

# Risk stratification of cutaneous melanoma reveals carcinogen metabolism enrichment and immune inhibition in high-risk patients

Xia Li<sup>1</sup>, Yunpeng Cai<sup>1</sup>

<sup>1</sup>Research Center for Biomedical Information Technology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, P.R. China

**Correspondence to:** Xia Li, Yunpeng Cai; **email:** [xia.li@siat.ac.cn](mailto:xia.li@siat.ac.cn), [bioinflux@126.com](mailto:bioinflux@126.com); [yp.cai@siat.ac.cn](mailto:yp.cai@siat.ac.cn)

**Keywords:** cutaneous melanoma, biomarker, driver gene, metabolism, immune cell type

**Received:** April 9, 2020

**Accepted:** July 7, 2020

**Published:** August 28, 2020

**Copyright:** Li et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

Cutaneous melanoma (CM) is the most lethal form of skin cancer. Risk assessment should facilitate stratified surveillance and guide treatment selection. Here, based on the mRNA-seq data from 419 CM patients in the Cancer Genome Atlas (TCGA), we developed a prognostic 21-gene signature to distinguish the outcomes of high- and low-risk patients, which was further validated in two external cohorts. The signature achieved a higher C-index as compared with other known biomarkers and clinical characteristics in both the TCGA and validation cohorts. Notably, in high-risk patients the expression levels of three driver genes, *BRAF*, *NRAS*, and *NF1* in the MAPK pathway, were lower but exhibited a stronger positive correlation as compared with low-risk patients. Moreover, the genes involved in nicotinamide adenine dinucleotide metabolism were negatively correlated with the expression of *BRAF* in the high-risk group. Function analysis revealed that the upregulated genes in the high-risk group were enriched in the cytochrome P450-mediated metabolism of chemical carcinogens. Furthermore, the low-risk group had high levels of gamma delta T cells infiltration, while regulatory T cells were accumulated in the high-risk group. The present study offers a promising new prognostic signature for CM, and provides insight into the mechanisms of melanoma progression.

## INTRODUCTION

Cutaneous melanoma (CM), the most common type of melanoma, is a lethal form of skin cancer. Although it comprises only 3–5% of all skin cancers, it contributes to approximately 75% of all skin cancer-related deaths [1, 2]. Over the last 20 years, the incidence of CM has increased by almost 50% in the United States, from 15/100,000 individuals/year to 22.8/100,000/year [3]. The prognosis of CM patients remains very poor. By the time the tumor has metastasized to the lymph nodes or distant tissues, the 5-year survival rate is only 15–20% [4, 5]. The 10-year survival rate may be as low as 40% when the disease becomes increasingly penetrative of the skin and/or develops local ulceration [6]. Due to its high potential for rapid progression and metastasis [7], risk assessment in CM patients could lead to earlier identification of high-risk disease. It could provide

useful information to evaluate prognosis and facilitate appropriate surveillance for the prevention of recurrence or metastasis. Alternatively, identifying patients with low-risk disease would help to avoid unnecessary treatment costs and reduce anxiety.

Clinicopathological features, such as Breslow thickness (mm), ulceration, and microsatellite metastasis, have been traditionally applied to assess and determine overall risk. However, simple classification of tumors based on phenotypic features does not always represent the intrinsic biology of individual tumors, and is limited in its ability to provide an accurate prediction of individual tumor prognosis [8, 9]. This has led to an interest in the identification of molecular biomarkers that can offer alternative tumor risk stratification, in addition to providing insight into complex tumor cell biology.

Gene expression profiling represents a standard preservation approach for purifying RNA from tumor tissue. It has advanced into the clinical setting to provide a robust and reproducible platform for the simultaneous evaluation of a large number of genes. In recent years, a variety of genes have been proposed as prognostic markers in different types of cancer, such as liver, ovarian, and melanoma [10–13]. A previous study in 2014 identified a panel of genes for use as a gene expression signature to define high- and low-risk groups of melanoma patients [12]. RNA from a total of 40 primary Stage II–III melanomas were analyzed to evaluate the expression of 446 immune- or melanoma-related genes, finally obtaining a 53-immune-related gene signature. The signature was validated in a cohort of 48 Stage II–III melanomas, but further validation studies in other populations have not yet been published. Another gene signature of 31 genes was reported as a prognostic marker in 2015 [13]. The prognostic genes were selected from a comparative review of several previous microarray studies, including genes that were associated with metastasis or exhibited expression differences in primary tumors compared with metastatic tumors. The risk for patients was predicted in 164 melanoma tumors, and the ability to segregate risk has since been assessed in additional retrospective and prospective studies [14, 15]. Although it has important clinical use, additional validation in larger populations could allow better performance of the signature.

However, there exist two major limitations regarding the use of these biomarkers in determining the risk for CM. Firstly, as described above, these biomarkers were obtained using a small sample size with a lack of sufficient validation. Subsequent validation in a greater number of patients should ensure reproducibility and reliability. The second limitation lies in the gene set in which the final prognostic genes are distilled. The current signature was determined from a specific gene list, such as immune- or metastasis-related genes. Of note, biomarkers used for the diagnosis of other cancer types were also obtained in a similar way [10, 11], such as based on the differentially expressed genes between tumor and control samples. An important concern in these approaches is missing other new biomarker candidates from the whole transcriptome. Furthermore, genes with differences in expression between tumor and control groups, or metastatic and primary tumors, may also exhibit notable expression differences among tumor samples or metastatic patients. Thus, selection solely from differentially expressed genes may lack consideration of interpatient heterogeneities, and consequently fail to provide adequate prognostic information regarding patient outcomes. To date, few biomarkers provide sufficient prognostic value;

therefore, the discovery of new markers for the evaluation of patient prognosis is paramount. In addition, the molecular mechanism contributing to an increased risk in CM patients remains poorly understood. Exploration of the molecular differences between risk groups should advance our knowledge of the underlying risk factors. Precise risk stratification and targeting of high-risk factors could be of great benefit to the prevention and treatment of CM.

In the present study, based on the whole transcriptome profiles of 419 patients with CM from the Cancer Genome Atlas (TCGA), we developed a new prognostic signature that successfully classified patients into high- and low-risk groups with significantly different outcomes. Further analysis confirmed that the signature had superior ability in predicting the overall survival (OS) of patients than other known biomarkers and clinical characteristics. The prognostic value of the signature was further validated in two external cohorts. Moreover, we observed a different expression pattern of the driver genes in the two risk groups and the major functional differences existed in the enrichment of metabolic pathways and levels of infiltration of different immune cell types.

## RESULTS

### Identification of high- and low-risk patients using the 21-gene signature

A total of 419 patients with CM, for whom both transcriptome data and clinical information were available, were included in the present study. Among 19,620 protein-coding genes, 1,086 were associated with OS (Log-rank test,  $P < 0.01$ ) according to univariate Cox regression analysis. A lasso-penalized Cox proportional hazards model was applied to choose a gene set with the best prognostic value (Figure 1A). Consequently, a new signature of 21 genes was identified and selected, including *SATB1*, *HN1L*, *CCL8*, *TTC39C*, *HPDL*, *OLFML2A*, *LMNTD2*, *ATP11A*, *SLC5A3*, *HEYL*, *BOK*, *RBCK1*, *CCT6B*, *ABTB1*, *CLEC18A*, *MRPS6*, *NXT2*, *SPEF2*, *KLK13*, *SPAG8*, and *COL22A1* (Supplementary Table 1). None of these 21 genes have been reported as CM biomarkers in previous studies.

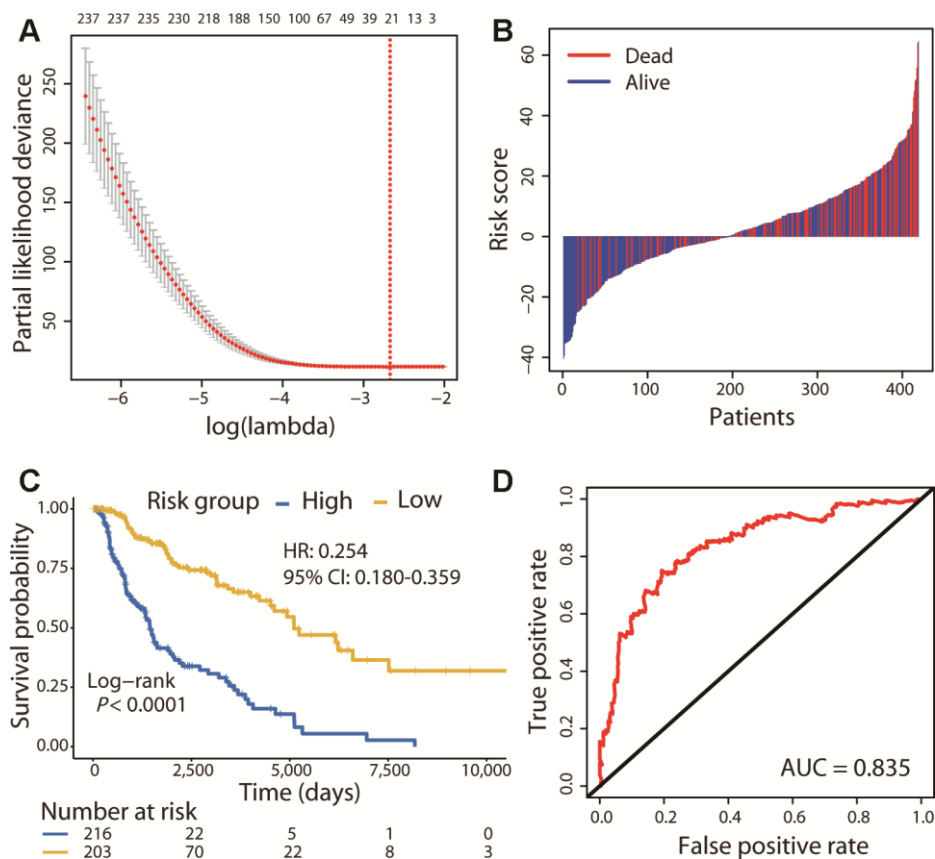
A risk score for each patient was calculated using the expression value of the 21 genes and their regression coefficients from Cox analysis, similar to a previously reported approach [16]. The risk score distribution with survival status was shown in Figure 1B. Patients were divided into high-risk ( $n = 216$ ) and low-risk ( $n = 203$ ) groups using the cut-off value calculated by the survminer R package. Survival analysis found that

patients in the high-risk group had a significantly shorter OS than those in the low-risk group (Hazard ratio [HR]: 0.254, 95% confidence interval [CI]: 0.180–0.359; Log-rank test,  $P < 0.01$ ; Figure 1C). The concordance statistic (C-index) was used to evaluate the predictive accuracy; the C-index of the 21-gene signature for the prediction of OS was 0.679. Moreover, receiver operating characteristic (ROC) analysis was also performed, and the AUC (area under the ROC curve) for the 5-year OS was 0.835 (Figure 1D). These results indicated that our prognostic gene model achieved good performance in distinguishing between high- and low-risk patients.

