



OPEN

## Identification of BHLHE40 expression in peripheral blood mononuclear cells as a novel biomarker for diagnosis and prognosis of hepatocellular carcinoma

Pattapon Kunadirek<sup>1</sup>, Chaiyaboot Ariyachet<sup>1</sup>, Supachaya Sriphoosanaphan<sup>2</sup>, Nutcha Pinjaroen<sup>3</sup>, Pongserath Sirichindakul<sup>4</sup>, Intawat Nookaew<sup>5</sup>, Natthaya Chuaypen<sup>1</sup>✉ & Pisit Tangkijvanich<sup>1</sup>✉

Novel and sensitive biomarkers is highly required for early detection and predicting prognosis of hepatocellular carcinoma (HCC). Here, we investigated transcription profiles from peripheral blood mononuclear cells (PBMCs) of 8 patients with HCC and PBMCs from co-culture model with HCC using RNA-Sequencing. These transcription profiles were cross compared with published microarray datasets of PBMCs in HCC to identify differentially expressed genes (DEGs). A total of commonly identified of 24 DEGs among these data were proposed as cancer-induced genes in PBMCs, including 18 upregulated and 6 downregulated DEGs. The KEGG pathway showed that these enriched genes were mainly associated with immune responses. Five up-regulated candidate genes including BHLHE40, AREG, SOCS1, CCL5, and DDIT4 were selected and further validated in PBMCs of 100 patients with HBV-related HCC, 100 patients with chronic HBV infection and 100 healthy controls. Based on ROC analysis, BHLHE40 and DDIT4 displayed better diagnostic performance than alpha-fetoprotein (AFP) in discriminating HCC from controls. Additionally, BHLHE40 and DDIT4 had high sensitivity for detecting AFP-negative and early-stage HCC. BHLHE40 was also emerged as an independent prognostic factor of overall survival of HCC. Together, our study indicated that BHLHE40 in PBMCs could be a promising diagnostic and prognostic biomarker for HBV-related HCC.

Hepatocellular carcinoma (HCC), one of the malignant tumors with high heterogeneity, is a leading cause of cancer-related deaths worldwide, especially where chronic hepatitis B virus (HBV) infection is common<sup>1</sup>. Accurate assessment of HCC risk at the early state is very important to optimize the opportunity in receiving curative health care interventions, such as surgical and ablative therapies. In addition, the prognosis of patients with HCC remains unsatisfactory due to the aggressiveness and high recurrence rates of the cancer<sup>1</sup>. At present, alpha-fetoprotein (AFP), a fetal-specific glycoprotein is the most widely used biomarker for HCC screening. Despite its routine use in clinical practice, AFP provides a low sensitivity of 58–68% in detecting an early-stage HCC and its level might be elevated in non-malignant chronic liver disease<sup>2</sup>. Therefore, obtaining reliable serum biomarkers for early diagnosis and prognostic prediction are highly required to improve the outcome and overall survival of patients with HCC.

<sup>1</sup>Department of Biochemistry, Faculty of Medicine, Center of Excellence in Hepatitis and Liver Cancer, Chulalongkorn University, Bangkok 10330, Thailand. <sup>2</sup>Division of Gastroenterology, Department of Medicine, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand. <sup>3</sup>Department of Radiology, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand. <sup>4</sup>Department of Surgery, Faculty of Medicine, Chulalongkorn University, Bangkok 10330, Thailand. <sup>5</sup>Department of Biomedical Informatics, College of Medicine, University of Arkansas for Medical Sciences, Little Rock 72205, USA. ✉email: natthaya.c@chula.ac.th; pisittkvn@yahoo.com

