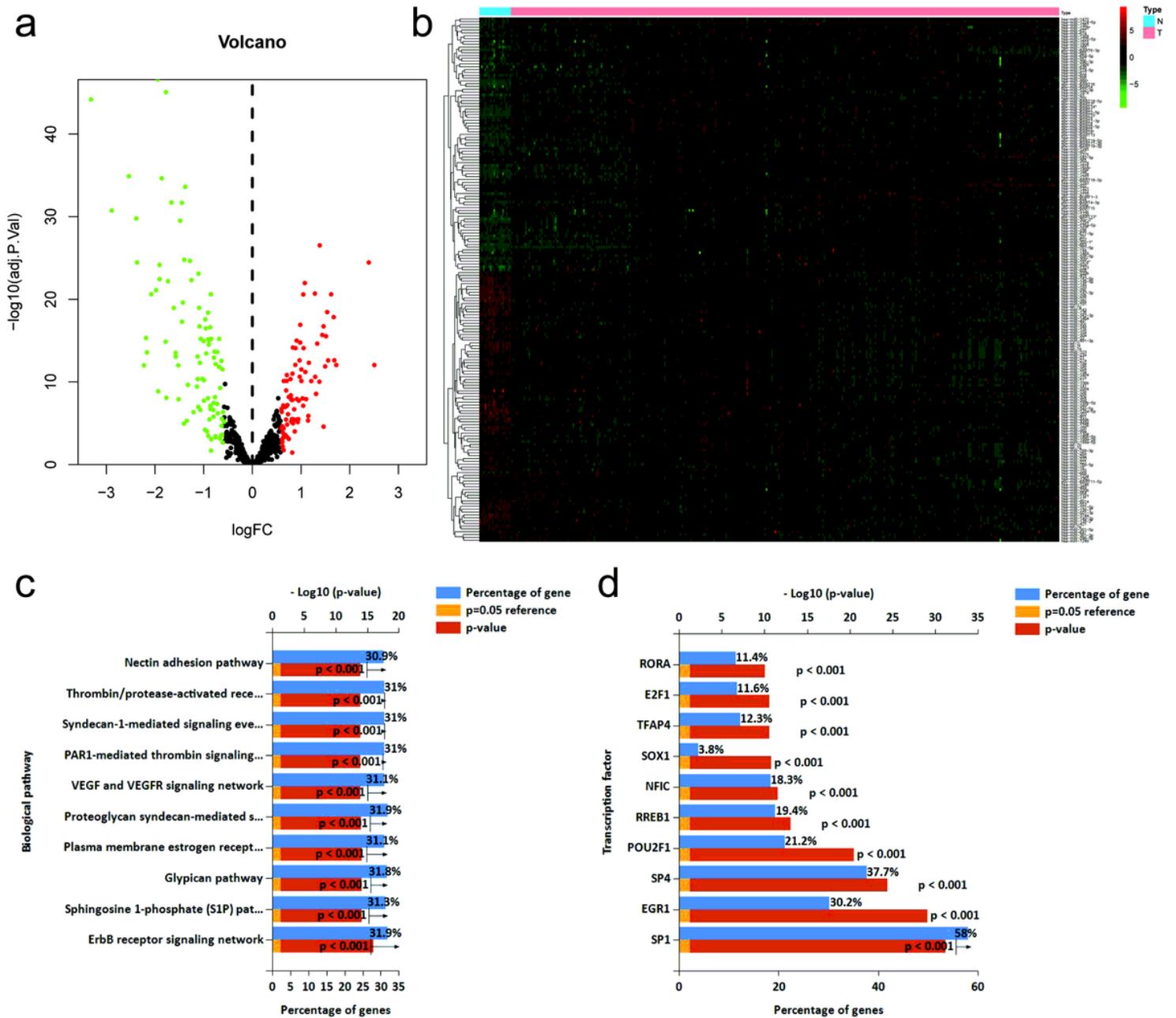


Figure 1

Identification of DEMs and enrichment analysis. a. The volcano of DEMs. b. The heatmap of DEMs. c. Biological pathway analysis for DEMs using FunRich software. d. Transcription factor analysis for DEMs using FunRich software.

