Hepatology International (2019) 13:715–725
https://doi.org/10.1007/s12072-019-09978-9

ORIGINAL ARTICLE

Clinical and morpho-molecular classifiers for prediction of hepatocellular carcinoma prognosis and recurrence after surgical resection

Xiuming Zhang[1] · Yanfeng Bai[1] · Lei Xu[2,3] · Buyi Zhang[4] · Shi Feng[1] · Liming Xu[1] · Han Zhang[1] · Linjie Xu[1] · Pengfei Yang[2,3] · Tianye Niu[2,3] · Shusen Zheng[5,6,7] · Jimin Liu[8]

Received: 13 May 2019 / Accepted: 6 August 2019 / Published online: 17 September 2019
© Asian Pacific Association for the Study of the Liver 2019

Abstract

Background Approximately 50% hepatocellular carcinoma (HCC) patients die within 5 year after surgical resection. The present staging systems do not fully allow to accurately predict the HCC prognosis and recurrence. This study aimed to identify clinicopathological characteristics and molecular markers to establish classifiers to predict the 5-year overall survival (OS) and the 3-year recurrence in HCC patients post-operatively.

Methods We enrolled 647 HCC patients from two institutions, underwent surgical resection and divided the patients into one training and two validation cohorts. Clinicopathologic characteristics and tumor protein expression of 29 biomarkers by immunohistochemical (IHC) analysis were used to develop and validate a prognostic and a recurrent classifier, using the maximum relevance minimum redundancy algorithm jointly with the multivariable regression method.

Results The prognostic classifier distinguished HCC patients into high- and low-probability survival groups with significant differences in 5-year OS rate in all three cohorts (training cohort: 57.36% vs. 22.97%; $p < 0.0001$; internal validation cohort: 61.90% vs. 28.85%; $p < 0.0001$; independent validation cohort: 64.28% vs. 22.45%; $p < 0.0001$). The recurrent classifier also demonstrated good discrimination in all three cohorts.

Conclusion This study presented a prognostic classifier and a recurrent classifier using clinicopathologic and IHC characteristics. The developed classifiers stratified HCC patients into high- and low-probability survival or recurrent groups, which can help clinicians judge whether adjuvant therapy is beneficial post-operatively.

Keywords Hepatocellular carcinoma · Prognosis · Recurrence · Predicting classifiers · Immunomarkers

Introduction

Hepatocellular carcinoma is a highly prevalent cancer worldwide [1]. Surgical resection is the main therapy for HCC patients; however, this treatment is applicable only for a subset of patients. Most patients subsequently develop tumor recurrence, which results in poor clinical outcomes [2, 3]. Because of high recurrent rate and poor prognosis, accurate recurrent and prognostic assessments for the patients with HCC are imperative.

Several prognostic staging systems based on clinicopathological features have been developed for HCC, such as Tumor-Node-Metastasis (TNM) Staging System [4], the Japan Integrated Staging (JIS) score [5, 6], Barcelona Clinic Liver Cancer (BCLC) staging system [7], the Tokyo Score [8], Cancer of the Liver Italian Program (CLIP) score [9], and Okuda score [10]. Although these staging systems classify patients into distinct groups with various clinical outcomes, neither provides substantial prognostic information nor predicts the recurrence for the HCC patients accurately. New methods for more precise prognostic and recurrent prediction are urgent needed.

Xiuming Zhang, Yanfeng Bai, and Lei Xu contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12072-019-09978-9) contains supplementary material, which is available to authorized users.

✉ Shusen Zheng
shusenzheng@zju.edu.cn

✉ Jimin Liu
nliu@haltonhealthcare.com

Extended author information available on the last page of the article

Variable expressions in gene, mRNA, and protein profiles are reported to be important molecular prognostic markers [11, 12]. Recently, the predictive power of gene expression profiles in HCC has been explored [13]. Over the past decades, immunomarkers reflecting protein profile have been reported to be valuable in tumor detection, differentiation, prognosis, and early recurrence. These studies demonstrated the feasibility of developing classifiers by combining multiple clinicopathologic features and immunomarkers to predict the prognosis and recurrence for HCC patients. Therefore, this study aimed to develop prognostic and recurrent classifiers to predict the 5-year OS and 3-year recurrence using both immunomarkers and clinicopathologic characteristics for patients with surgically resectable HCC.

Materials and methods

Patient population

This retrospective multicenter study was approved by the research ethics committees at the First Affiliated Hospital, Zhejiang University School of Medicine (Institution I) and the Second Affiliated Hospital, Zhejiang University School of Medicine (Institution II). Hospital medical records were reviewed to identify patients with primary and biopsy-confirmed HCC at both institutions. The inclusion and exclusion criteria of patient cohorts are provided in the Supplement. A total number of 542 patients diagnosed between October 2010 and December 2013 were enrolled in this study from the Institution I. 105 patients diagnosed between February 2011 and January 2013 were identified from the Institution II. The patient cohort from the Institution I was randomly separated into four partitions. 406/542 patients (three partitions) were in the training cohort, while 136/542 patients (one partition) were in the internal validation cohort. 105 patients from the Institution II were used as an independent validation cohort. All patients were followed up by telephone until death or deadline of April 30, 2018.

The clinical outcome of 5-year OS might partially reflect the long-term effect of tumor development after the treatment. Therefore, the 5-year OS status was used as the primary predictive outcome in this study. In addition, we developed a predictive model to evaluate the 3-year tumor recurrence, since more than half HCC patients developed recurrence in the initial 3 years after the surgery. The schematic for the prognostic classifier and recurrent classifier is presented in Supplementary Fig. S1.

Tissue microarray analysis and IHC markers

Tissue microarray analysis was conducted for tissue samples from all patients as described previously [14]. In this study, we selected 29 immunomarkers identified in the previous studies for IHC stain based on their role in HCC tumorigenesis, invasiveness, and metastasis [11, 13]. The name of these immunomarkers and their biologic nature, antibody clone, and antibody dilution are summarized in Supplementary Table S1.

Representative histopathologic findings of BAX, CD34, CEA, E-cadherin, FAS, HRas, MMP-9, PCNA, p53, S100A9, and TYSY on IHC stain are presented in Fig. 1.