Biomarkers produced by cancer cells including altered gene expression and methylation could be identified in the adjacent body fluid or blood circulation, leading to a new approach for a minimally invasive early cancer detection<sup>3,4</sup>. Indeed, growing evidence has revealed an important role of peripheral blood mononuclear cells (PBMCs) as novel circulating sources that are closely correlated with the pathogenesis of various malignancies<sup>5</sup>. In this context, recent studies demonstrated that changes of gene expression and methylation profiles in PBMCs were observed in patients with non-small cell lung cancer, renal cell carcinoma and breast cancer<sup>6–8</sup>. Regarding HCC, it was shown that the expression profiles of PBMCs differed significantly between patients with or without cancer and could be used as a surrogate approach for the assessment of tumor infiltrating lymphocytes<sup>9</sup>. In a previous study, a co-culture model was performed to investigate the alteration of PBMCs in HCC and identified the alteration of checkpoint inhibitor marker on PBMCs as a prognostic marker for HCC<sup>10</sup>. In addition, gene profiling in PBMCs detected by RNA-sequencing (RNA-Seq) could provide a potential tool for the diagnosis of advanced HCC with metastasis<sup>11</sup>. Together, these data have indicated that a co-culture model might be used to mimic the interaction of cancer and PBMCs, and this may access and identify altered genes in PBMCs from communicating with HCC as a useful marker in the detection and prognostication of HCC.

In this study, we examined whether secretion from cancer cells could induce gene expression in circulating WBCs of patients with HCC. To this end, we investigated transcriptional profiles of PBMCs derived from co-culture model to identify differential genes and cross comparison with previous studies<sup>12</sup>, resulted novel diagnostic biomarkers. The performance of these biomarkers was further validated in PBMCs of patients with HBV-related HCC in comparison with non-cancer controls by qRT-PCR. Finally, the prognostic role of these candidate genes in terms of overall survival of patients with HCC was also investigated.