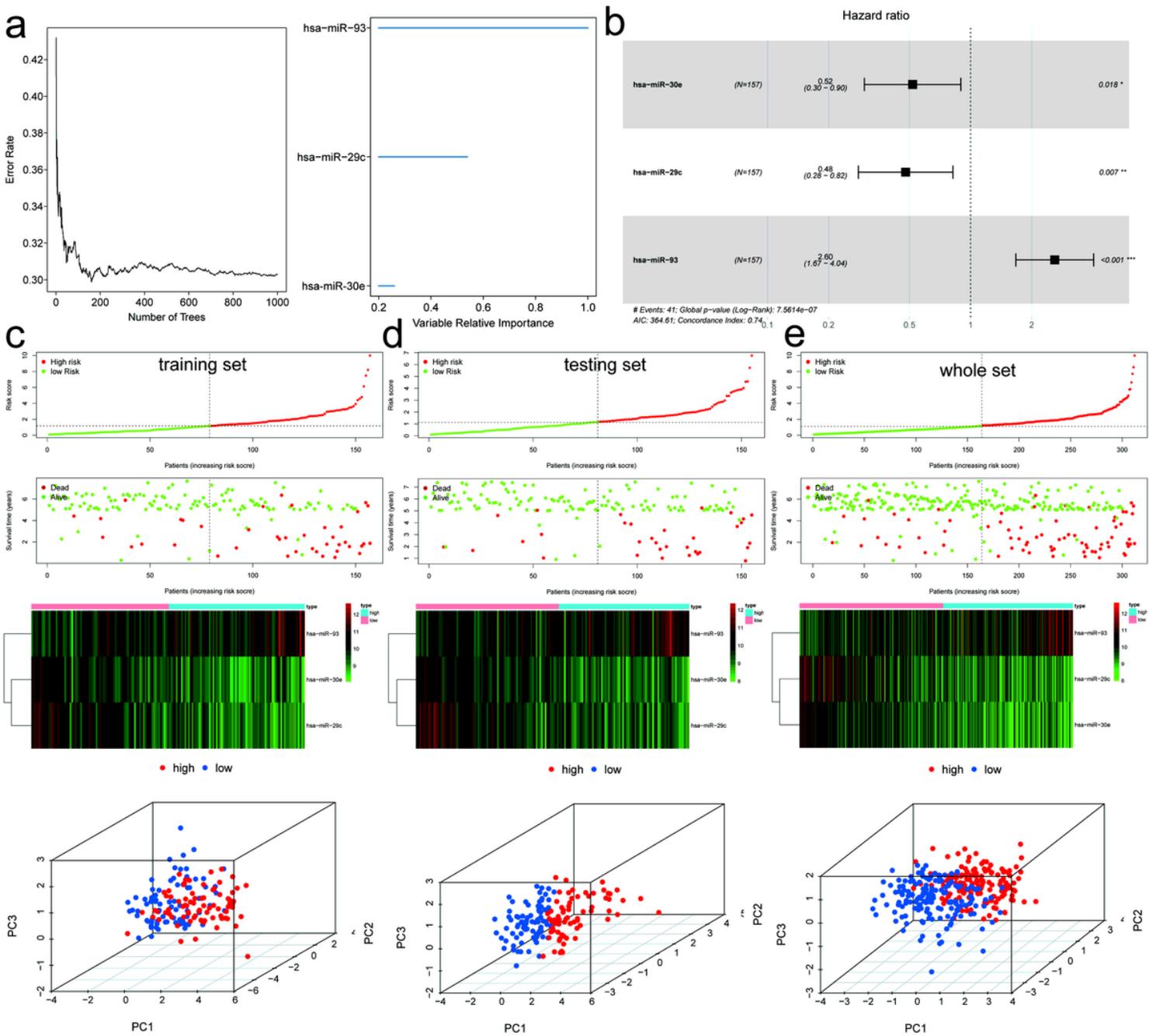


Figure 2

Construction and verification of the signature. a. Decision tree constructed using random forest algorithm and screening of relative importance genes. b. Three-miRNA signature was built by step multivariate Cox regression. Distribution of Risk score, survival time, miRNAs expression, and Principal Component Analysis (PCA) between the high-risk and low-risk groups in training set (c), testing set (d), and whole set (e).