### Comparison of the 21-gene signature with other known prognostic biomarkers

We compared the patient survival prediction ability of the 21-gene signature with other known biomarkers, 31-

gene [13] and 53-gene [12] signatures identified in previous studies. Moreover, a recent study reported methylation at 4 specific sites as a prognostic biomarker for CM, demonstrating superior ability to predict the OS of patients than numerous known prognostic markers [17]; thus, this methylation signature was also included in our comparison. The 419 CM patients were divided into high- and low-risk groups using each signature, and survival analysis showed that patients in the high-risk group were significantly associated with a poorer OS (Log-rank test,  $P < 0.01$ ; Figure 2A–2C). The C-index of the 31-gene, 53-gene, and methylation signatures were 0.595, 0.573, and 0.638, respectively, each of which was lower than that of the 21-gene signature (C-index: 0.679). Furthermore, the AUC for the 21-gene signature was much higher than that for the other three signatures (Figure 2D). These results demonstrated the powerful ability of the 21-gene signature to predict prognosis and outperform the other signatures.



**Figure 1. Risk stratification of CM patients using the 21-gene signature.** (A) Cross-validation error plot for tuning parameter selection in the proportional hazards model. The R package glmnet returns a sequence of lambda values and cross-validation chooses the optimal value. The plot includes the cross-validation curve (red dotted line) and upper and lower standard deviation (error bars) along the lambda sequence. We used the lambda value with a minimum standard error of the mean (red vertical line) across 100 runs to choose the 21-gene signature. (B) Risk score distribution with patient survival status. The x axis is sorted by risk score values. Red indicates dead patients and blue indicates those still alive. (C) The Kaplan–Meier survival analysis of the signature. Patients were divided into high- and low-risk groups using the cut-off value estimated by the R package survminer. (D) ROC analysis of the signature in predicting the OS of patients.

## External validation of the risk prediction ability of the 21-gene signature

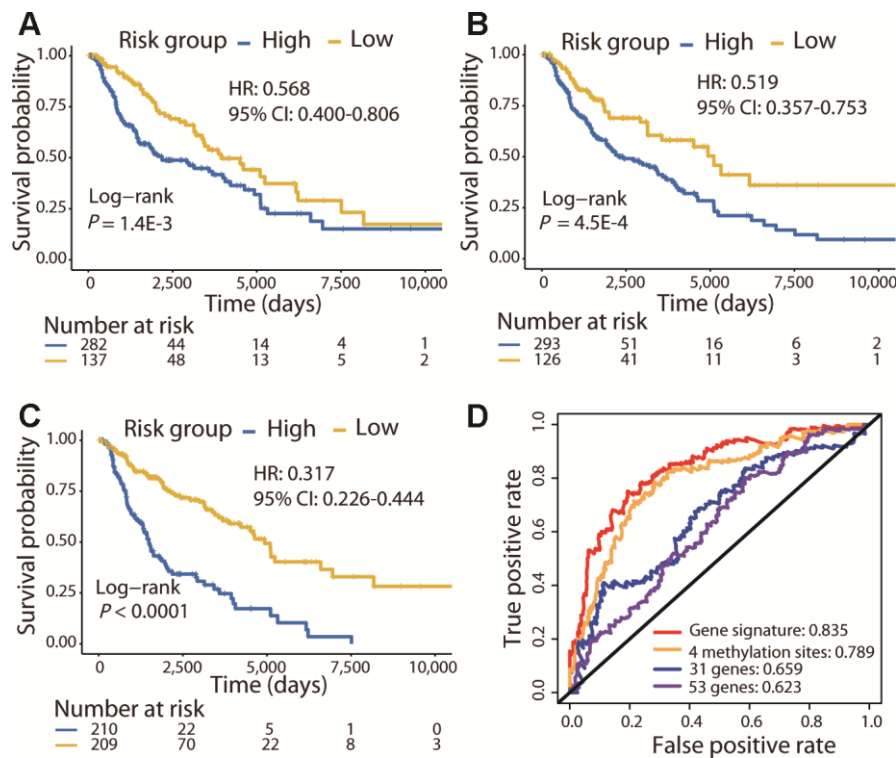
To validate the 21-gene signature in other populations, we calculated the risk score for melanoma patients in GSE54467 ( $n = 79$ ) and GSE65904 ( $n = 210$ ) using the same formula and performed risk stratification in the same way as with the TCGA cohort. Consistent with the results of the TCGA cohort, patients in the high-risk group exhibited a significantly poorer OS than those in the low-risk group in both validation populations (Log-rank test,  $P < 0.01$ ; Figure 3A). In addition, we used the 31-gene and 53-gene signatures to separate patients into high- and low-risk groups in these two validation cohorts. Both of these signatures successfully categorized patients, and the survival curves showed that the low-risk group had a longer OS than the high-risk group (Log-rank test,  $P < 0.01$ ; Figure 3B, 3C).

Subsequently, we compared the prognostic value of the 21-gene signature with the two signatures in the two validation cohorts. The C-index of the 21-gene signature had a higher value (GSE54467: 0.644; GSE65904: 0.616) compared with that of the 31-gene (GSE54467: 0.626; GSE65904: 0.573) and 53-gene (GSE54467: 0.599; GSE65904: 0.565) signatures in the

validation cohorts. Moreover, the 21-gene signature also achieved a higher AUCs than the two signatures in GSE54467 and GSE65904 (Figure 3D). Collectively, the prognostic capacity of the 21-gene signature was not only good in the TCGA cohort but also in the validation dataset, exhibiting superior ability to the two other known signatures.

## Comparison of risk prediction with other clinical factors

We further performed univariate and multivariate Cox regression analyses to evaluate the prognostic independence of the 21-gene signature. Clinical features including gender, age, Breslow thickness, ulceration, pathological stage, tumor site, and metastasis were compared with the 21-gene signature. The results showed that the 21-gene signature (risk score) and pathological stage were both significantly correlated with patient OS, independent of other factors (Table 1). However, the HR of the risk score indicated a 74.6% reduction in the risk of death in the low-risk group, but only a 46.8% reduction in the early stage (Stages 0, I, and II) group. The C-index of the pathological stage was 0.594, which was also lower than that of the 21-gene signature (0.679). In addition, we also compared



**Figure 2. Comparison of the 21-gene signature with known biomarkers in predicting the OS in the TCGA cohort.** Kaplan–Meier analysis was performed for patients classified by the 31-gene (A), 53-gene (B), and methylation (C) signatures. (D) ROC curves of the three known signatures and the 21-gene signature are shown. The AUC values of each signature demonstrate their ability to predict the patient OS.

their performance in the two validation datasets (Supplementary Table 2). In GSE54467, only the risk score was associated with OS (Log-rank test,  $P < 0.01$ ). Both the risk score and tumor stage were significantly correlated with prognosis in GSE65904; however, only the risk score was significantly associated with OS in the multivariate analysis (Log-rank test,  $P < 0.01$ ). Again, the 21-gene signature had a higher C-index (0.616) than that of the tumor stage (0.61) in GSE65904. These results demonstrated better risk prediction of the 21-gene signature than currently used clinicopathological prognostic factors.

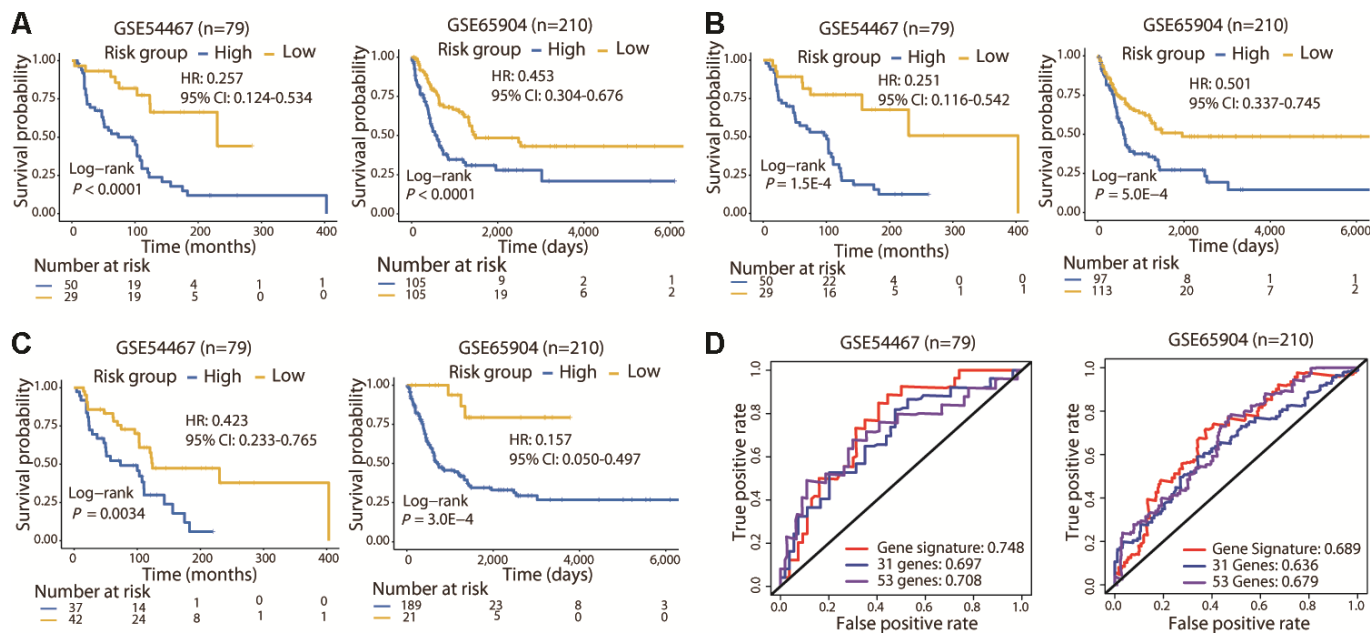
### Strong correlation of *BRAF-NRAS-NFI* expression in the high-risk group and weak correlation in the low-risk group

The above results demonstrated the robust performance of the 21-gene signature in prognosis prediction. Known driver genes, including *BRAF*, *NRAS*, and *NFI*, have been previously identified for CM; therefore, we compared mutation differences between the two risk groups using the exome-sequencing data of the same patients. Only *BRAF* showed a significantly high mutation load in the low-risk group (89/215 vs. 114/202, Fisher's exact test,  $P = 2.389E-3$ ). Gene expression comparison revealed that the expression level of *BRAF* was also significantly higher in low-risk patients (Student's *t* test, adjusted  $P$  value =  $4.15E-13$ ; Figure

4A). The other two driver genes, *NRAS* and *NFI*, also had higher expression levels in the low-risk group. We calculated the Spearman's rank correlation coefficient ( $\rho$ ) between the expression levels of the driver genes in all patients and found good correlation (Figure 4B). However, interestingly, there was a strong correlation in the high-risk group but a weak correlation in the low-risk group, despite these genes being upregulated in the latter (Figure 4B). The strong correlation of the expression levels of these driver genes in the high-risk group may be associated with poor prognosis in these patients.

### The genes involved in nicotinamide adenine dinucleotide (NAD) metabolism exhibited a negative correlation with *BRAF* expression

To obtain other correlation patterns, we further calculated the Spearman's  $\rho$  values between the expression level of *BRAF* and those of all genes in the transcriptome of patients in each risk group. We defined a strong positive correlation as  $\rho > 0.5$ , a strong negative correlation as  $\rho < -0.5$ , and a weak correlation as  $-0.3 < \rho < 0.3$ . When examining all  $\rho$  values, there were two sets of genes with correlation patterns: 1) those with a strong positive correlation with *BRAF* in the high-risk group and a weak correlation in the low-risk group; and 2) those with a strong negative correlation with *BRAF* in the high-risk group and a weak correlation in the low-risk group (Figure 5A).