## Results

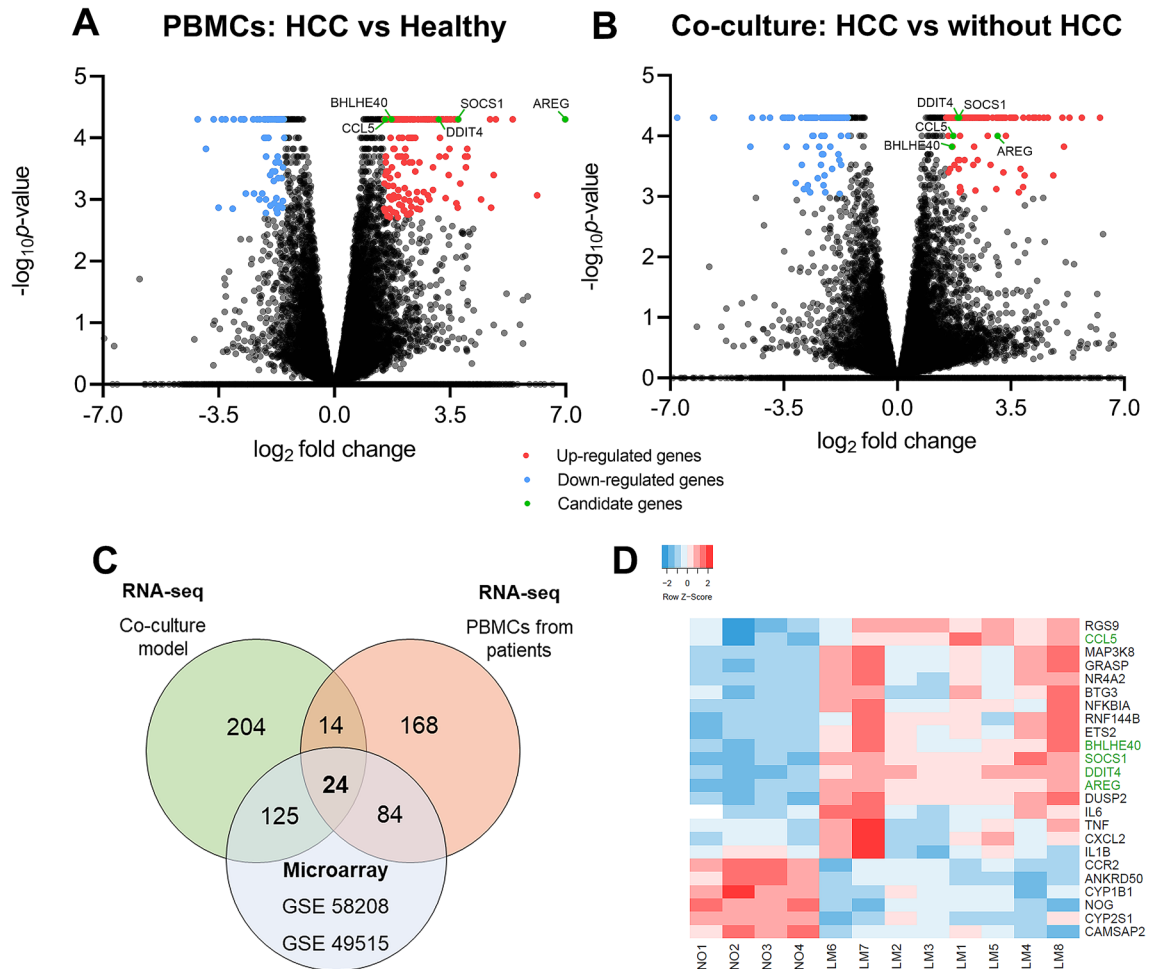
**Integrated gene expression analysis of PBMCs.** To investigate the transcription profiles, PBMCs from 8 patients with HCC and 4 healthy controls, as well as PBMCs from 3 healthy controls co-cultured with HCC cancer cells (Huh7) were collected to perform RNA-Seq. Baseline characteristics of patients and healthy controls were presented in Supplementary Table S1. Our results showed that a total of 290 genes were identified as differentially expressed genes (DEGs) in PBMCs of patients with HCC compared with healthy controls, which included 213 up-regulated ( $P < 0.05$ ,  $\log_2FC > 1.5$ ) and 77 down-regulated genes ( $P < 0.05$ ,  $\log_2FC < 1.5$ ; Fig. 1A). In co-culture of PBMCs with Huh7 cells, we found a total of 367 DEGs with 222 up-regulated genes and 145 down-regulated genes in PBMC co-culture with HCC compared with PBMCs without HCC (Fig. 1B).

To identify whether the DEGs of RNA-Seq data from PBMCs of HCC represent as cancer-induced genes, DEGs from patients with HCC and co-culture model were compared with published microarray data sets (GSE 58208 and 49519)<sup>12</sup> using Connection Up-and Down-Regulation Expression Analysis of Microarrays eXtension (CU-DREAMX)<sup>13</sup>. The intersection among three data sets were shown in Venn diagram (Fig. 1C). Our results showed that a total of 24 DEGs with 18 up-regulated and 6 down-regulated genes were overlapped among these data (Fig. 1C), and intersected genes across data were listed in Supplementary Table S2. To assess whether the above-mentioned 24 intersected genes differed significantly between patients with HCC and healthy controls, a hierarchical clustering using heatmap analysis was further performed. In this respect, the results demonstrated that these two groups could be clearly discriminated by these genes (Fig. 1D). In addition, heatmap analysis of 24 intersected genes in PBMCs from co-culture model and published microarray data were shown in Supplementary Fig. S1. Together, these 24 DEGs in PBMCs were speculated as cancer-induced genes due to the alteration of genes when interacting with HCC in the co-culture model and presenting in PBMCs of patients with HCC.