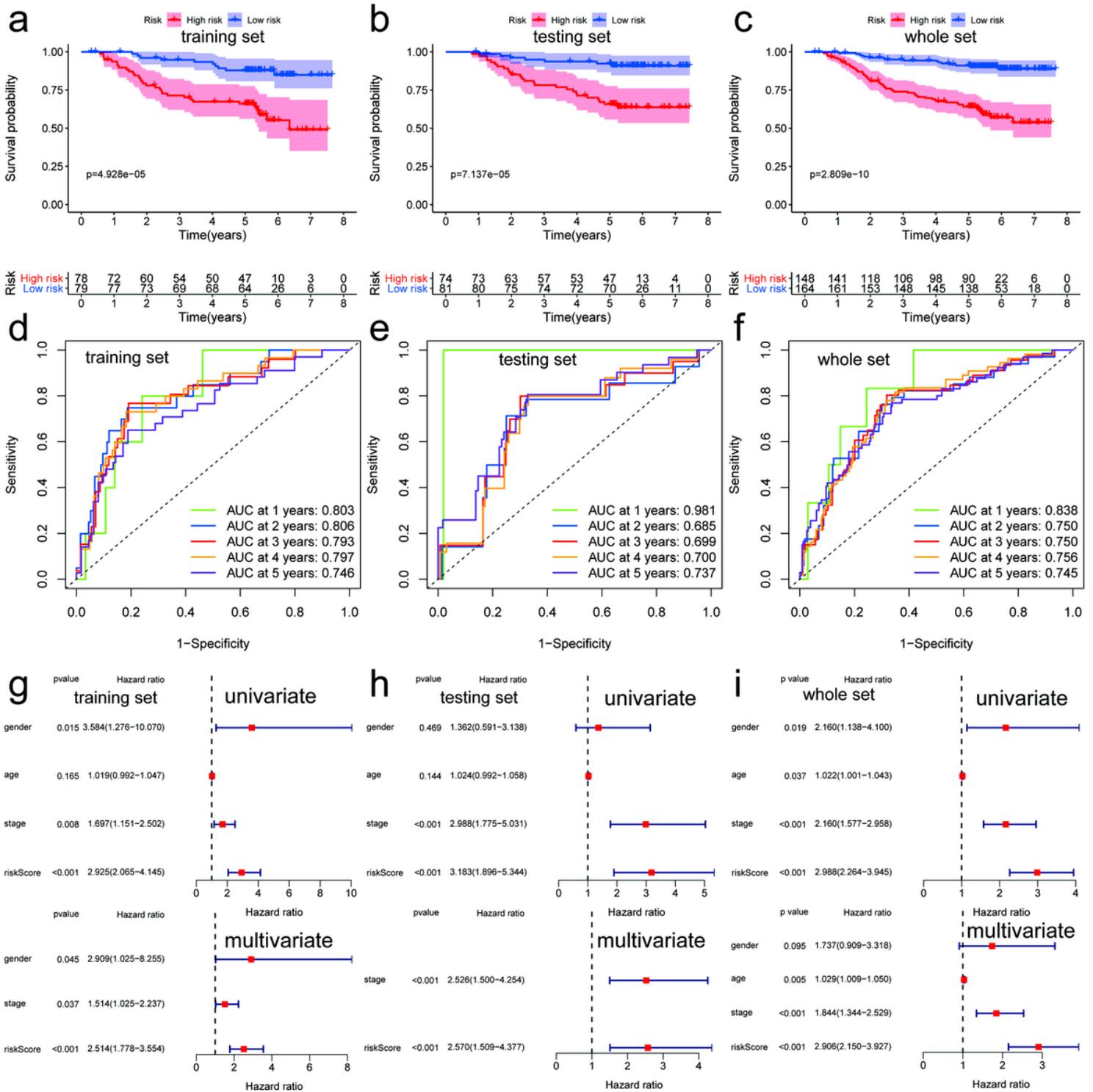


Figure 3

Kaplan-Meier plotter and ROC analysis and Independent prognostic factors determined. Overall survival analysis of the three-miRNA signature in training set (a), testing set (b), and whole set (c). The ROC analysis of the three-miRNA signature for 1-, 2-, 3-, 4-, and 5-years in training set (d), testing set (e), and whole set (f). Independent prognostic factors determined by the Univariate and multivariate Cox regression analyses in raining set (g), testing set (h), and whole set (i).