**Figure 3. Survival and ROC curves of the different signatures in the validation cohorts.** Survival analysis of patients classified by the 21-gene (A), 31-gene (B), and 53-gene (C) signatures in the two validation cohorts. (D) ROC curves and AUC values of the three signatures in the two validation cohorts. In GSE65904, the AUC value for 3-year OS was calculated, since few patients lived for 5 years.

**Table 1. Comparison of the risk prediction ability between the 21-gene signature and clinical factors.**

Variable	Group	Univariate		Multivariate	
		HR (95% CI)	P value	HR (95% CI)	P value
Risk score	High (n = 216)	1	<b>1.11E-16</b>	1	<b>1.36E-7</b>
	Low (n = 203)	0.254 (0.180-0.359)		0.253 (0.152-0.422)	
Gender	Male (n = 264)	1.129 (0.796-1.601)	0.496		
	Female (n = 155)	1			
Breslow thickness (mm)	≤ 1 (n = 53)	1		1	
	1.01-2 (n = 77)	0.925 (0.529-1.619)		0.552 (0.259-1.177)	0.124
	2.01-4 (n = 67)	1.558 (0.886-2.739)	<b>3.75E-6</b>	0.550 (0.253-1.195)	0.131
	≥ 4 (n = 122)	2.962 (1.735-5.056)		0.724 (0.326-1.609)	0.428
Ulceration	Yes (n = 146)	1.953 (1.290-2.955)	<b>0.001</b>	1.605 (0.952-2.704)	0.076
	No (n = 132)	1		1	
Age	≤ 60 (n = 233)	0.575 (0.410-0.807)	<b>0.001</b>	0.590 (0.371-0.939)	0.026
	> 60 (n = 186)	1		1	
Pathological stage	Stage 0, I, and II (n = 203)	1	<b>4.84E-4</b>	1	<b>9.36E-4</b>
	Stage III and IV (n = 180)	1.880 (1.312-2.694)		2.326 (1.411-3.836)	
	Extremities (n = 181)	1			
Tumor site	Head and neck (n = 31)	1.451 (0.784-2.685)	0.111		
	Trunk (n = 145)	0.847 (0.586-1.225)			
	Other (n = 13)	1.921 (0.878-4.200)			
	Primary tumor (n = 78)	4.657 (1.538-4.657)		1.953 (0.682-5.590)	0.212
Metastasis	Regional lymph node (n = 208)	1.022 (0.678-1.539)	<b>8.51E-6</b>	0.895 (0.511-1.567)	0.698
	Regional cutaneous or subcutaneous tissue (n = 70)	0.784 (0.457-1.343)		0.483 (0.228-1.021)	0.057
	Distant metastasis (n = 60)	1		1	

Function analysis revealed that the positively correlated genes were involved in mRNA processing, and the negatively correlated genes were enriched in metabolic pathways (Figure 5B). Specifically, we noticed that four genes, *NDUFA3*, *NDUFB7*, *NDUFS8*, and *NDUFB11*, were present in all the metabolic pathways. Interestingly, these four genes encode the subunits of NADH dehydrogenase that catalyzes the formation of NAD (Figure 5D). Each of these genes exhibited a strong negative correlation with *BRAF* (Figure 5C) and was highly expressed in the high-risk group (Supplementary Figure 1). These data suggested that the NAD metabolic pathway, perhaps in conjunction with *BRAF* expression, may be important in CM patient survival.

### Cytochrome P450-mediated metabolic pathways were highly enriched in the high-risk group

To determine the global functional differences, we next investigated the differentially expressed genes between the two risk groups and compared their enriched biological functions. Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis revealed

that immune pathways were upregulated in the low-risk group and the chemical carcinogenesis function was significantly enriched in the high-risk group (Figure 6A).

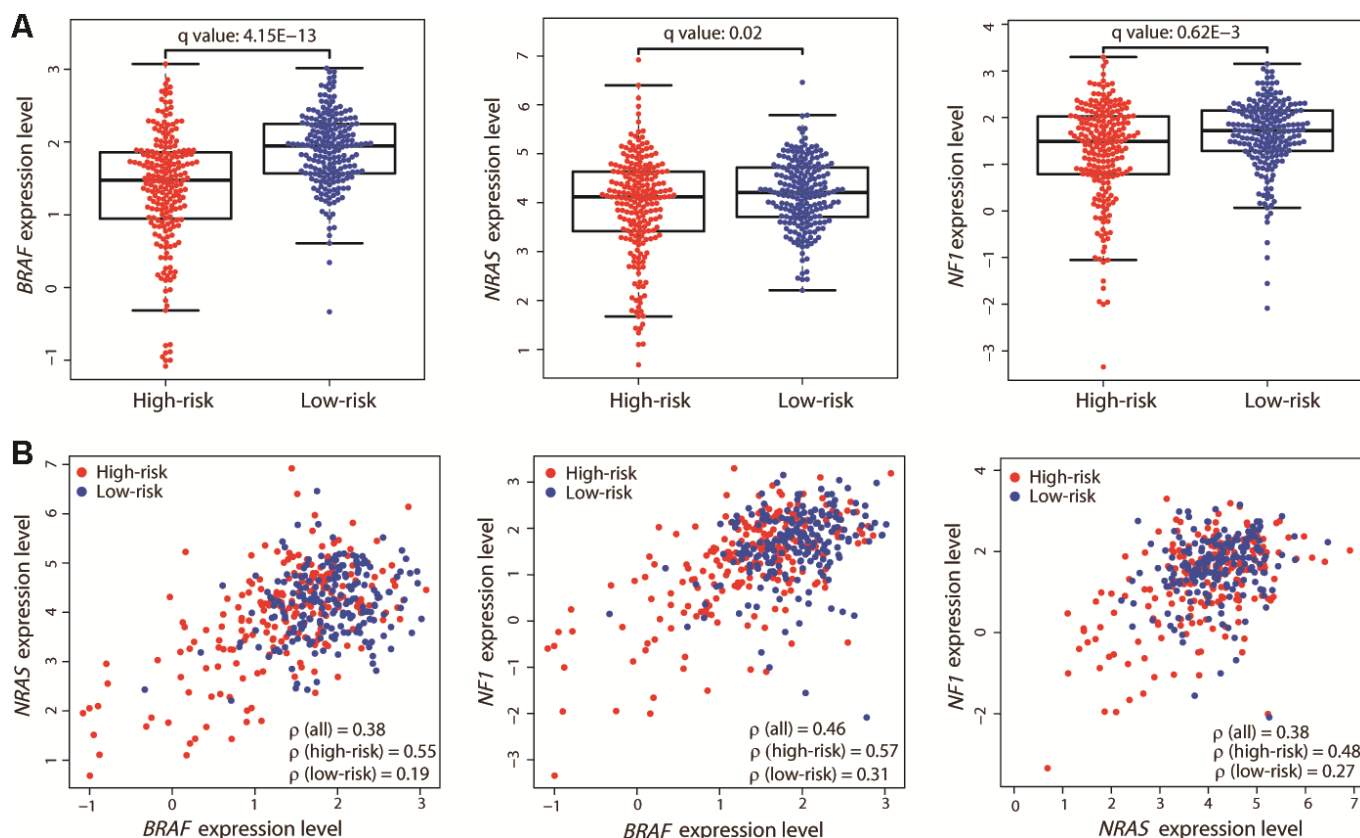
We carefully examined the components that were disordered in the chemical carcinogenesis function and found that the altered signal-transduction required cytochrome P450 enzymes. Specifically, the altered genes were members of three cytochrome P450-mediated metabolic pathways. The first pathway involved the metabolism of the environmental carcinogen benzo[a]pyrene (BaP) (Figure 6B), a cancer-causing agent, exposure to which can increase the risk of skin cancer. Three genes were upregulated in this pathway: *EPHX1*, *GSTO2*, and *GSTP1*. Activation of this pathway converts BaP to B[a]P-7,8-dihydrodiol-9,10-epoxide, a DNA-reactive intermediate. *EPHX1* is also a member of another enriched pathway, the metabolism of DMBA (Figure 6C); DMBA is a widely used chemical compound to induce skin cancer in animal models [18]. Similarly, this pathway activates the metabolism of DMBA via cytochrome P450, resulting in the formation of DNA adducts. These two

pathways both contribute to the development of skin cancer, and exposure to these carcinogens is associated with a greater risk of its occurrence. Thus, these results verified the rationale for our risk stratification. Moreover, we noticed that three upregulated genes in the high-risk group, *ALDH3A1*, *ALDH3B2*, and *ADH7*, participate in the cytochrome P450-mediated metabolism of trichloroethylene (Figure 6D); trichloroethylene is an organic chemical, exposure to which can cause cancer. Our result suggested that it may also be associated with an increased risk of CM.

Furthermore, we also performed single sample gene set enrichment analysis (ssGSEA) and the results confirmed the significant enrichment of cytochrome P450-mediated metabolic pathways in the high-risk group (Figure 6E). Collectively, these results demonstrated that the upregulation of cytochrome P450-mediated metabolic pathways had a deleterious effect on CM patient survival, highlighting the necessity to examine patient exposure to chemical carcinogens.

### The high-risk group displayed immune inhibition and the low-risk group was enriched in gamma delta T cells

Immune system-related genes were upregulated in the low-risk group (Figure 6A); consistently, the immune pathways in the ssGSEA analysis were significantly enriched in the low-risk group (Figure 7A–7G). We further investigated the cellular composition of immune infiltrates by quantitating the fraction of different immune cell types in the two risk groups. The cell fractions of a total of 22 immune cell types were estimated using the CIBERSORT algorithm [19]. The results showed a similar pattern of immune composition of the majority of immune cell types between the high- and low-risk groups (adjusted *P* value > 0.01; Figure 7H). However, in the high-risk group, we observed a depletion of gamma delta and activated memory CD4 T cells, and a significant enrichment of regulatory T (Treg) cells and T follicular helper cells (adjusted *P* value < 0.01). Tregs