**Functional gene annotation and pathway enrichment analysis.** Functional enrichment analyses have shown to play an important role in the identification of biological characteristics in transcriptome data. In this study, we performed Gene Ontology (GO) and gProfiler analysis to identify the functional and signaling pathway of DEGs. The majority of DEGs were significantly enriched in the molecular function such as cytokine activity and biological process such as positive regulation of leucocyte cell–cell adhesion, which involved in immune regulation (Fig. 2). Furthermore, enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were similarly involved in immunological response, which included tumor necrosis factor (TNF), pro-inflammatory interleukin (IL)-17 and toll-like receptor signaling pathways.

**Selection and validation of candidate biomarkers in PBMCs of HCC patients.** To investigate whether the identified DEGs in PBMCs could be used as biomarkers in clinical setting, 5 upregulated candidate genes that were consistently up-regulated across PBMCs samples in our RNA-seq analysis were selected. These genes including BHLHE40, AREG, SOCS1, CCL5 and DDIT4 (green labelled in Fig. 1A,B,D) were chosen for validation by qRT-PCR method. The validated cohort consisted of 100 patients with HBV-related HCC, 100 patients with chronic hepatitis B (CHB) and 100 healthy controls. Baseline characteristics in each group were shown in Table 1. Among these individuals, healthy controls and the non-HCC group were matched for age and gender with patients with HCC. Patients with HCC had significantly higher in serum albumin, aspartate aminotransferase (AST), alkaline phosphatase and AFP levels than patients without HCC. Additionally, patients with HCC had a higher proportion of cirrhosis and lower levels of platelet count compared with the non-HCC group. There was no significant difference between groups in terms of serum alanine aminotransferase (ALT) and total bilirubin levels.

Our results demonstrated that BHLHE40, AREG, SOCS1, CCL5 and DDIT4 expression in patients with HCC were higher than those detected in the non-HCC group and healthy controls (Fig. 3). The average relative expression level of BHLHE40 in the HCC group ( $6.86 \pm 4.89$ ) were significantly higher than the non-HCC group ( $0.62 \pm 3.65$ ,  $P < 0.001$ ) and healthy controls ( $0.00 \pm 3.44$ ,  $P < 0.001$ ). Similarly, AREG levels in the HCC group