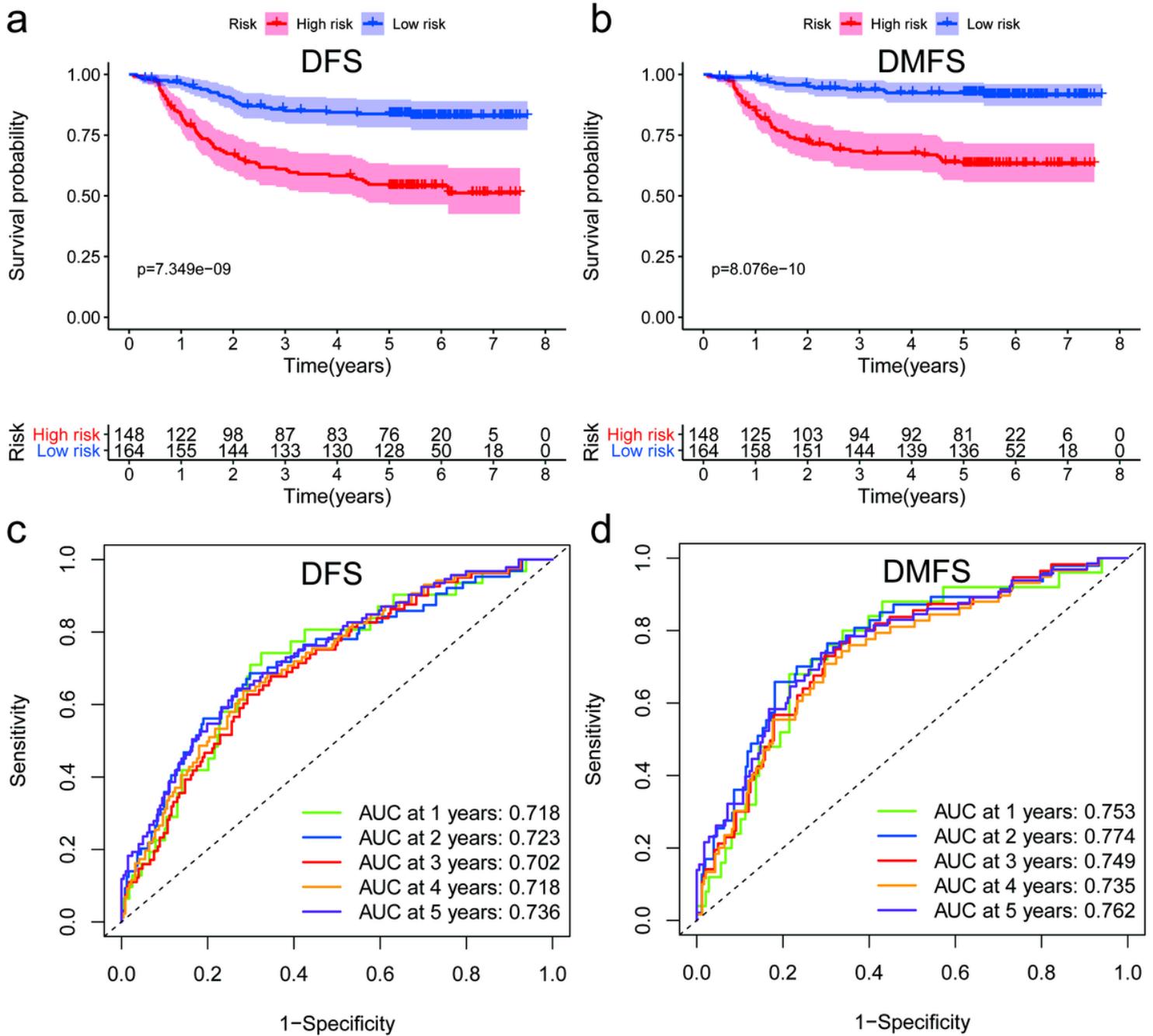


Figure 4

Kaplan-Meier plotter and ROC analysis. Disease-free survival (DFS) (a) and distant-metastasis-free survival (DMFS) (b) analysis of the three-miRNA signature in whole set. ROC curve for the prediction of DFS (c) and DMFS (d) in patients with NPC.