The results of IHC were evaluated using a four-category scoring system, based on percentage of positive-stained cells and staining intensity in each sample. The specific description for the IHC scoring is presented in the Supplement.

Development of prognostic classifier

To build a prognostic classifier using the most survival-related factors, we applied the maximum relevance minimum redundancy (mRMR) algorithm to select the optimal factors [15]. Then, a prognostic classifier was developed using the multivariable regression method. The specific description for the model construction is presented in the Supplement.

Validation of prognostic classifier

The proposed prognostic classifier was validated in both internal and independent validation cohorts. A prognostic probability (called prognostic score) was calculated for each patient to reflect the 5-year survival probability according to the prognostic classifier. The classifier performance was evaluated in discrimination and calibration [16]. The receiver operating characteristic (ROC) curve and area under the ROC curve (AUC) measured the discrimination of the proposed classifier [17]. DeLong test was used to evaluate the difference in ROC curves [18]. To divide patients into the high- or low-probability survival groups, we defined the optimal threshold value in prognostic score based on the training cohort. Then, the threshold was used to stratify patients in the validation cohorts. The calibration curve accompanied with the Hosmer–Lemeshow test was used to measure the calibration of the proposed classifier [19]. A non-significant statistic result indicated a favorable model fit.

Survival curves were estimated for the predicted high- and low-probability survival groups according to the proposed prognostic classifier and threshold using the Kaplan–Meier method [20]. The differences between the two survival curves were measured using the log-rank test.

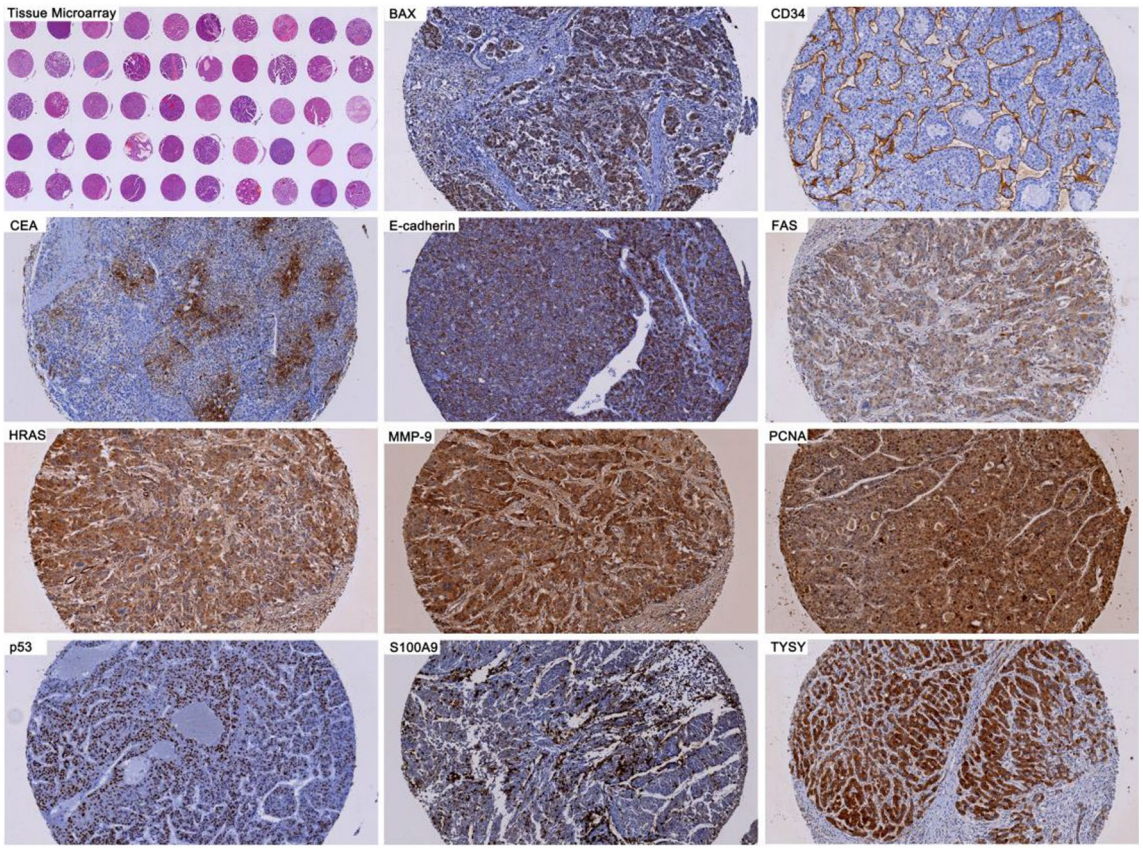

Fig. 1 Representative findings of BAX, CD34, CEA, E-cadherin, FAS, HRas, MMP-9, PCNA, p53, S100A9, and TYSY on IHC stain of 29 molecular markers used in this study

Development and validation of recurrent classifier

For the 3-year recurrence evaluation, we applied similar patient inclusion and exclusion criteria. In total, 677 patients from the Institute I, 507 in the training cohort and 170 in the validation cohort as well as 105 patients, an independent validation cohort, from the Institution II were enrolled in the study.

We used the mRMR algorithm to select the most recurrence-related factors. A recurrent classifier was developed using the multivariable regression algorithm. The model training and validation procedures were the same as the procedure for the prognostic classifier. A recurrent score was calculated for each patient to reflect the 3-year recurrent risk using the recurrent classifier. The performance was measured concerning its discrimination and calibration.

Statistical analysis

The univariate analysis for the prognostic evaluation was performed to evaluate the predictive ability of individual clinicopathologic characteristic and immunomarker using the Cox proportional hazard regression analysis and ROC curve analysis for patients in the training cohort. The univariate analysis for factors in the recurrent evaluation was performed using the ROC curve analysis in the training cohort.

R Software (version 3.4.1; http://www.R-project.org) and MedCalc Statistical Software (version 15.2.2; http://www.medcalc.org) were used for statistical analyses. A p value of less than 0.05 was defined as the significant difference in two-tailed analyses.