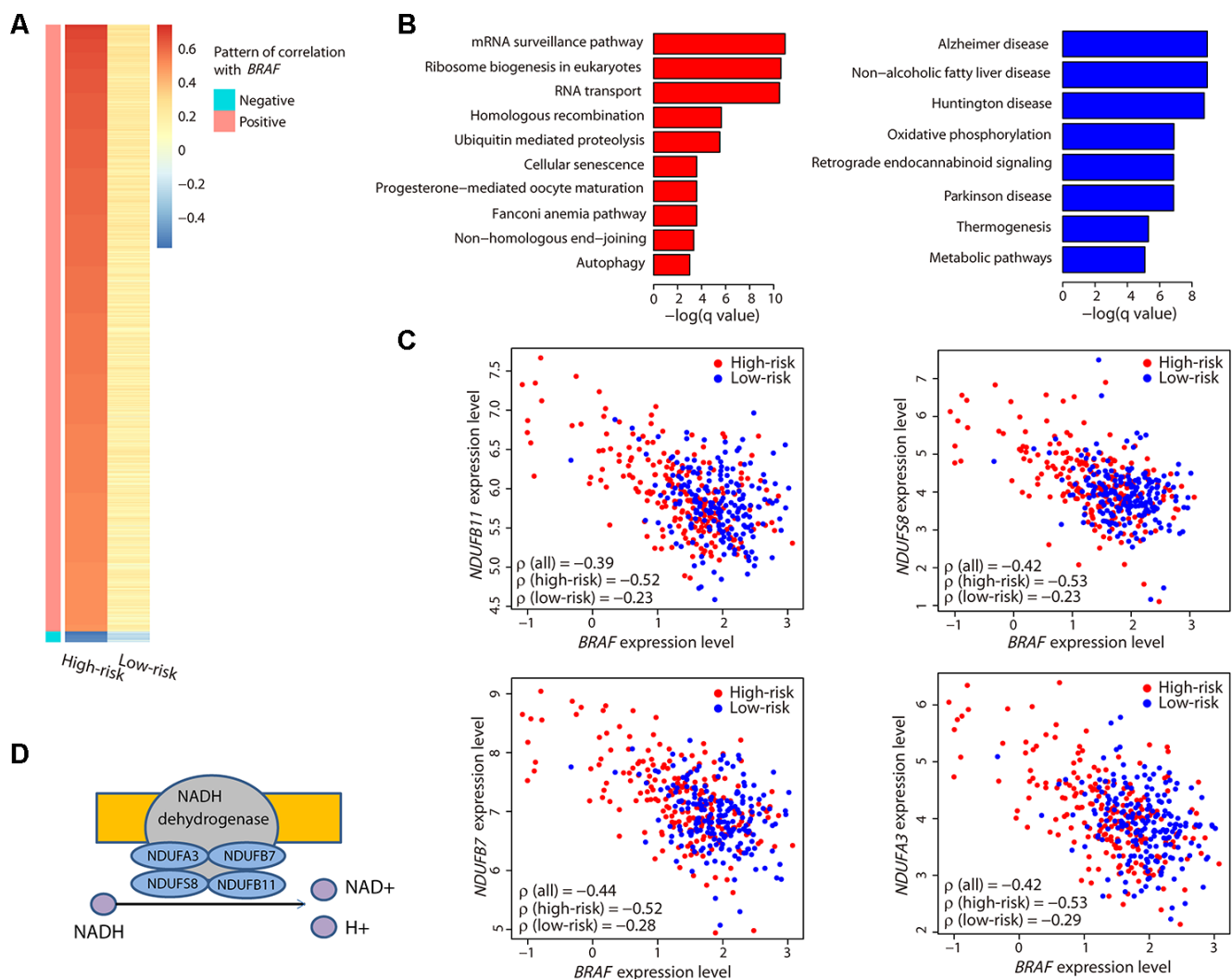


**Figure 4. Expression differences and correlation of the three driver genes between the two risk groups.** (A) Distribution of the expression levels of *BRAF*, *NRAS*, and *NF1* in the two risk groups. The expression levels are represented by log<sub>2</sub>-transformed RPKM values. The *P* value was calculated by a two-sided Student's *t* test and adjusted using the Benjamini–Hochberg method (*q* value). (B) Correlation between the expression levels of the three driver genes. Each dot represents one patient (high-risk, red dot; low-risk, blue dot). The correlation coefficient  $\rho$  was calculated by Spearman's rank analysis. Coefficient values between two genes among all patients, high-risk patients, and low-risk patients are shown.

have immunosuppressive functions in cancer, such as inhibiting recognition and clearance of tumor cells [20]. Gamma delta T cells, defined by the expression of heterodimeric T-cell receptors composed of  $\gamma$  and  $\delta$  chains, show tissue-specific localization and are enriched in skin and mucosal tissues. Evidence has shown that the infiltration of gamma delta T cells in a tumor was the best predictor of a favorable outcome [21]. The enrichment of gamma delta T cells in the low-risk group was consistent with a better prognosis, further supporting the rationale for our risk prediction. Therefore, our results indicated that immune inhibition existed in the high-risk group and immune activation existed in the low-risk group.

## DISCUSSION

The clinical features of CM have traditionally been used to assess patient risk; however, this method faces limitations since it does not consider the intrinsic heterogeneity of CM. Some patients diagnosed with thin melanomas (Breslow thickness < 1 mm) display variable 10-year survival outcomes ranging from 85% to 99%, and conversely, some patients with thicker melanomas may be cured by surgical management [22]. Molecular stratification of CM, such as gene signatures based on mRNA expression [12, 13], has shown great potential for prognosis prediction and disease management. The development of new biomarkers to



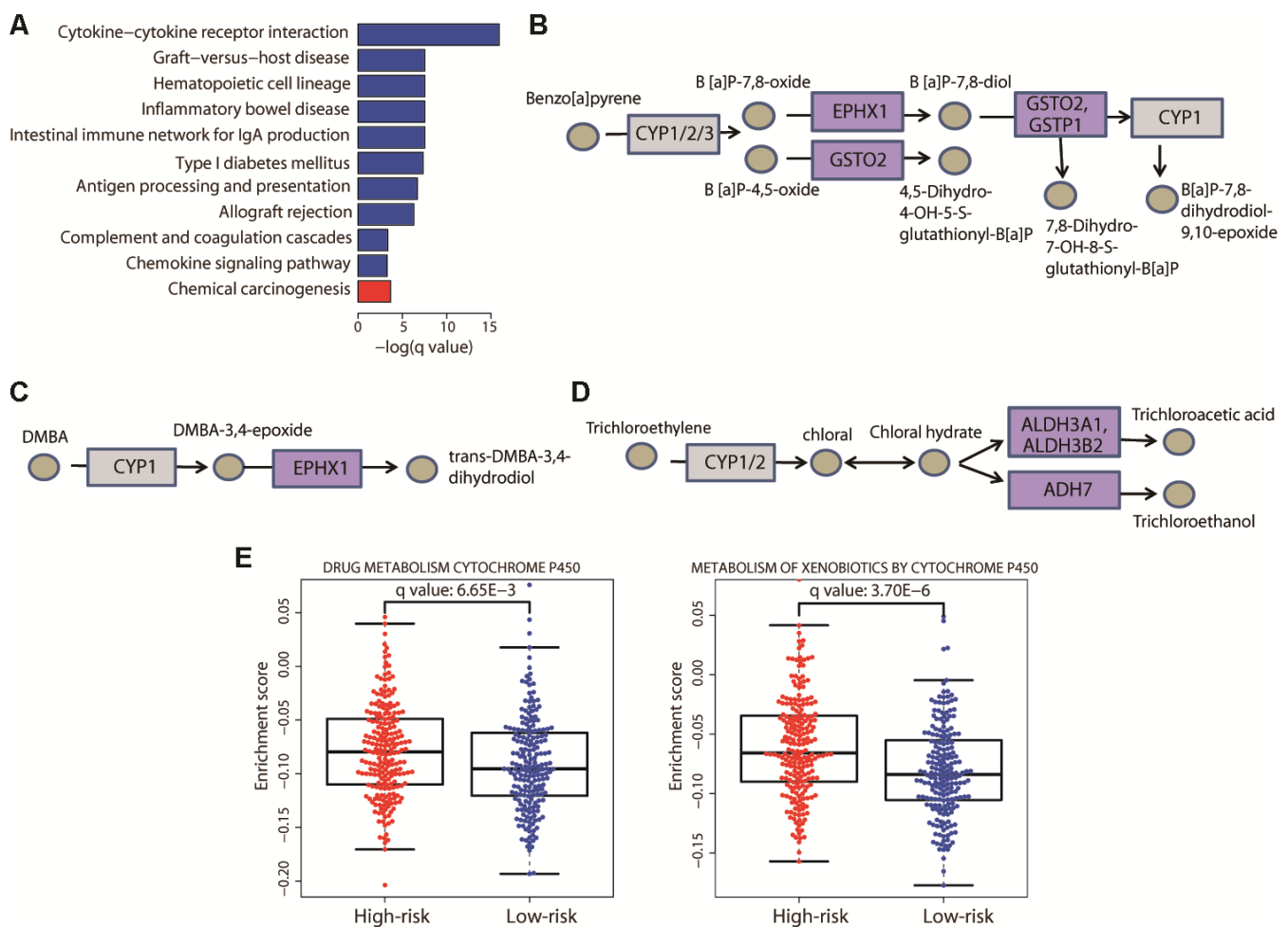
**Figure 5. Characterization of the *BRAF*-correlated genes.** (A) The heatmap illustrates all  $\rho$  values of genes that had a strong correlation with *BRAF*. Only two sets of genes exhibiting a strong positive or negative correlation in the high-risk group and a weak correlation in the low-risk group were obtained. (B) Bar plots showing the significantly (adjusted  $P$  value < 0.05) enriched KEGG pathways related to the positively (red) and negatively (blue) correlated genes. (C) Correlation between the expression levels of the four genes encoding the subunits of NADH dehydrogenase and *BRAF*. Coefficient values between two genes among all patients, high-risk patients, and low-risk patients are shown. (D) Illustration showing the four genes encoding the subunits of NADH dehydrogenase and their function.



precisely assess tumor prognosis would help us to better understand the underlying biology of high-risk patients, leading to improved treatment options.

In the present study, using the transcriptome profiles of 419 CM patients, we developed a new 21-gene prognostic signature with a powerful ability to stratify patients into high- and low-risk groups. The signature showed consistent prognosis prediction in two external validation cohorts and exhibited superior ability to other known signatures and clinicopathological factors. Among the 21 genes, several have been reported as biomarkers for other cancer types or diseases. *ATP11A* and *BOK* are new biomarkers for colorectal cancer [23, 24]. Interestingly, the BCL-2 family protein BOK has

been proposed to act in a pro-apoptotic pathway [25]. Links between *BOK* expression and patient outcome have not been reported in melanoma to date. In our data, increased expression of *BOK* was associated with poor OS in CM patients (HR: 2.337, 95% CI: 1.476–3.700; Log-rank test,  $P < 0.01$ ; Supplementary Table 1). When examining the differentially expressed genes between the two risk groups, we found that *BOK* was significantly upregulated in the high-risk group (adjusted  $P$  value  $< 0.01$ ). Owing to its role in apoptosis in tumor cells, these findings contradict our expectations. It has been reported that the expression level of *BCL-2* is increased during the development and progression of melanoma [25]. Notably, the function of *BOK* in mammalian cells has not been well