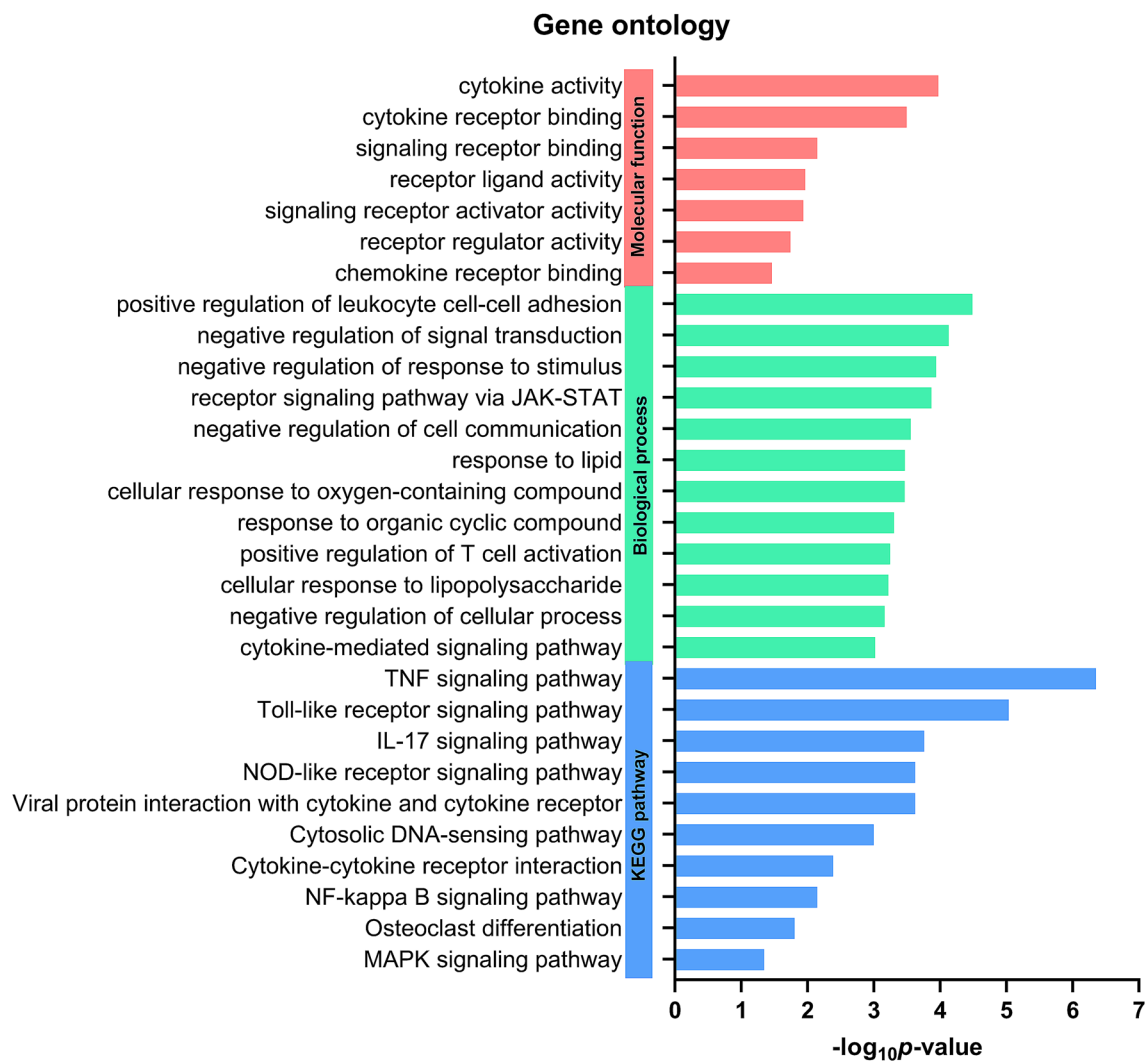


**Figure 1.** Transcriptome profiling of cancer-induced genes and integration transcription profiles of PBMCs. **(A)** volcano plot presents differentially expressed genes in PBMCs compared with healthy controls. **(B)** Volcano plot presents differentially expressed genes in PBMCs of co-culture compared with control. Volcano plot shows  $-\log_{10}P$ -values on the y-axis and fold change expressed as  $\log_2$  on the x-axis. Red scatter dots represent up-regulated genes with  $P \leq 0.05$  and  $\log_2$  fold change  $\geq 1.5$ . Blue scatter dots represent down-regulated genes with  $P \leq 0.05$  and  $\log_2$  fold change  $\leq -1.5$ . Green scatter dots represent a candidate of cancer-induced genes for validation in PBMCs. **(C)** Venn diagram represents intersecting genes between 3 transcription profiles of PBMCs including PBMCs from patients with HCC, PBMCs from co-culture model and published microarray data. **(D)** Heatmap of cancer-induced genes expression in PBMCs of patients with HCC. Candidate of cancer-induced genes for validation are labels in green.

were significantly higher than the non-HCC group ( $2.34 \pm 2.53$  vs  $0.41 \pm 3.92$ ,  $P < 0.001$ ) and healthy controls ( $0.00 \pm 4.88$ ,  $P < 0.001$ ). For SOCS1 expression level, there was significantly higher in patients with HCC than healthy controls ( $2.55 \pm 3.28$  vs  $0.00 \pm 5.39$ ,  $P < 0.001$ ), but did not reach statistical significance when compared with the non-HCC group ( $1.13 \pm 4.94$ ,  $P = 0.053$ ). For CCL5, the levels of this gene in the HCC group ( $4.32 \pm 4.81$ ) were significantly different from the non-HCC ( $1.25 \pm 4.46$ ,  $P < 0.001$ ) and healthy controls ( $0.00 \pm 4.19$ ,  $P < 0.001$ ). A similar trend was found for DDIT4 expression in the HCC group ( $6.62 \pm 4.48$ ) when compared with the non-HCC group ( $0.16 \pm 3.42$ ,  $P < 0.001$ ) and healthy controls ( $0.00 \pm 3.37$ ,  $P < 0.001$ ). There was no statistically significant difference of these 5 genes between patients without HCC and healthy controls.