Results

Patient characteristics

The demographic and clinical characteristics for 5-year OS evaluation of patients in the training, internal validation, and independent validation cohorts are presented in Table 1. The univariate analysis of clinicopathologic characteristics and immunomarkers for 5-year OS evaluation of patients in the training cohort is listed in Table 2. Using the univariate analysis of Cox regression, seven clinicopathologic characteristics and four immunomarkers were considered statistically significant with $p < 0.05$ (post-operative HBV DNA

Table 1 Characteristics for 5-year OS evaluation of HCC patients in the training, internal validation and independent validation cohorts

Characteristic	Training cohort ($n=406$)		Internal validation cohort ($n=136$)		Independent validation cohort ($n=105$)	
	No.	%	No.	%	No.	%
Gender						
Female	59	14.53	18	13.24	10	9.52
Male	347	85.47	118	86.76	95	90.48
Age (years)						
Median	56		54		55	
Range	20–89		25–81		21–78	
Follow-up (months)						
Median	50		60		53	
Range	0.25–87		0.5–90		0.5–93	
Location						
Left	97	23.89	44	32.35	26	24.76
Right	297	73.15	84	61.77	78	74.29
Both	12	2.96	8	5.88	1	0.95
HBV DNA level						
$<1\times10^3$copies/ml	123	30.3	47	34.56	40	38.1
$\geq1\times10^3$copies/ml	283	69.7	89	65.44	65	61.9
Clinical stage						
I	326	80.29	109	80.15	80	76.19
II	28	6.90	12	8.82	11	10.48
III	52	12.81	15	11.03	13	12.38
IV	0	0	0	0	1	0.95
Tumor differentiation						
I	22	5.42	14	10.29	6	5.72
II	180	44.34	55	40.44	52	49.52
III	167	41.13	50	36.77	39	37.14
IV	37	9.11	17	12.50	8	7.62
Capsular invasion						
Yes	93	22.9	41	30.1	34	32.4
No	313	77.1	95	69.9	71	67.6
Tumor number						
Single	350	86.21	111	81.62	94	89.52
Multiple	56	13.79	25	18.38	11	10.48
Tumor diameter (cm)						
Median	4.0		4.0		4.0	
Range	1.0–17.5		1.0–14		1.0–20	
AFP level						
Normal	148	36.45	58	42.65	45	42.86
Abnormal	258	63.55	78	57.35	60	57.14
5-year Survival	182	44.83	67	49.26	47	44.76

AFP alpha-fetoprotein

level in serum, clinical stage, tumor differentiation, capsular invasion, tumor number, tumor diameter, AFP level; BAX, CD34, CK19, and MMP-9). For the ROC analysis, ten factors showed prognostic significance, including five clinicopathologic characteristics (HBV DNA level, clinical stage, capsular invasion, tumor number, and tumor diameter) and five immunomarkers (BAX, CD34, FAS, MMP-9, and TYSY).

Table 2 Univariate analysis for 5-year OS evaluation of clinicopathologic characteristics and immunomarkers for patients in the training cohort

Characteristic	Hazard ratio	95% CI	*p*	AUC	95% CI	*p* (area = 0.5)
Gender	1.0517	0.7282–1.5191	0.7891	0.502	0.453–0.552	0.8911
Age (years)	0.9955	0.9841–1.0069	0.4383	0.512	0.462–0.562	0.6745
Tumor location	1.0296	0.7753–1.3672	0.8412	0.507	0.457–0.556	0.7627
HBV DNA level	1.346	1.0125–1.6586	0.0086*	0.569	0.524–0.618	0.0011*
Clinical stage	1.7001	1.4395–2.0079	< 0.0001*	0.585	0.536–0.634	< 0.0001*
Tumor differentiation	1.2437	1.0430–1.4830	0.0157*	0.541	0.491–0.590	0.1224
Capsular invasion	1.7102	1.2361–2.3129	0.0068*	0.560	0.508–0.610	0.0012*
Tumor number	2.2587	1.6251–3.1394	< 0.0001*	0.565	0.515–0.614	0.0001*
Tumor diameter (cm)	1.1652	1.1171–1.2153	< 0.0001*	0.682	0.634–0.727	< 0.0001*
AFP level	1.4016	1.0613–1.8511	0.0179*	0.543	0.493–0.592	0.0738
BAX	1.1785	1.0302–1.3483	0.0172*	0.547	0.571–0.668	< 0.0001*
β-catenin	0.9407	0.8635–1.1457	0.9407	0.518	0.468–0.567	0.5152
Caspase-9	0.877	0.7654–1.0049	0.0601	0.541	0.491–0.591	0.1282
CD34	0.6681	0.5331–0.8374	0.0005*	0.579	0.529–0.627	0.0006*
CDKN1A	1.039	0.8980–1.2021	0.6093	0.508	0.458–0.557	0.7789
CDKN1B	1.0328	0.8470–1.2594	0.751	0.507	0.457–0.557	0.7483
CEA	0.8875	0.7222–1.0906	0.2588	0.539	0.489–0.588	0.0803
CK19	1.14	1.0105–1.2861	0.0341*	0.533	0.483–0.582	0.1247
c-Myc	0.9557	0.6676–1.3681	0.8055	0.505	0.455–0.554	0.6483
Cyclin-D1	1.0679	0.7402–1.5406	0.7266	0.501	0.451–0.551	0.9255
E-cadherin	0.9461	0.8294–1.0792	0.4118	0.531	0.481–0.580	0.2668
EMA	1.0978	0.9384–1.2844	0.2462	0.517	0.468–0.567	0.3694
FAS	0.8716	0.7507–1.0118	0.0725	0.556	0.506–0.605	0.0344*
HRas	0.8818	0.7703–1.0093	0.0693	0.549	0.500–0.599	0.0666
Ki67	1.0193	0.8585–1.2103	0.8277	0.504	0.455–0.554	0.8615
MLH1	0.6847	0.8092–1.3813	1.0572	0.502	0.452–0.552	0.9382
MMP-2	1.0971	0.9488–1.2685	0.2135	0.536	0.486–0.586	0.1713
MMP-9	0.8314	0.7212–0.9584	0.0113*	0.568	0.519–0.617	0.0110*
MSH2	1.0035	0.7724–1.3039	0.979	0.508	0.459–0.558	0.7343
PCNA	1.1058	0.9210–1.3276	0.2835	0.515	0.465–0.564	0.5745
PDCD4	1.0269	0.8959–1.1771	0.7047	0.503	0.453–0.552	0.921
PTEN	0.9232	0.7503–1.1359	0.4522	0.508	0.459–0.558	0.7355
p53	1.052	0.9320–1.1875	0.4144	0.508	0.458–0.557	0.7504
S100A4	0.975	0.8019–1.1854	0.8006	0.501	0.451–0.550	0.9755
S100A9	1.161	0.9293–1.4504	0.191	0.52	0.470–0.569	0.171
Survivin	1.0137	0.9149–1.1233	0.7954	0.506	0.456–0.556	0.8272
TGF-β	1.0492	0.8949–1.2301	0.5558	0.505	0.455–0.555	0.8558
TYSY	0.8975	0.7868–1.0238	0.1092	0.554	0.504–0.603	0.0471*
VEGF	0.9774	0.8483–1.1262	0.7535	0.509	0.460–0.559	0.728