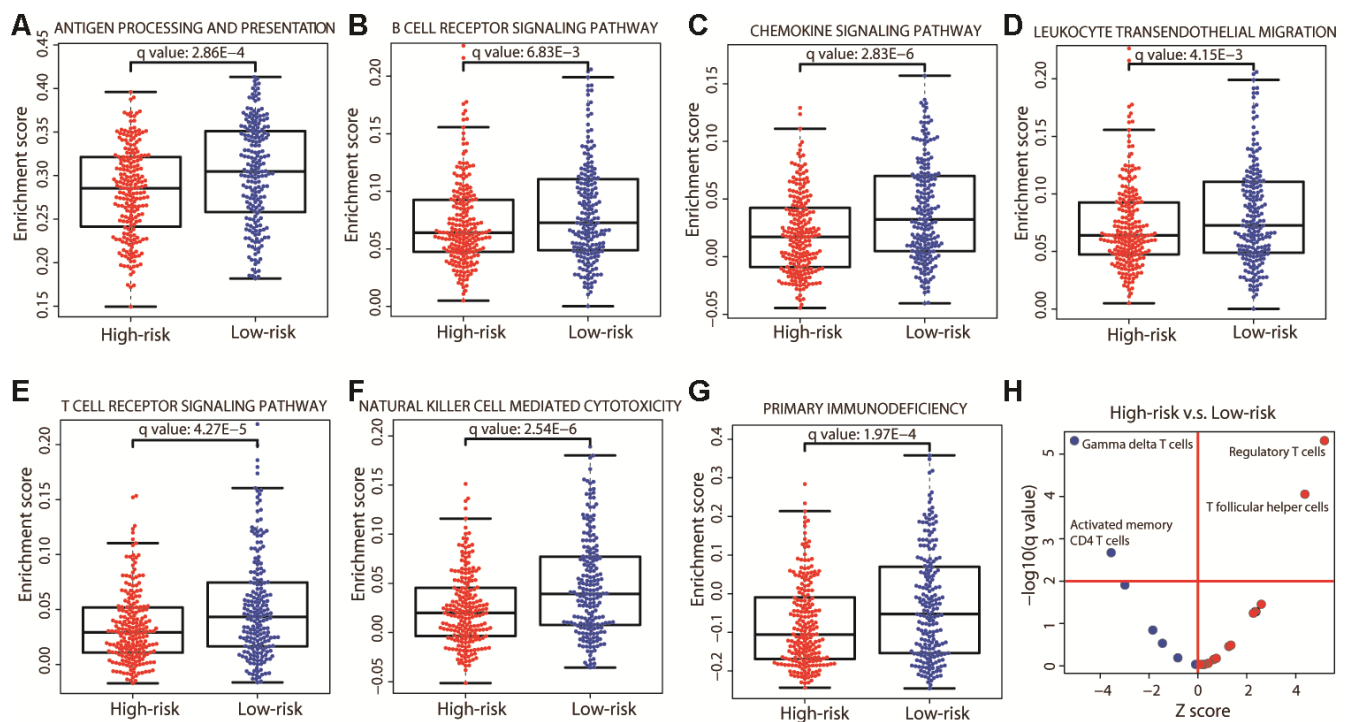


**Figure 6. Enrichment of the cytochrome P450-mediated metabolic pathways in the high-risk group.** (A) Bar plot showing the significantly (adjusted  $P$  value  $< 0.05$ ) enriched KEGG pathways in the high-risk (red) and low-risk (blue) patients. The chemical carcinogenesis function was significantly enriched in high-risk patients. The specific chemical metabolic pathways are illustrated in (B–D); in these three pathways, each circle represents one chemical compound and each rectangle represents an enzyme. The cytochrome P450 enzymes are marked in gray and the enzymes with altered gene expression are marked in purple. (E) Distribution of the enrichment scores of the cytochrome P450-mediated pathways that showed significant differences (adjusted  $P$  value  $< 0.01$ ) between the two risk groups. Each dot represents one patient. The enrichment score of each patient was calculated by ssGSEA. The  $P$  value was calculated by a two-sided Student’s  $t$  test and adjusted using the Benjamini–Hochberg method.

characterized and various other functions, such as a role in metabolism, have also been reported [26]. Its function and regulation in melanoma warrant further investigation. The upregulation of *COL22A1* has been proposed as a useful prognostic predictor in patients with squamous cell carcinoma of the head and neck [27]. The expression of *CLEC18* has been demonstrated as a potential biomarker in patients with chronic hepatitis B infection [28]. Different expression levels of *KLK13* have been shown to be associated with patient prognosis in esophageal squamous cell carcinoma [29] and bladder cancer [30]. The transcription regulator *HN1L*, identified as a new oncogene, is considered a potential biomarker for non-small cell lung cancer [31, 32]. Our study demonstrates that these genes also serve as a prognostic signature in CM, adding new prognostic value.

Intriguingly, the expression level of the most common oncogene in melanoma, *BRAF*, was significantly lower in the high-risk group as compared with that in the low-risk group. *BRAF* inhibitors have proven to be highly effective in targeting the oncogenic *BRAF* protein in melanoma patients [33]. However, the response to *BRAF* inhibitors varies among patients, as does the

expression of *BRAF*, and there is no correlation between its expression and response to *BRAF* inhibitors or survival [34]. Wilmott et al. suggested that the expression of *BRAF* did not predict the response or survival of patients, and hypothesized that patients with a low expression of *BRAF* would have a reduced survival rate as compared with those with a high expression [34]. This hypothesis is consistent with our observation that the high-risk patients actually harbored a low expression level of *BRAF*. The lack of correlation may be explained by the fact that melanoma survival is determined by complex molecular mechanisms. Our data showed that the expression levels of three driver genes, *BRAF*, *NRAS*, and *NF1*, were strongly positively correlated in the high-risk patients and weakly correlated in the low-risk patients. The interesting correlation pattern indicated that it may be the correlation pattern of each driver gene, not the expression level, that contributes to the high-risk tendency. On the other hand, our study reveals for the first time that the expression of genes involved in the NAD metabolic pathways are negatively correlated with the expression of *BRAF* in high-risk CM patients. Moreover, the functions of the genes that were upregulated in the high-risk group were enriched in



**Figure 7. Immune differences between the two risk groups.** (A–G) Distribution of the enrichment scores of the immune pathways that were significantly different (adjusted  $P$  value  $< 0.01$ ) between the two risk groups. Each dot represents one patient. The enrichment score of each patient was calculated by ssGSEA. The  $P$  value was calculated by a two-sided Student's  $t$  test and adjusted using the Benjamini–Hochberg method. (H) Immune cell composition differences between the high- and low-risk groups. The  $P$  values were calculated by the Wilcoxon rank-sum test and adjusted using the Benjamini–Hochberg method. The red horizontal line represents an adjusted  $P$  value of 0.01 and the red vertical line indicates a Z score of 0. The analysis was performed for all 22 immune cell types but only those that were significant are labeled on the plot.

cytochrome P450-mediated metabolic pathways. Taken together, these findings support a connection between metabolic dysregulation and high-risk CM patients.

Metabolic adaptation of cancer cells is required to support proliferation, growth and survival [35]. The present study shows that the expression levels of four genes encoding subunits of NADH dehydrogenase were negatively correlated with the expression of *BRAF* and upregulated in high-risk patients. NADH dehydrogenase catalyzes the oxidation of NADH to NAD, the latter of which is a key cofactor for energy transduction in metabolic processes. Increased levels of NAD result in metabolic alterations in cancer cells [36, 37]. It has been hypothesized that drugs interfering with the NAD biosynthetic enzyme would stop tumor growth [38]. Our results suggest that NADH dehydrogenase represents a new therapeutic target in CM patients. Interestingly, the inhibition of *BRAF* in melanoma cell lines has been reported to achieve high levels of NAD, which were activated by the overexpression of nicotinamide phosphoribosyltransferase (NAMPT), the most important NAD biosynthetic enzyme [39]. These observations are in accordance with our data demonstrating a negative correlation between the expression of *BRAF* and NADH dehydrogenase. *BRAF* may act as an important regulator in the metabolic alterations of CM cells.

Function analysis of the genes that were significantly upregulated in the high-risk group showed enrichment of cytochrome P450-mediated metabolic pathways, which also suggests metabolic alterations in high-risk patients. Many studies have reported the presence of cytochrome P450 enzymes in tumors and their role in the promotion of cancer progression [40, 41]; however, the involvement of cytochrome P450 in melanoma has not been well studied. Our observation that genes involved in this pathway were expressed at higher levels in the high-risk group offer a potential therapeutic option using these enzymes as drug targets. The enrichment of pathways metabolizing substrates of cytochrome P450 enzymes, such as BaP, DMBA, and trichloroethylene, in the present study, suggest that exposure to any toxic substance, such as smoking or drinking contaminated water, should be considered a lifestyle risk factor for screening CM patients in the future.

Melanoma has long been a core focus of ongoing immunotherapy research [42], and several therapeutic strategies have been approved by the FDA for clinical use [43–45]. However, resistance has been reported in a proportion of patients, with studies demonstrating that the existence of multiple immunosuppressive pathways in the tumor microenvironment is intrinsically responsible for failure

of immunotherapy [46–48]. Here, we reveal that immunosuppressive Tregs were accumulated in high-risk patients, suggesting that this group may acquire resistance to immunotherapy. Notably, metabolic dysregulation has been reported to occur with T cells, which consequently favors the development of Tregs [49]. Interestingly, the NAD pathway enzymes are receiving increasing attention due to their roles in several aspects of immune cell fate and function [50], once again highlighting that the inhibition of NAD synthesis could restore metabolic balance in the tumor microenvironment. Therefore, approaches inhibiting Treg functions or removing these cells directly from the tumor microenvironment could serve as new immunotherapeutic strategies for CM. On the other hand, our study shows that gamma delta T cells exhibited a high level of infiltration in low-risk patients, suggesting another promising treatment for CM utilizing these cells. In fact, there has been great interest in exploring the therapeutic potential of gamma delta T cells in different types of tumors, given the safety and favorable efficacy displayed in clinical trials [21, 51, 52]. Our study suggests that CM patients may benefit from immunotherapy that increases gamma delta T cells or inhibits Treg functions.

In summary, the present study developed a new 21-gene prognostic signature for CM risk assessment by considering genes from the entire transcriptome and deciphered the underlying mechanisms contributing to risk. The use of this signature may promote further biomarker discovery for the improvement of patient OS, and the uncovered mechanisms can aid the development of new therapeutic targets for CM.

## MATERIALS AND METHODS

### Data source

We collected data regarding 419 patients with skin cutaneous melanoma (SKCM), whose gene expression profile and clinical information were available, from Genomic Data Commons (GDC) (<https://portal.gdc.cancer.gov/>) (April 12, 2018). Both the normalized gene expression values (RPKM [reads per kilobase per million mapped reads]) and the raw reads counts data of the same patients were downloaded from the mRNA-seq platform (Illumina HiSeq 2000 RNA Sequencing). The log<sub>2</sub>-transformed RPKM values were used to represent the expression levels. Entrez IDs were used to represent genes, and a total of 19,620 protein-coding genes were retained for downstream analysis. For validation purposes, two sets of gene expression microarrays, GSE54467 (n = 79) and GSE65904 (n = 210), were obtained from the National Center for Biotechnology Information Gene Expression Omnibus

(NCBI GEO, <https://www.ncbi.nlm.nih.gov/geo/>). The corresponding clinical information was also acquired.

### Gene signature selection and patient risk classification

For each gene, the top quartile of patients with higher expression levels were selected as the high expression group and the bottom quartile of patients were selected as the low expression group. Genome-wide survival analysis was performed using the univariate Cox proportional hazards regression model in the survival package of the R platform (version 3.5.1) (R Core Team, Vienna, Austria). Genes with  $P$  values less than 0.01 were determined as significantly correlated with patient survival.