In subgroup analysis, patients with CHB were divided into the cirrhotic ( $n = 11$ ) and non-cirrhotic ( $n = 89$ ) groups. In Supplementary Fig. S2, there was no significant difference in the expression of all studied genes between the cirrhotic and non-cirrhotic groups. Of note, the mean relative expression levels of BHLHE40 and DDIT4 in the HCC group were significantly higher than the cirrhotic group ( $P < 0.001$ ). These data might indicate that the expression levels of BHLHE40 and DDIT4 in PBMCs could effectively distinguish HCC from cirrhosis.

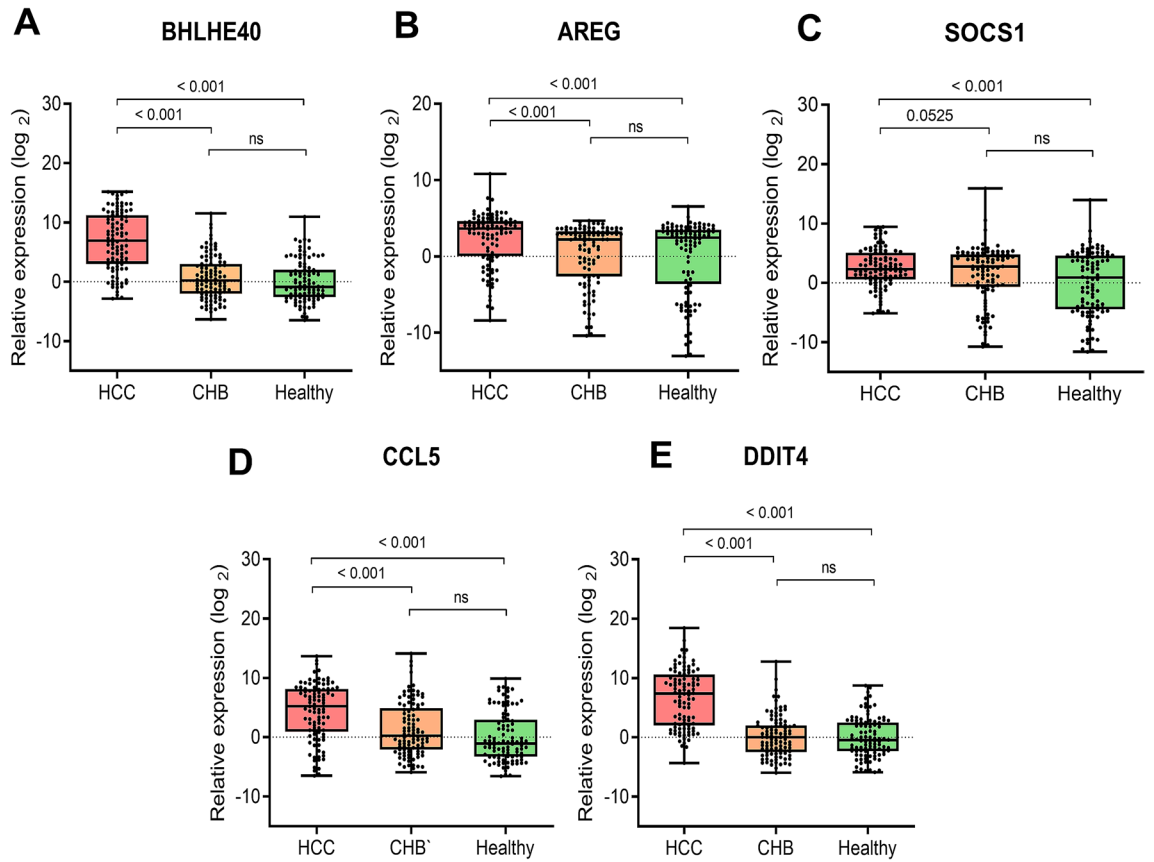
**Selected candidate genes as diagnostic markers of HCC.** To investigate a diagnostic performance of candidate genes in discriminating HCC from non-HCC (including patients with CHB and healthy controls), the ROC curves were calculated. The area under the ROC curve (AUROC) was 0.83 [95% confidence interval (CI); 0.78–0.89,  $P < 0.001$ ] for BHLHE40, 0.69 (95% CI 0.62–0.77,  $P < 0.001$ ) for AREG, 0.54 (95% CI 0.46–0.62,  $P = 0.363$ ) for SOCS1, 0.69 (95% CI 0.61–0.76,  $P < 0.001$ ) for CCL5, 0.85 (95% CI 0.80–0.90,  $P < 0.001$ ) for DDIT4 and 0.81 (95% CI 0.75–0.87,  $P < 0.001$ ) for AFP (Fig. 4). The ROC curves of HCC vs healthy controls and CHB