AUC area under the ROC curve, *CI* confidence interval

*Significant difference, *p* value lower than 0.05

Development and validation of prognostic classifier

A prognostic classifier was developed by incorporating three clinicopathological characteristics of age, tumor number, tumor diameter, and seven immunomarkers, including BAX, CD34, CEA, E-cadherin, MMP-9, S100A9, and TYSY. The prognostic score calculation formula is provided in the Supplement.

The ROC curves for the proposed prognostic classifier are presented in Fig. 2. For the evaluation of patients in training cohort, the AUC for the proposed prognostic classifier (0.726, 95% confidence interval [CI]: 0.680–0.769) had significantly improvement than each individual factor from the DeLong test (Fig. 2a; Table 2). The proposed prognostic classifier showed favorable discrimination with an AUC of 0.731 (95% CI: 0.648–0.804) for the internal validation

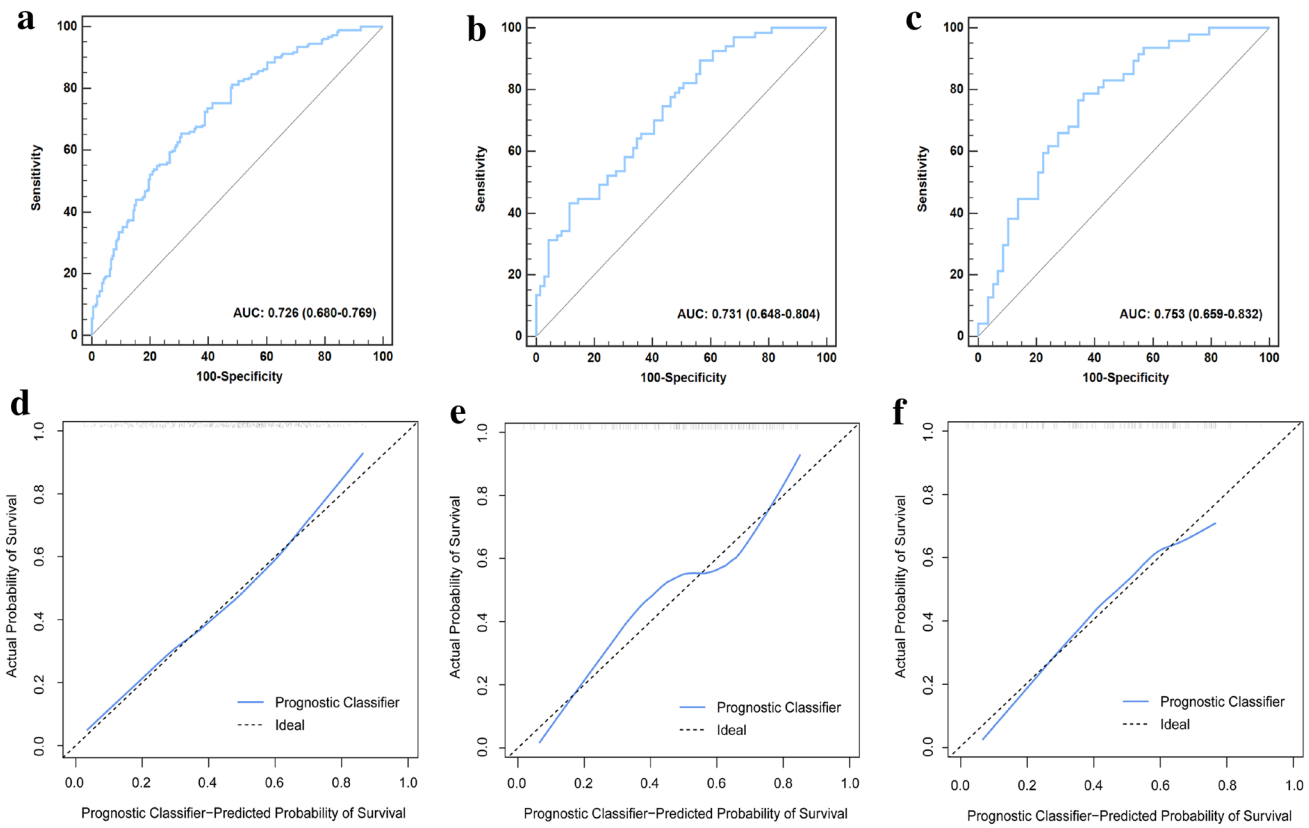

Fig. 2 ROC curves for the proposed prognostic classifier in the training (**a**), internal validation (**b**), and independent validation (**c**) cohorts. Calibration curves for the proposed prognostic classifier in the training (**d**), internal validation (**e**), and independent validation (**f**) cohorts. Calibration curves illustrate the calibration of the proposed classifier regarding the agreement between the predicted survival probability and the actual survival probability. The vertical axis indicates the actual survival probability, which was acquired using the bootstrapping method. The horizontal axis indicates the predicted survival probability. The black diagonal line indicates a perfect prediction by an ideal model, in which the predicted survival probability and the actual survival probability are exactly equal. The blue line indicates the performance of the proposed prognostic classifier, of which a less departure from the diagonal line represents a better prediction

cohort (Fig. 2b) and 0.753 (95% CI: 0.659–0.832) for the independent validation cohort (Fig. 2c).