Subsequently, significant genes were entered into a lasso-penalized Cox proportional hazards model using the R package glmnet to select the optimal prognostic genes. A 10-fold cross-validation and a maximum number of 100,000 iterations were applied. We performed 100 repeat runs to avoid randomization of the results, obtaining a lambda value with the minimum standard error of the mean, which was used to extract the gene signature. The risk score for each patient was calculated as the sum of the expression level of each gene multiplied by its corresponding regression coefficient, as previously reported [16]:

risk score =  $\sum_{i=1}^k \beta_i E_i$ , where  $\beta_i$  was the regression coefficient from the Cox analysis, and  $E_i$  was the expression level of the  $i$ th gene.

Based on the risk score, an optimal cut-off value was estimated by the R package survminer, and the patients were classified into high- and low-risk groups according to the threshold. A survival curve was generated using the Kaplan–Meier method, and differences were evaluated using the Log-rank (Mantel–Cox) test. The HR and corresponding 95% CI were obtained using the Cox proportional hazards model.

The concordance C-index value was applied to evaluate the predictive accuracy. A larger C-index indicated a more accurate predictive ability of the model. In addition, ROC curve analysis was used to evaluate the predictive value of the risk score using the R package survivalROC. The AUC value for 5-year OS was calculated.

For the purpose of comparison with the methylation signature, the methylation data (Illumina Human Methylation 450 platform, beta values) of the same CM patients were downloaded from the GDC (April 12, 2018). Beta values were measured as the ratio of the methylated probe intensity over all methylation probe

intensities and were used to represent the relative methylation level. The predictive value of the risk score was investigated in the same manner in both the TCGA cohort and the validation cohorts.

### Comparison of the signature with clinical factors

Clinical parameters including gender, Breslow thickness, ulceration, age, pathological stage, tumor site, and metastasis were included for comparison and independent investigation. Patients were divided into different groups according to clinical characteristics: gender (male, female), Breslow thickness ( $\leq 1$  mm, 1.01–2 mm, 2.01–4 mm, and  $\geq 4$  mm), ulceration (yes, no), age ( $\leq 60$ ,  $> 60$ ), pathological stage (early stage [Stage 0, I, and II], later stage [Stage III and IV]), tumor site (extremities, head and neck, trunk, and other), and metastasis (primary tumor, regional lymph node, regional cutaneous or subcutaneous tissue, and distant metastasis). Univariate and multivariate analyses were performed using the Cox regression model. The validation datasets followed the same grouping and analysis procedure.

### Analysis of driver genes in both risk groups

We downloaded the exome-sequencing data of the same SKCM patients from GDC (April 12, 2018) and obtained the nonsynonymous mutations. For comparison of the gene mutation burden between the two risk groups, we calculated the number of patients in each group harboring the mutated gene, and Fisher's exact test was applied to evaluate the difference. A  $P$  value less than 0.01 was considered significant.

The differences in the gene expression levels between the two groups were evaluated by two-sided Student's  $t$  tests. The  $P$  value was adjusted using the Benjamini–Hochberg method. Spearman's rank analysis was performed to assess the correlation of gene expression with the whole cohort and with the high- and low-risk groups.

### Correlation analysis of the expression of genes in the entire transcriptome with that of *BRAF*

Based on genes in the entire transcriptome, we performed Spearman's rank analysis to assess the association of their expression levels with that of *BRAF* in each risk group. We defined three categories of correlation: a strong positive correlation as  $\rho > 0.5$ , a strong negative correlation as  $\rho < -0.5$ , and a weak correlation as  $-0.3 < \rho < 0.3$ . We then extracted genes that fell into six correlation patterns: 1) a strong positive correlation in the high-risk group and a weak correlation in the low-risk group; 2) a strong positive correlation in

the high-risk group and a strong negative correlation in the low-risk group; 3) a weak correlation in the high-risk group and a strong positive correlation in the low-risk group; 4) a weak correlation in the high-risk group and a strong negative correlation in the low-risk group; 5) a strong negative correlation in the high-risk group and a strong positive correlation in the low-risk group; and 6) a strong negative correlation in the high-risk group and a weak correlation in the low-risk group. Finally, we obtained two sets of genes: 1) those with a strong positive correlation with *BRAF* in the high-risk group and a weak correlation in the low-risk group; and 2) those with a strong negative correlation with *BRAF* in the high-risk group and a weak correlation in the low-risk group. Function analysis of these two sets of genes was performed using WebGestalt [53]. The *P* value was adjusted using the Benjamini–Hochberg method. A pathway was considered significantly enriched at an adjusted *P* value less than 0.05.

### Analysis of gene expression and functional differences between the two risk groups

We applied the R package DESeq [54] using the mRNA-Seq raw reads count data to identify the differentially expressed genes. The *P* value was adjusted using the Benjamini–Hochberg method. We defined genes as differentially expressed at an absolute log<sub>2</sub> fold change larger than 1 and an adjusted *P* value less than 0.01. Gene function analysis of these differentially expressed genes was performed using WebGestalt, and the *P* value was adjusted using the Benjamini–Hochberg method. The pathway was considered significantly enriched at an adjusted *P* value less than 0.05.

To further assess the enrichment of the functional pathways in each patient, we performed ssGSEA using the expression level data. All pathways were derived from the KEGG database and downloaded from the Molecular Signatures Database (MSigDB) (version 6.2). The enrichment score for each pathway in our sample was calculated by the R package GSVA [55]. A two-sided Student's *t* test was used to compare the difference in the enrichment score between the two risk groups. The adjusted *P* value for each pathway was calculated using the Benjamini–Hochberg method. The pathway was considered significantly enriched at an adjusted *P* value less than 0.01.

### Analysis of differences in immune cell composition between the two risk groups

The cell fractions of 22 immune cell types in each patient were based on a deconvolution approach, CIBERSORT, and obtained directly from The Cancer

Immunome Atlas (<https://tcia.at/>) (August 26, 2019). The difference in the relative composition of immune cell populations between the two risk groups was calculated by the Wilcoxon rank-sum test. *P* values were adjusted using the Benjamini–Hochberg method. Immune cell types with an adjusted *P* value less than 0.01 were deemed significant result.

### AUTHOR CONTRIBUTIONS

X.L. developed the concept, designed the research, collected data, performed all analyses, and wrote the manuscript; Y.C. provided computing hardware and financial support. All authors read and approved the final manuscript.

### ACKNOWLEDGMENTS

The authors would like to thank the TCGA project for data sharing.

### CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

### FUNDING

This work was supported by the National Natural Science Foundation of China under Grant 31501065 (to X.L.), and the Shenzhen Science and Technology Innovation Project under Grant JSGG20170823 144843046 (to Y.C.).

### REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin.* 2016; 66:7–30. <https://doi.org/10.3322/caac.21332> PMID:[26742998](https://pubmed.ncbi.nlm.nih.gov/26742998/)
2. Sladden MJ, Balch C, Barzilai DA, Berg D, Freiman A, Handiside T, Hollis S, Lens MB, Thompson JF. Surgical excision margins for primary cutaneous melanoma. *Cochrane Database Syst Rev.* 2009; 4:CD004835. <https://doi.org/10.1002/14651858.CD004835.pub2> PMID:[19821334](https://pubmed.ncbi.nlm.nih.gov/19821334/)
3. Noone AM, Howlader N, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen Hs, Feuer EJ, Cronin KA (Eds). SEER Cancer Statistics Review, 1975–2015, National Cancer Institute. Bethesda, MD, [https://seer.cancer.gov/csr/1975\\_2015/](https://seer.cancer.gov/csr/1975_2015/)
4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin.* 2018; 68:7–30. <https://doi.org/10.3322/caac.21442> PMID:[29313949](https://pubmed.ncbi.nlm.nih.gov/29313949/)

5. Weiss SA, Hanniford D, Hernando E, Osman I. Revisiting determinants of prognosis in cutaneous melanoma. *Cancer*. 2015; 121:4108–23.  
<https://doi.org/10.1002/cncr.29634>  
PMID:[26308244](https://pubmed.ncbi.nlm.nih.gov/26308244/)
6. Gershenwald JE, Scolyer RA, Hess KR, Sondak VK, Long GV, Ross MI, Lazar AJ, Faries MB, Kirkwood JM, McArthur GA, Haydu LE, Eggermont AM, Flaherty KT, et al, and for members of the American Joint Committee on Cancer Melanoma Expert Panel and the International Melanoma Database and Discovery Platform. Melanoma staging: evidence-based changes in the American joint committee on cancer eighth edition cancer staging manual. *CA Cancer J Clin*. 2017; 67:472–92.  
<https://doi.org/10.3322/caac.21409>  
PMID:[29028110](https://pubmed.ncbi.nlm.nih.gov/29028110/)
7. Homsy J, Kashani-Sabet M, Messina JL, Daud A. Cutaneous melanoma: prognostic factors. *Cancer Control*. 2005; 12:223–29.  
<https://doi.org/10.1177/107327480501200403>  
PMID:[16258493](https://pubmed.ncbi.nlm.nih.gov/16258493/)
8. Kattan MW, Hess KR, Amin MB, Lu Y, Moons KG, Gershenwald JE, Gimotty PA, Guinney JH, Halabi S, Lazar AJ, Mahar AL, Patel T, Sargent DJ, et al, and members of the AJCC Precision Medicine Core. American joint committee on cancer acceptance criteria for inclusion of risk models for individualized prognosis in the practice of precision medicine. *CA Cancer J Clin*. 2016; 66:370–74.  
<https://doi.org/10.3322/caac.21339>  
PMID:[26784705](https://pubmed.ncbi.nlm.nih.gov/26784705/)
9. Hyams DM, Cook RW, Buzaid AC. Identification of risk in cutaneous melanoma patients: prognostic and predictive markers. *J Surg Oncol*. 2019; 119:175–86.  
<https://doi.org/10.1002/iso.25319>  
PMID:[30548543](https://pubmed.ncbi.nlm.nih.gov/30548543/)
10. Zhang R, Ye J, Huang H, Du X. Mining featured biomarkers associated with vascular invasion in HCC by bioinformatics analysis with TCGA RNA sequencing data. *Biomed Pharmacother*. 2019; 118:109274.  
<https://doi.org/10.1016/j.biopha.2019.109274>  
PMID:[31545220](https://pubmed.ncbi.nlm.nih.gov/31545220/)
11. Shen S, Wang G, Zhang R, Zhao Y, Yu H, Wei Y, Chen F. Development and validation of an immune gene-set based prognostic signature in ovarian cancer. *EBioMedicine*. 2019; 40:318–26.  
<https://doi.org/10.1016/j.ebiom.2018.12.054>  
PMID:[30594555](https://pubmed.ncbi.nlm.nih.gov/30594555/)
12. Sivendran S, Chang R, Pham L, Phelps RG, Harcharik ST, Hall LD, Bernardo SG, Moskalenko MM, Sivendran M, Fu Y, de Moll EH, Pan M, Moon JY, et al. Dissection of immune gene networks in primary melanoma tumors critical for antitumor surveillance of patients with stage II-III resectable disease. *J Invest Dermatol*. 2014; 134:2202–11.  
<https://doi.org/10.1038/jid.2014.85>  
PMID:[24522433](https://pubmed.ncbi.nlm.nih.gov/24522433/)
13. Gerami P, Cook RW, Wilkinson J, Russell MC, Dhillon N, Amaria RN, Gonzalez R, Lyle S, Johnson CE, Oelschlager KM, Jackson GL, Greisinger AJ, Maetzold D, et al. Development of a prognostic genetic signature to predict the metastatic risk associated with cutaneous melanoma. *Clin Cancer Res*. 2015; 21:175–83.  
<https://doi.org/10.1158/1078-0432.CCR-13-3316>  
PMID:[25564571](https://pubmed.ncbi.nlm.nih.gov/25564571/)
14. Greenhaw BN, Zitelli JA, Brodland DG. Estimation of prognosis in invasive cutaneous melanoma: an independent study of the accuracy of a gene expression profile test. *Dermatol Surg*. 2018; 44:1494–500.  
<https://doi.org/10.1097/DSS.0000000000001588>  
PMID:[29994951](https://pubmed.ncbi.nlm.nih.gov/29994951/)
15. Hsueh EC, DeBloom JR, Lee J, Sussman JJ, Covington KR, Middlebrook B, Johnson C, Cook RW, Slingluff CL Jr, McMasters KM. Interim analysis of survival in a prospective, multi-center registry cohort of cutaneous melanoma tested with a prognostic 31-gene expression profile test. *J Hematol Oncol*. 2017; 10:152.  
<https://doi.org/10.1186/s13045-017-0520-1>  
PMID:[28851416](https://pubmed.ncbi.nlm.nih.gov/28851416/)
16. Chen HY, Yu SL, Chen CH, Chang GC, Chen CY, Yuan A, Cheng CL, Wang CH, Terng HJ, Kao SF, Chan WK, Li HN, Liu CC, et al. A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med*. 2007; 356:11–20.  
<https://doi.org/10.1056/NEJMoa060096>  
PMID:[17202451](https://pubmed.ncbi.nlm.nih.gov/17202451/)
17. Guo W, Zhu L, Zhu R, Chen Q, Wang Q, Chen JQ. A four-DNA methylation biomarker is a superior predictor of survival of patients with cutaneous melanoma. *Elife*. 2019; 8:e44310.  
<https://doi.org/10.7554/eLife.44310>  
PMID:[31169496](https://pubmed.ncbi.nlm.nih.gov/31169496/)
18. Xu C, Huang MT, Shen G, Yuan X, Lin W, Khor TO, Conney AH, Kong AN. Inhibition of 7,12-dimethylbenz(a)anthracene-induced skin tumorigenesis in C57BL/6 mice by sulforaphane is mediated by nuclear factor E2-related factor 2. *Cancer Res*. 2006; 66:8293–96.  
<https://doi.org/10.1158/0008-5472.CAN-06-0300>  
PMID:[16912211](https://pubmed.ncbi.nlm.nih.gov/16912211/)
19. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015; 12:453–57.