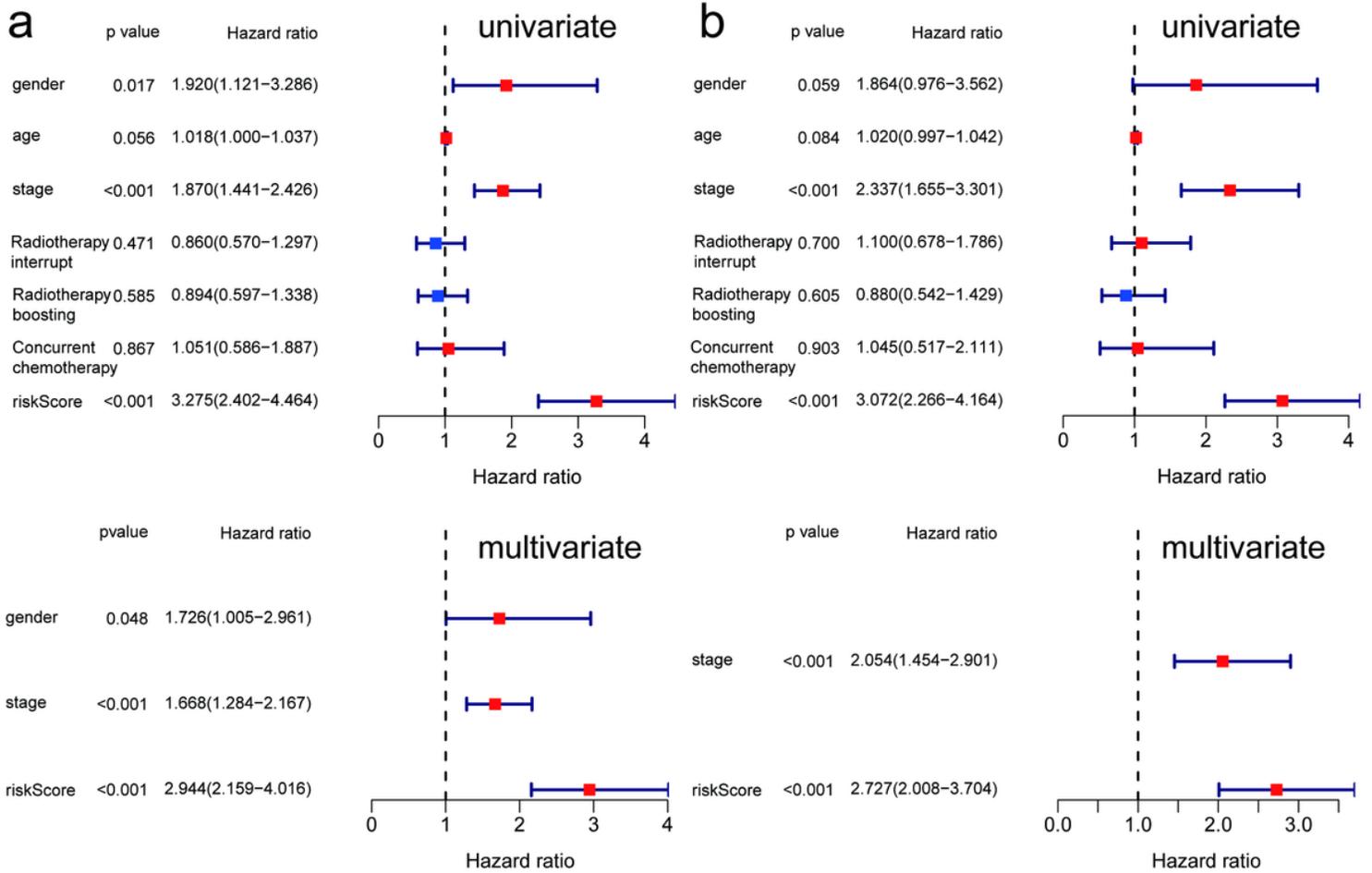


Figure 5

Independent prognostic prediction analysis. The three-miRNA signature was an independent prognostic factor for Disease-free survival (a) and distant-metastasis-free survival (b) by Univariate and multivariate Cox regression analyses.