The optimal threshold in the prognostic score was defined as −0.4471. Patients with prognostic scores ≥ −0.4471 were defined as the high-probability survival group, while patients with scores < −0.4471 were the low-probability survival group. In the training cohort, we identified 258 patients with high-probability of 5-year OS and 148 patients with low-probability. The mean survival month for the high-probability survival group was higher than that for the low-probability survival group (51.86 ± 19.83 vs. 34.60 ± 22.09). The 5-year OS rate differed statistically significantly between patients in the high-probability survival group and low-probability survival group (the high-probability survival group, 57.36%; the low-probability survival group, 22.97%; $p < 0.0001$) (Fig. 3a).

For the internal validation cohort, 84 patients were separated into the high-probability survival group, while 52 patients remained in the low-probability survival group. The mean survival month for the high-probability survival group was higher than the low-probability survival group (53.46 ± 20.43 vs. 38.30 ± 23.40). The 5-year OS rate was 61.90% and 28.85% for the high-probability survival group and low-probability survival group, respectively. A statistical difference was noticed between the two survival curves ($p < 0.0001$) (Fig. 3b).

For patients in the independent validation cohort, 56 patients were identified as high-probability survival, while 49 patients were low-probability survival. Patients in the high-probability survival group generally had longer survival time than patients in the low-probability survival group (63.00 ± 22.74 vs. 41.31 ± 21.97). The 5-year OS rate differed strikingly between the predicted high- and low-probability survival groups (64.28% vs. 22.45%; $p < 0.0001$) (Fig. 3c).

According to Cox regression analysis, the prognostic classifier was significantly highly related to the actual 5-year OS status in all cohorts (hazard ratio for predicted

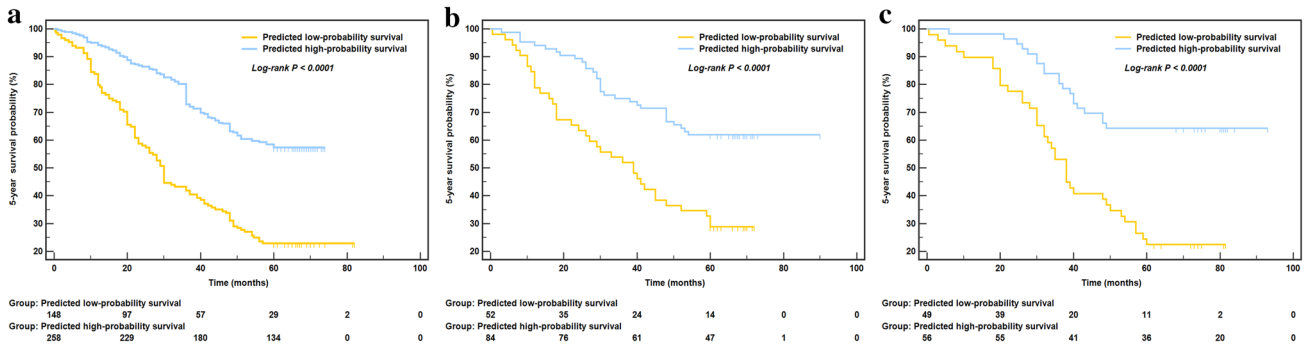

Fig. 3 Kaplan–Meier survival curves for the predicted high-probability survival and predicted low-probability survival in the training (**a**), internal validation (**b**), and independent validation (**c**) cohorts using the proposed prognostic classifier

high-probability survival vs. predicted low-probability survival; the training cohort: 2.7237, 95% CI: 2.0318–3.6512, $p < 0.0001$; the internal validation cohort: 2.5699, 95% CI: 1.5431–4.2798, $p < 0.0001$; the independent validation cohort: 3.1041, 95% CI: 1.8265–5.2752, $p < 0.0001$).

The calibration curves of the prognostic classifier demonstrated a good consistency between the predicted survival probability and the actual survival probability in the training cohort (Fig. 2d). The non-significant statistic from Hosmer–Lemeshow test ($p = 0.7760$) suggested no departure from the perfect fit. Good consistency was observed for the internal validation cohort (Fig. 2e), and independent validation cohort (Fig. 2f), as well as non-significant statistics from the Hosmer–Lemeshow test (the internal validation cohort, $p = 0.7030$; the independent validation cohort, $p = 0.9100$).

Development and validation of recurrent classifier

The demographic and clinical characteristics of the patients for the 3-year recurrence evaluation in the training, internal validation, and independent validation cohorts are presented in Supplementary Table S2. The univariate analysis of all clinicopathologic characteristics and immunomarkers for 3-year recurrence evaluation in the training cohort is listed in Supplementary Table S3. Using the ROC univariate analysis, 14 factors showed prognostic significance, including six clinicopathologic characteristics (clinical stage, differentiation level, capsular invasion, tumor number, tumor diameter, and AFP level) and eight immunomarkers (CD34, CDKN1A, E-cadherin, HRas, PCNA, p53, TGF-β, and VEGF).

The recurrent classifier was constructed by integrating five clinicopathologic characteristics of age, tumor location, tumor number, tumor diameter, AFP level, and six immunomarkers, including CD34, E-cadherin, HRas, PCNA, p53, and S100A9. The recurrent score calculation formula is provided in the Supplement. The recurrent classifier yielded a higher AUC of 0.734 (95% CI: 0.693–0.772) than individual factor for the training cohort (Fig. 4a; Supplementary Table S3). Favorable discrimination was also observed for the internal validation cohort (0.749, 95% CI: 0.677–0.812) (Fig. 4b) and the independent validation cohort (0.730, 95% CI: 0.635–0.812) (Fig. 4c).

The calibration curves suggested a good consistency between the predicted recurrent probability and the actual recurrent probability for all cohorts used in this study (Fig. 4d–f). The non-significant statistics of the Hosmer–Lemeshow test showed a desirable model fit (the training cohort: $p = 0.7028$; the internal validation cohort, $p = 0.1137$; the independent validation cohort, $p = 0.6041$).