- <https://doi.org/10.1038/nmeth.3337>  
PMID:25822800
20. Hsieh CS, Lee HM, Lio CW. Selection of regulatory T cells in the thymus. *Nat Rev Immunol*. 2012; 12:157–67.  
<https://doi.org/10.1038/nri3155>  
PMID:22322317
21. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD, Diehn M, West RB, Plevritis SK, Alizadeh AA. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med*. 2015; 21:938–45.  
<https://doi.org/10.1038/nm.3909>  
PMID:26193342
22. Balch CM, Gershenwald JE, Soong SJ, Thompson JF, Atkins MB, Byrd DR, Buzaid AC, Cochran AJ, Coit DG, Ding S, Eggermont AM, Flaherty KT, Gimotty PA, et al. Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol*. 2009; 27:6199–206.  
<https://doi.org/10.1200/JCO.2009.23.4799>  
PMID:19917835
23. Miyoshi N, Ishii H, Mimori K, Tanaka F, Nagai K, Uemura M, Sekimoto M, Doki Y, Mori M. ATP11A is a novel predictive marker for metachronous metastasis of colorectal cancer. *Oncol Rep*. 2010; 23:505–10.  
<https://doi.org/10.3892/or.00000662>  
PMID:20043114
24. Carberry S, D’Orsi B, Monsefi N, Salvucci M, Bacon O, Fay J, Rehm M, McNamara D, Kay EW, Prehn JH. The BAX/BAK-like protein BOK is a prognostic marker in colorectal cancer. *Cell Death Dis*. 2018; 9:125.  
<https://doi.org/10.1038/s41419-017-0140-2>  
PMID:29374142
25. Bush JA, Li G. The role of bcl-2 family members in the progression of cutaneous melanoma. *Clin Exp Metastasis*. 2003; 20:531–39.  
<https://doi.org/10.1023/a:1025874502181>  
PMID:14598887
26. Srivastava R, Cao Z, Nedeva C, Naim S, Bachmann D, Rabachini T, Gangoda L, Shahi S, Glab J, Menassa J, Osellame L, Nelson T, Fernandez-Marrero Y, et al. BCL-2 family protein BOK is a positive regulator of uridine metabolism in mammals. *Proc Natl Acad Sci USA*. 2019; 116:15469–74.  
<https://doi.org/10.1073/pnas.1904523116>  
PMID:31311867
27. Misawa K, Kanazawa T, Imai A, Endo S, Mochizuki D, Fukushima H, Misawa Y, Mineta H. Prognostic value of type XXII and XXIV collagen mRNA expression in head and neck cancer patients. *Mol Clin Oncol*. 2014; 2:285–91.  
<https://doi.org/10.3892/mco.2013.233> PMID:24649348
28. Tsai TY, Peng CY, Yang HI, Huang YL, Tao MH, Yuan SS, Lai HC, Hsieh SL. The human c-type lectin 18 is a potential biomarker in patients with chronic hepatitis B virus infection. *J Biomed Sci*. 2018; 25:59.  
<https://doi.org/10.1186/s12929-018-0460-2>  
PMID:30055605
29. Nohara K, Yamada K, Yamada L, Hagiwara T, Igari T, Yokoi C, Soma D, Yamashita S, Dohi T, Kawamura YI. Expression of kallikrein-related peptidase 13 is associated with poor prognosis in esophageal squamous cell carcinoma. *Gen Thorac Cardiovasc Surg*. 2018; 66:351–57.  
<https://doi.org/10.1007/s11748-018-0910-5>  
PMID:29582368
30. Tokas T, Avgeris M, Alamanis C, Scorilas A, Stravodimos KG, Constantinides CA. Downregulated KLK13 expression in bladder cancer highlights tumor aggressiveness and unfavorable patients’ prognosis. *J Cancer Res Clin Oncol*. 2017; 143:521–32.  
<https://doi.org/10.1007/s00432-016-2301-6>  
PMID:27858162
31. Li L, Zeng TT, Zhang BZ, Li Y, Zhu YH, Guan XY. Overexpression of HN1L promotes cell Malignant proliferation in non-small cell lung cancer. *Cancer Biol Ther*. 2017; 18:904–15.  
<https://doi.org/10.1080/15384047.2017.1385678>  
PMID:29053395
32. Liu Y, Choi DS, Sheng J, Ensor JE, Liang DH, Rodriguez-Aguayo C, Polley A, Benz S, Elemento O, Verma A, Cong Y, Wong H, Qian W, et al. HN1L promotes triple-negative breast cancer stem cells through LEPR-STAT3 pathway. *Stem Cell Reports*. 2018; 10:212–27.  
<https://doi.org/10.1016/j.stemcr.2017.11.010>  
PMID:29249663
33. Sosman JA, Kim KB, Schuchter L, Gonzalez R, Pavlick AC, Weber JS, McArthur GA, Hutson TE, Moschos SJ, Flaherty KT, Hersey P, Kefford R, Lawrence D, et al. Survival in BRAF V600-mutant advanced melanoma treated with vemurafenib. *N Engl J Med*. 2012; 366:707–14.  
<https://doi.org/10.1056/NEJMoa1112302>  
PMID:22356324
34. Wilmott JS, Menzies AM, Haydu LE, Capper D, Preusser M, Zhang YE, Thompson JF, Kefford RF, von Deimling A, Scolyer RA, Long GV. BRAF(V600E) protein expression and outcome from BRAF inhibitor treatment in BRAF(V600E) metastatic melanoma. *Br J Cancer*. 2013; 108:924–31.  
<https://doi.org/10.1038/bjc.2013.29> PMID:23403819
35. Audrito V, Managò A, Gaudino F, Deaglio S. Targeting metabolic reprogramming in metastatic melanoma: the key role of nicotinamide phosphoribosyltransferase (NAMPT). *Semin Cell Dev Biol*. 2020; 98:192–201.  
<https://doi.org/10.1016/j.semcd.2019.05.001>