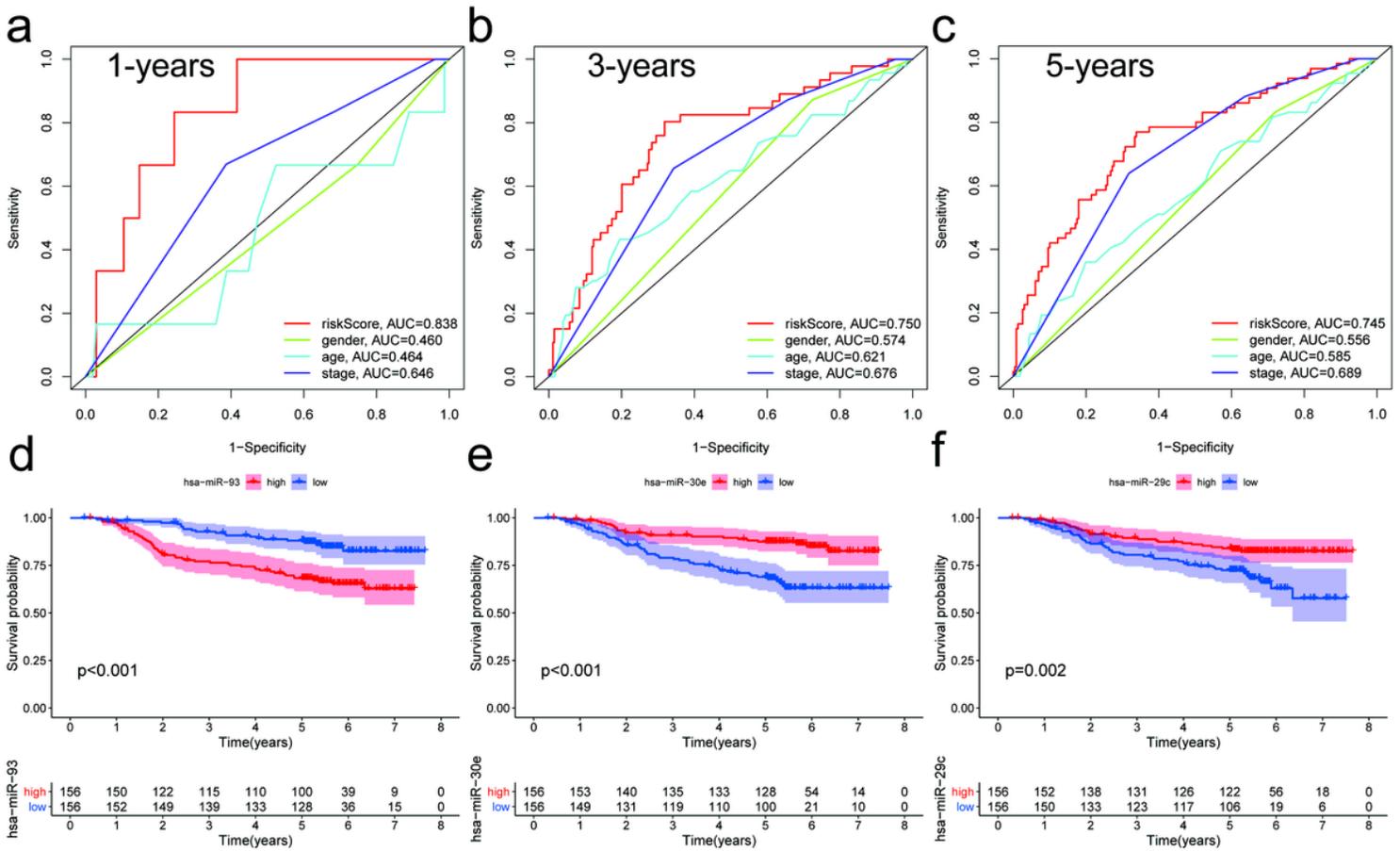


Figure 6

ROC analysis of clinical feature and risk score, and Kaplan-Meier survival curve for three miRNAs. ROC analysis of risk score, gender, age, and TNM stage for 1- (a), 3- (b), and 5-years (c) overall survival in the whole set. The overall survival analysis of has-miR-93 (d), has-miR-30e (e), and has-miR-29c (f) in the whole set.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.tif](#)
- [FigureS2.tif](#)
- [FigureS3.tif](#)