Discussion

Accurate prediction of patient survival and tumor recurrence is crucial in modern oncology treatments [21]. Considerable efforts have been devoted to establishing prognostic models for HCC using clinical and pathological evidence [22]. Currently, the clinicopathological-based TNM staging system is the most acceptable and well-used method worldwide. The TNM staging provides an assessment of solid tumors based on clinical and pathological features, i.e., tumor size, vascular invasion and the extent of tumor invasion, etc., whereas IHC stain of tumor protein expression is not required. The heterogeneous nature of the tumor might produce different clinical outcomes among patients with the same TNM stage [13]. The proposed prognostic and recurrent classifiers were developed combining both clinicopathological characteristics and immunomarkers of HCC protein profile. The most important clinical question is whether the inclusion of various IHC parameters into the prognostic classifiers could improve the predictive ability of prognosis as compared to the classifiers which are constructed by the conventional clinicopathological variables. Our results indicated that the prognostic classifier showed

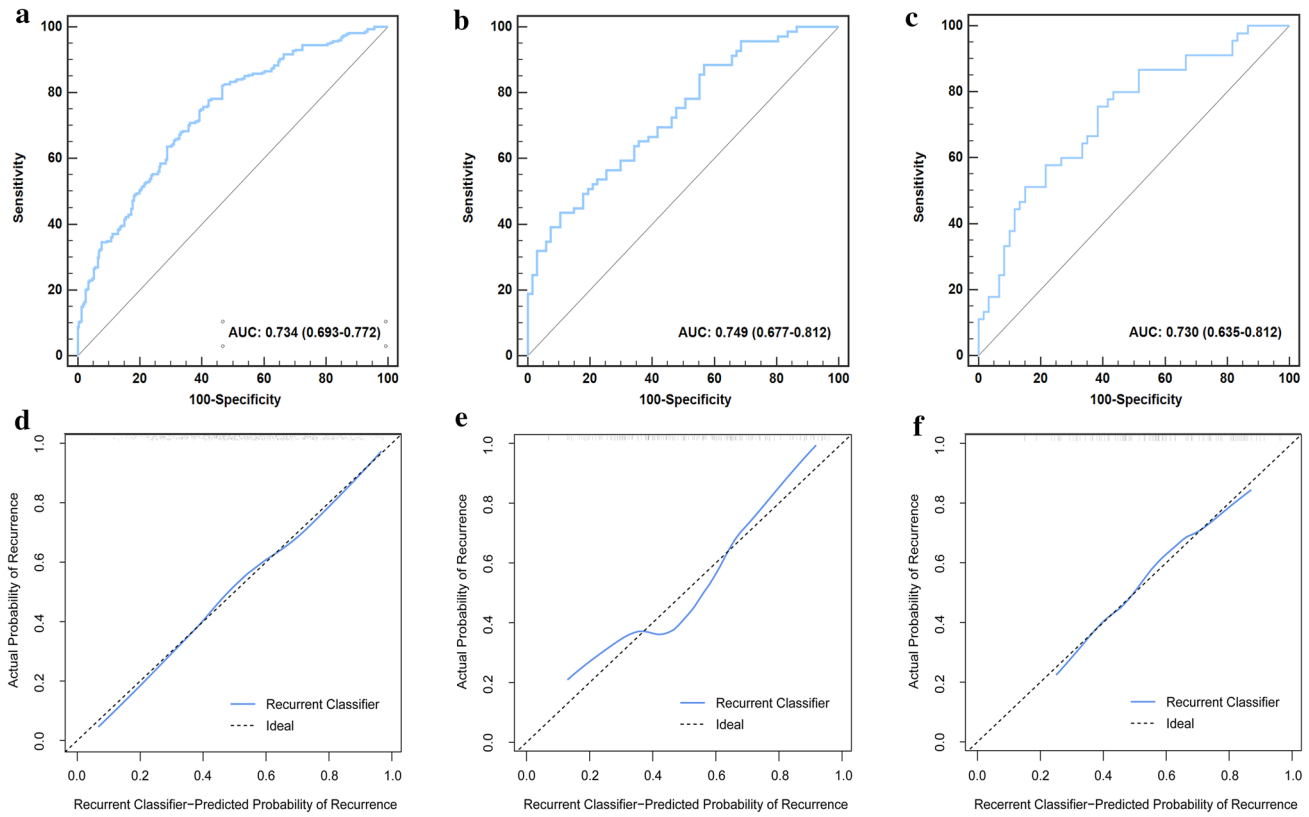

Fig. 4 ROC curves for the proposed recurrent classifier in the training (**a**), internal validation (**b**), and independent validation (**c**) cohorts. Calibration curves for the proposed recurrent classifier in the training (**d**), internal validation (**e**), and independent validation (**f**) cohorts

a higher prediction accuracy for 5-year OS than the TNM staging (0.726, 95% CI: 0.680–0.769 vs. 0.675, 95% CI: 0.627–0.720). The recurrent classifier also showed a higher prediction accuracy for 3-year recurrence than the TNM staging (0.734, 95% CI: 0.693–0.772 vs. 0.655, 95% CI: 0.613–0.710). Compared with BCLC staging system commonly used in clinical practice, our classifiers could also better predict the prognosis (0.726, 95% CI: 0.680–0.769 vs. 0.594, 95% CI: 0.549–0.637) and recurrence (0.734, 95% CI: 0.693–0.772 vs. 0.576, 95% CI: 0.531–0.619) of HCC. Furthermore, the results showed that the proposed clinicopathological characteristics and immunomarkers-based prognostic and recurrent classifiers showed better prediction performance than the classifiers developed based on clinicopathological characteristics alone in all cohorts. Overall, the developed prognostic and recurrent classifiers could stratify patients into high- and low-probability survival groups and better predict clinical outcome and HCC recurrence than the classifiers which are constructed by the conventional clinicopathological variables, and would facilitate the individualized HCC patient management accurately.

Stratifying survival and tumor recurrence in HCC patients using gene expression profiling has already been studied previously. However, until now, no reliable and consistent molecular profile signature for predicting survival and recurrence has been reported [23]. On the other hand, uncertain reproducibility, high testing cost, and utilization of fresh tissue restrict its routine clinical application [24]. By contrast, IHC staining has the advantages of being convenient, rapid, cost-effective, and user-friendly.

Several studies have been published on the prediction of prognosis or recurrence in HCC. Two studies evaluated the prognosis and tumor recurrence in patients with HCC using the emerging radiomics method [25, 26]. Nevertheless, the reproducibility and stability of the radiomics study are still questionable currently. Srivastava et al. [24] proposed a morpho-molecular prognostic model to predict the survival for HCC patients. The model novelly combined the expression of several HCC proteins and clinicopathological features. The AUC for the morpho-molecular prognostic model was lower than the prognostic classifier developed in our current study. Besides that study, only assessed 13 protein IHC expression markers. The newer markers developed recently were not included in that study.