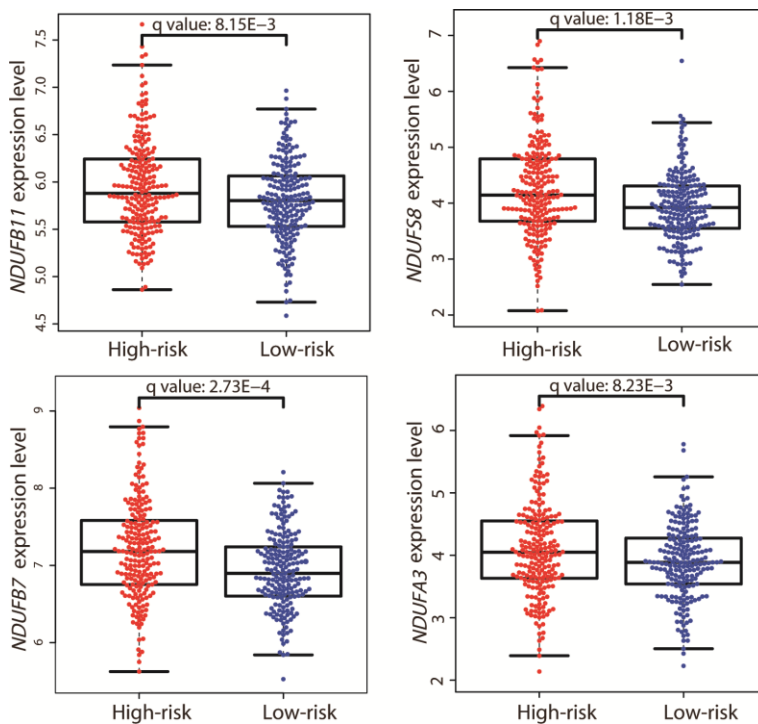
- PMID:[31059816](#)
36. Chiarugi A, Dölle C, Felici R, Ziegler M. The NAD metabolome—a key determinant of cancer cell biology. *Nat Rev Cancer*. 2012; 12:741–52.  
<https://doi.org/10.1038/nrc3340>  
PMID:[23018234](#)
  37. Garten A, Schuster S, Penke M, Gorski T, de Giorgis T, Kiess W. Physiological and pathophysiological roles of NAMPT and NAD metabolism. *Nat Rev Endocrinol*. 2015; 11:535–46.  
<https://doi.org/10.1038/nrendo.2015.117>  
PMID:[26215259](#)
  38. Sampath D, Zabka TS, Misner DL, O'Brien T, Dragovich PS. Inhibition of nicotinamide phosphoribosyltransferase (NAMPT) as a therapeutic strategy in cancer. *Pharmacol Ther*. 2015; 151:16–31.  
<https://doi.org/10.1016/j.pharmthera.2015.02.004>  
PMID:[25709099](#)
  39. Audrito V, Managò A, La Vecchia S, Zamporlini F, Vitale N, Baroni G, Cignetto S, Serra S, Bologna C, Stingi A, Arruga F, Vaisitti T, Massi D, et al. Nicotinamide Phosphoribosyltransferase (NAMPT) as a Therapeutic Target in BRAF-Mutated Metastatic Melanoma. *J Natl Cancer Inst*. 2018; 110: 10.1093/jnci/djx198.  
<https://doi.org/10.1093/jnci/djx198>  
PMID:[29309612](#)
  40. Mitsui Y, Chang I, Fukuhara S, Hiraki M, Arichi N, Yasumoto H, Hirata H, Yamamura S, Shahryari V, Deng G, Wong DK, Majid S, Shiina H, et al. CYP1B1 promotes tumorigenesis via altered expression of CDC20 and DAPK1 genes in renal cell carcinoma. *BMC Cancer*. 2015; 15:942.  
<https://doi.org/10.1186/s12885-015-1951-0>  
PMID:[26626260](#)
  41. Murray GI, Patimalla S, Stewart KN, Miller ID, Heys SD. Profiling the expression of cytochrome P450 in breast cancer. *Histopathology*. 2010; 57:202–11.  
<https://doi.org/10.1111/j.1365-2559.2010.03606.x>  
PMID:[20716162](#)
  42. Hu-Lieskovan S, Mok S, Homet Moreno B, Tsoi J, Robert L, Goedert L, Pinheiro EM, Koya RC, Graeber TG, Comin-Anduix B, Ribas A. Improved antitumor activity of immunotherapy with BRAF and MEK inhibitors in BRAF(V600E) melanoma. *Sci Transl Med*. 2015; 7:279ra41.  
<https://doi.org/10.1126/scitranslmed.aaa4691>  
PMID:[25787767](#)
  43. Eggermont AM, Chiarion-Sileni V, Grob JJ, Dummer R, Wolchok JD, Schmidt H, Hamid O, Robert C, Ascierto PA, Richards JM, Lebbé C, Ferraresi V, Smylie M, et al. Adjuvant ipilimumab versus placebo after complete resection of high-risk stage III melanoma (EORTC 18071): a randomised, double-blind, phase 3 trial. *Lancet Oncol*. 2015; 16:522–30.  
[https://doi.org/10.1016/S1470-2045\(15\)70122-1](https://doi.org/10.1016/S1470-2045(15)70122-1)  
PMID:[25840693](#)
  44. Queirolo P, Dozin B, Morabito A, Banelli B, Piccioli P, Fava C, Leo C, Carosio R, Laurent S, Fontana V, Ferrucci PF, Martinoli C, Cocorocchio E, et al. Corrigendum: association of CTLA-4 gene variants with response to therapy and long-term survival in metastatic melanoma patients treated with ipilimumab: an italian melanoma intergroup study. *Front Immunol*. 2018; 9:403.  
<https://doi.org/10.3389/fimmu.2018.00403>  
PMID:[29531523](#)
  45. Weber J, Mandala M, Del Vecchio M, Gogas HJ, Arance AM, Cowey CL, Dalle S, Schenker M, Chiarion-Sileni V, Marquez-Rodas I, Grob JJ, Butler MO, Middleton MR, et al, and CheckMate 238 Collaborators. Adjuvant nivolumab versus ipilimumab in resected stage III or IV melanoma. *N Engl J Med*. 2017; 377:1824–35.  
<https://doi.org/10.1056/NEJMoa1709030>  
PMID:[28891423](#)
  46. Larkin J, Hodi FS, Wolchok JD. Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *N Engl J Med*. 2015; 373:1270–1.  
<https://doi.org/10.1056/NEJMc1509660>  
PMID:[26398076](#)
  47. Wolchok JD, Chiarion-Sileni V, Gonzalez R, Rutkowski P, Grob JJ, Cowey CL, Lao CD, Wagstaff J, Schadendorf D, Ferrucci PF, Smylie M, Dummer R, Hill A, et al. Overall survival with combined nivolumab and ipilimumab in advanced melanoma. *N Engl J Med*. 2017; 377:1345–56.  
<https://doi.org/10.1056/NEJMoa1709684>  
PMID:[28889792](#)
  48. Ramapriyan R, Caetano MS, Barsoumian HB, Mafra AC, Zambalde EP, Menon H, Tsouko E, Welsh JW, Cortez MA. Altered cancer metabolism in mechanisms of immunotherapy resistance. *Pharmacol Ther*. 2019; 195:162–71.  
<https://doi.org/10.1016/j.pharmthera.2018.11.004>  
PMID:[30439456](#)
  49. Zhang L, Romero P. Metabolic control of CD8<sup>+</sup> T cell fate decisions and antitumor immunity. *Trends Mol Med*. 2018; 24:30–48.  
<https://doi.org/10.1016/j.molmed.2017.11.005>  
PMID:[29246759](#)
  50. Audrito V, Managò A, Gaudino F, Sorci L, Messana VG, Raffaelli N, Deaglio S. NAD-biosynthetic and consuming enzymes as central players of metabolic regulation of innate and adaptive immune responses in cancer. *Front Immunol*. 2019; 10:1720.



- <https://doi.org/10.3389/fimmu.2019.01720>  
PMID:[31402913](https://pubmed.ncbi.nlm.nih.gov/31402913/)
51. Kong Y, Cao W, Xi X, Ma C, Cui L, He W. The NKG2D ligand ULBP4 binds to TCRgamma9/delta2 and induces cytotoxicity to tumor cells through both TCRgammadelta and NKG2D. *Blood*. 2009; 114:310–17. <https://doi.org/10.1182/blood-2008-12-196287>  
PMID:[19436053](https://pubmed.ncbi.nlm.nih.gov/19436053/)
52. Chiplunkar S, Dhar S, Wesch D, Kabelitz D. Gammadelta T cells in cancer immunotherapy: current status and future prospects. *Immunotherapy*. 2009; 1:663–78. <https://doi.org/10.2217/imt.09.27>  
PMID:[20635991](https://pubmed.ncbi.nlm.nih.gov/20635991/)
53. Wang J, Vasaikar S, Shi Z, Greer M, Zhang B. WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res*. 2017; 45:W130–37. <https://doi.org/10.1093/nar/gkx356>  
PMID:[28472511](https://pubmed.ncbi.nlm.nih.gov/28472511/)
54. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. 2010; 11:R106. <https://doi.org/10.1186/gb-2010-11-10-r106>  
PMID:[20979621](https://pubmed.ncbi.nlm.nih.gov/20979621/)
55. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013; 14:7. <https://doi.org/10.1186/1471-2105-14-7>  
PMID:[23323831](https://pubmed.ncbi.nlm.nih.gov/23323831/)

## SUPPLEMENTARY MATERIALS

### Supplementary Figure



**Supplementary Figure 1. Expression differences in the four NADH dehydrogenase genes between the two risk groups.** Distribution of *NDUFA3*, *NDUF7*, *NDUF58*, and *NDUFB11* expression in the two risk groups. The expression levels are represented by log2-transformed RPKM values. The *P* value was calculated by a two-sided Student's *t* test and adjusted using the Benjamini-Hochberg method.

## Supplementary Tables

Supplementary Table 1. Information regarding the 21 genes following survival analysis.

Gene	Entrez ID	P value	coefficient	HR	95% CI
<i>SATB1</i>	6304	9.01E-5	-0.893	0.409	0.258-0.649
<i>HN1L</i>	90861	2.44E-4	0.928	2.53	1.514-4.229
<i>CCL8</i>	6355	6.58E-5	-0.939	0.391	0.243-0.631
<i>TTC39C</i>	125488	2.55E-4	-0.861	0.423	0.263-0.680
<i>HPDL</i>	84842	2.03E-4	0.875	2.398	1.491-3.858
<i>OLFML2A</i>	169611	8.75E-4	0.824	2.28	1.385-3.755
<i>LMNTD2</i>	256329	6.211E-3	-0.672	0.511	0.313-0.834
<i>ATP11A</i>	23250	3.173E-3	0.709	2.032	1.257-3.283
<i>SLC5A3</i>	6526	7.56E-6	-1.14	0.32	0.19-0.539
<i>HEYL</i>	26508	5.497E-3	0.677	1.968	1.209-3.201
<i>BOK</i>	666	2.01E-4	0.849	2.337	1.476-3.700
<i>RBCK1</i>	10616	1.056E-3	-0.792	0.453	0.278-0.736
<i>CCT6B</i>	10693	2.863E-3	-0.759	0.468	0.282-0.779
<i>ABTB1</i>	80325	5.107E-3	-0.643	0.526	0.333-0.831
<i>CLEC18A</i>	348174	5.438E-3	-0.612	0.542	0.35-0.841
<i>MRPS6</i>	64968	7.88E-4	-0.822	0.44	0.269-0.719
<i>NXT2</i>	55916	7.08E-5	-0.926	0.396	0.247-0.635
<i>SPEF2</i>	79925	2.208E-3	-0.708	0.492	0.310-0.782
<i>KLK13</i>	26085	6.35E-4	0.814	2.258	1.398-3.647
<i>SPAG8</i>	26206	9.86E-4	-0.765	0.465	0.292-0.742
<i>COL22A1</i>	169044	9.29E-4	-0.758	0.469	0.296-0.742

Supplementary Table 2. Risk prediction comparison with the validation cohorts.

GSE54467		Univariate		Multivariate	
Variable	Group	HR (95% CI)	P value	HR (95% CI)	P value
Risk score	High (n = 50)	1	<b>9.00E-5</b>		
	Low (n = 29)	0.257 (0.124-0.534)			
Age	≤ 60 (n = 47)	0.549 (0.309-0.975)	0.040		
	> 60 (n = 32)	1			
Gender	Male (n = 50)	1.090 (0.600-1.981)	0.800		
	Female (n = 29)	1			
Pathological stage	Stage I and II (n = 58)	1	0.100		
	Stage III (n = 20)	1.774 (0.860-3.658)			
GSE65904		Univariate		Multivariate	
Variable	Group	HR (95% CI)	P value	HR (95% CI)	P value
Risk score	High (n = 105)	1	<b>7.00E-5</b>		
	Low (n = 105)	0.453 (0.304-0.676)			
Gender	Male (n = 124)	1.335 (0.885-2.016)	0.200		
	Female (n = 86)	1			
	General (n = 23)	1			
Tumor stage	In transit (n = 15)	0.736 (0.333-1.628)	<b>1.00E-4</b>		
	Local (n = 11)	0.316 (0.106-0.947)			
	Primary (n = 15)	0.052 (0.007-0.394)			
	Regional (n = 139)	0.354 (0.204-0.613)			
Tissue	Cutaneous (n = 21)	1	<b>0.010</b>		
	Lymph node (n = 130)	2.884 (0.905-9.185)			
	Subcutaneous (n = 33)	5.096 (1.513-17.160)			
	Visceral (n = 10)	6.984 (1.743-27.987)			