Post-operative HCC recurrence due to residual intrahepatic metastasis or multicentric origin occurs common in the initial 1–2 years [27]. It peaks early, and the risk of tumor recurrence decreases with prolongation of the post-operative

period [28]. Thus, adjuvant therapy such as sorafenib, radiofrequency ablation (RFA), transcatheter arterial chemoembolization (TACE), or other newly developed therapy could be potentially applied during this period. The recurrent classifier in our study provides clinicians a useful tool to predict recurrent probability for HCC patients. Thus, the high-predicted recurrent probability HCC patients, as noticed in the recurrent classifier, might benefit from more aggressive therapeutic regimens. In clinical work, pathologists could stain and score the immunomarkers included in the developed prognostic and recurrent classifiers and then bring the staining and clinicopathological characteristics index into the classifiers' formula for calculation. By stratifying patients into different risk groups according to the calculated results of classifiers, clinicians could decide whether these patients should receive any adjuvant therapy post-operatively.

Compared with other traditional machine-learning algorithms, e.g., decision trees, random forest, and artificial neural network, support vector machine (SVM), and multivariable regression methods were suitable for classifying problems with a relatively high-dimensional featured data set and low-dimensional patient sample size [29]. We performed a pilot study using 200 patients randomly selected from the overall patient population by constructing two prognostic classifiers using the SVM and multivariable regression methods, called SVM-p classifier and MR-p classifier. Prediction difference was observed between these two prognostic classifiers, although it was not significant (SVM-p classifier: AUC, 0.730, 95% CI: 0.668–0.786; MR-p classifier: AUC, 0.767, 95% CI: 0.707–0.819; $p = 0.0975$, Delong test). Thus, we used the MLR method as the final model construction algorithm in this study. To construct a prognostic classifier using the optimal factors, we used the mRMR feature selection algorithm to integrate clinicopathologic characteristics with immunomarkers. The mRMR algorithm was suitable to select the optimal factor combination from high-dimensional data.

In our study, we selected 29 immunomarkers related to different aspects of HCC tumorigenesis, development, and invasiveness, i.e., markers for angiogenesis, apoptosis, cell cycle regulation, DNA repair protein, mismatch repair protein, proto-oncogene, etc. These markers could be used as candidates to predict HCC prognosis and recurrence. In our study, the proposed prognostic classifier integrated seven immunomarkers: BAX, CD34, CEA, E-cadherin, MMP-9, S100A9, and TYSY; and three clinicopathologic characteristics. In contrast, the tumor recurrent classifier integrated six immunomarkers: CD34, E-cadherin, HRas, PCNA, p53, and S100A9, and five clinicopathologic characteristics. Among all, three markers, CD34, E-cadherin, and S100A9, were present in both prognostic and recurrent classifiers. The study indicates that CD34, E-cadherin, and S100A9 could probably be regarded as powerful predictors of HCC prognosis and recurrence.

Our study had several limitations. First, microvascular invasion (MVI) should be included and tested as an important clinicopathological feature affecting the prognosis of HCC. In our study, all cases occurred between October 2010 and December 2013. During this period, MVI was not a mandated indicator in pathology report in china until the evidence-based practice guidelines for standardized pathological diagnosis of primary liver cancer in China (2015 edition) was proposed [30]. Therefore, MVI could not be really reflected and was not included in the study. However, in view of the importance of MVI, we would further analyze the recent cases that sampled according to standardization and incorporate MVI into our classifiers in the future. Second, the proposed prognostic and tumor recurrent classifiers demonstrated satisfactory predictive accuracy; however, both classifiers were restricted by the current available immunomarker antibodies for the IHC stain. With the development of molecular techniques, newer markers with higher predictive value would be available in the near future. Thus, we should consider repeating this study to enhance our classifiers using newly valuable markers next time. Third, though the study included patients from two independent institutions, all patients were from one province in the country, where HBV infection is known as the leading cause of HCC. Therefore, our proposed prognostic and tumor recurrent classifiers should be further tested by more independent data from different regions, provinces, and even different countries, as well as non-HBV-induced HCC population.

In conclusion, the present study established prognostic and tumor recurrent classifiers for patients with HCC after surgical excision. The developed prognostic and recurrent classifiers would be useful in predicting patients' prognosis and tumor recurrence post-operatively. By stratifying patients into different risk groups, the proposed classifiers could help clinicians determine whether these HCC patients require any adjuvant therapy post-operatively.

Acknowledgements We thank the patients who participated in this study and the support from their families.

Author contributions XZ, YB, SZ, and JL conceived and designed the experiments. XZ, YB, BZ, LX, HZ, and LX performed all the experiments. LX, SF, PY, and TN analyzed the data. XZ, LX, and JL wrote the manuscript. All authors read and approved the final manuscript.

Funding This study was funded by the National S&T Major Project (no. 2017ZX10203205), National High-tech R&D Program for Young Scientists by the Ministry of Science and Technology of China (Grant no. 2015AA020917), National Key Research Plan by the Ministry of Science and Technology of China (Grant no. 2016YFC0104507), Natural Science Foundation of China (NSFC Grant no. 81871351), Zhejiang Medical and Health Science and Technology Project (no. 2018KY389).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval The study was approved by the Ethics Committee of the First Affiliated Hospital, Zhejiang University School of Medicine and the Second Affiliated Hospital, Zhejiang University School of Medicine. The study was in accordance with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Patients provided written informed consent before undertaking any study-related procedures.

References

1. El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology 2007;132:2557–2576
2. Forner A. Hepatocellular carcinoma surveillance with miRNAs. Lancet Oncol 2015;16:743–745
3. Chen XP, Qiu FZ, Wu ZD, Zhang ZW, Huang ZY, Chen YF. Long-term outcome of resection of large hepatocellular carcinoma. Br J Surg 2006;93:600–606
4. Vauthey JN, Lauwers GY, Esnaola NF, Do KA, Belghiti J, Mirza N, et al. Simplified staging for hepatocellular carcinoma. J Clin Oncol 2002;20:1527–1536
5. Kudo M, Chung H, Haji S, Osaki Y, Oka H, Seki T, et al. Validation of a new prognostic staging system for hepatocellular carcinoma: the JIS score compared with the CLIP score. Hepatology 2004;40:1396–1405
6. Kudo M, Chung H, Osaki Y. Prognostic staging system for hepatocellular carcinoma (CLIP score): its value and limitations, and a proposal for a new staging system, the Japan Integrated Staging Score (JIS score). J Gastroenterol 2003;38:207–215
7. Llovet JM, Bru C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. Semin Liver Dis 1999;19:329–338
8. Tateishi R, Yoshida H, Shiina S, Imamura H, Hasegawa K, Teratani T, et al. Proposal of a new prognostic model for hepatocellular carcinoma: an analysis of 403 patients. Gut 2005;54:419–425
9. Manghisi G, Elba S, Mossa A, Giorgio A, Aloisio V, Perrotta A, et al. A new prognostic system for hepatocellular carcinoma: a retrospective study of 435 patients: the Cancer of the Liver Italian Program (CLIP) investigators. Hepatology 1998;28:751–755
10. Okuda K, Ohtsuki T, Obata H, Tomimatsu M, Okazaki N, Hasegawa H, et al. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. Cancer 1985;56:918–928
11. Chaudhary K, Poirion OB, Lu L, Huang S, Ching T, Garmire LX. Multimodal meta-analysis of 1494 hepatocellular carcinoma samples reveals significant impact of consensus driver genes on phenotypes. Clin Cancer Res 2019;25:463–472
12. Zhu CQ, Shih W, Ling CH, Tsao MS. Immunohistochemical markers of prognosis in non-small cell lung cancer: a review and proposal for a multiphase approach to marker evaluation. J Clin Pathol 2006;59:790–800
13. Lim HY, Sohn I, Deng S, Lee J, Jung SH, Mao M, et al. Prediction of disease-free survival in hepatocellular carcinoma by gene expression profiling. Ann Surg Oncol 2013;20:3747–3753
14. Zhu ZH, Sun BY, Ma Y, Shao JY, Long H, Zhang X, et al. Three immunomarker support vector machines-based prognostic classifiers for stage IB non-small-cell lung cancer. J Clin Oncol 2009;27:1091–1099
15. De Jay N, Papillon-Cavanagh S, Olsen C, El-Hachem N, Bontempi G, Haibe-Kains B. mRMRe: an R package for parallelized mRMR ensemble feature selection. Bioinformatics 2013;29:2365–2368
16. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models a framework for traditional and novel measures. Epidemiology 2010;21:128–138
17. Zlobec I, Steele R, Terracciano L, Jass JR, Lugli A. Selecting immunohistochemical cut-off scores for novel biomarkers of progression and survival in colorectal cancer. J Clin Pathol 2007;60:1112–1116
18. Delong ER, Delong DM, Clarkepearson DI. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–845
19. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: the Hosmer-Lemeshow test revisited. Crit Care Med 2007;35:2052–2056
20. Goel MK, Khanna P, Kishore J. Understanding survival analysis: Kaplan-Meier estimate. Int J Ayurveda Res 2010;1:274–278
21. Liu CP, Zhang JH, Zheng SC, Liu J, Guo JC. A novel clinical multidimensional transcriptome signature predicts prognosis in bladder cancer. Oncol Rep 2018;40:2826–2835
22. Marrero JA, Fontana RJ, Barrat A, Askari F, Conjeevaram HS, Su GL, et al. Prognosis of hepatocellular carcinoma: comparison of 7 staging systems in an American cohort. Hepatology 2005;41:707–715
23. Lee JS, Chu IS, Heo J, Calvisi DF, Sun Z, Roskams T, et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology 2004;40:667–676
24. Srivastava S, Wong KF, Ong CW, Huak CY, Yeoh KG, Teh M, et al. A morpho-molecular prognostic model for hepatocellular carcinoma. Br J Cancer 2012;107:334–339
25. Zheng BH, Liu LZ, Zhang ZZ, Shi JY, Dong LQ, Tian LY, et al. Radiomics score: a potential prognostic imaging feature for postoperative survival of solitary HCC patients. BMC Cancer 2018;18:1148
26. Zhou Y, He L, Huang Y, Chen S, Wu P, Ye W, et al. CT-based radiomics signature: a potential biomarker for preoperative prediction of early recurrence in hepatocellular carcinoma. Abdom Radiol (NY) 2017;42:1695–1704
27. Hui T, Chuah T, Low H, Tan CH. Predicting early recurrence of hepatocellular carcinoma with texture analysis of preoperative MRI: a radiomics study. Clin Radiol 2018;73:11–16
28. Sakon M, Umeshita K, Nagano H, Eguchi H, Kishimoto S, Miyamoto A, et al. Clinical significance of hepatic resection in hepatocellular carcinoma: analysis by disease-free survival curves. Arch Surg 2000;135:1456–1459
29. Vapnik VN. An overview of statistical learning theory. IEEE Trans Neural Netw 1999;10:988–999
30. Cong WM, Bu H, Chen J, Dong H, Zhu YY, Feng LH, et al. Practice guidelines for the pathological diagnosis of primary liver cancer: 2015 update. World J Gastroenterol 2016;22:9279–9287

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Xiuming Zhang[1] · Yanfeng Bai[1] · Lei Xu[2,3] · Buyi Zhang[4] · Shi Feng[1] · Liming Xu[1] · Han Zhang[1] · Linjie Xu[1] · Pengfei Yang[2,3] · Tianye Niu[2,3] · Shusen Zheng[5,6,7] · Jimin Liu[8]

1 Department of Pathology, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310003, China

2 Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou 310020, China

3 Institute of Translational Medicine, Zhejiang University, Hangzhou 310029, China

4 Department of Pathology, The Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310009, China

5 Division of Hepatobiliary and Pancreatic Surgery, Department of Surgery, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310003, China

6 NHFPC Key Laboratory of Combined Multi-organ Transplantation, The First Affiliated Hospital, Zhejiang University, Hangzhou 310003, China

7 Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou 310003, China

8 Department of Pathology and Molecular Medicine, Faculty of Health Sciences, McMaster University, Hamilton, ON L8S 4K1, Canada