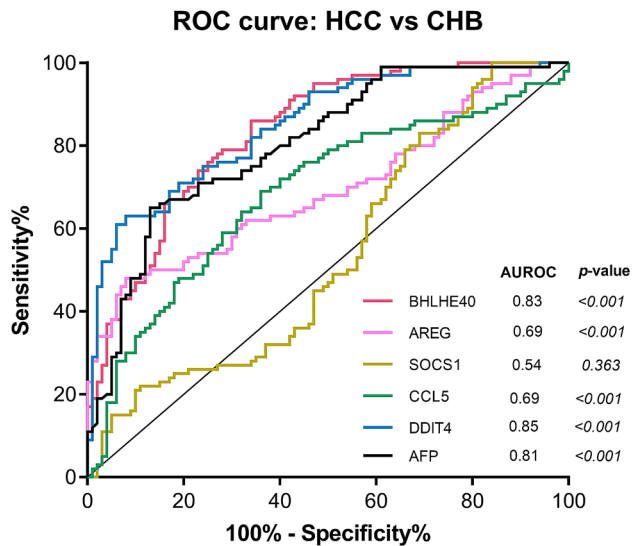
**Figure 2.** Gene Ontology (GO) analysis. Cancer-induced genes show molecular pathways and immune response patterns. Enriched molecular function, biological process and KEGG pathways with a  $p \leq 0.05$  are displayed.

Baseline characteristics	Healthy controls (n = 100)	Patients without HCC (n = 100)	Patients with HCC (n = 100)	P
Age (years)	56.3 ± 3.4	56.7 ± 8.4	57.7 ± 8.6	0.053
Gender (Male)	80 (80.0)	85 (85.0)	85 (85.0)	0.549
Total bilirubin (mg/dL)		0.9 ± 0.7	1.0 ± 0.7	0.729
Serum albumin (g/dL)		4.1 ± 0.6	3.6 ± 0.5	< 0.001*
Aspartate aminotransferase (IU/L)		36.6 ± 30.2	64.5 ± 53.5	< 0.001*
Alanine aminotransferase (IU/L)		42.9 ± 42.3	44.8 ± 25.8	0.709
Alkaline phosphatase (IU/L)		77.3 ± 50.6	125.3 ± 70.8	< 0.001*
Platelet count (10 <sup>9</sup> /L)		223.2 ± 74.6	176.0 ± 93.9	0.002*
Alpha fetoprotein (ng/mL)		8.5 ± 13.9	6928.1 ± 28,444.2	0.016*
Presence of cirrhosis		11 (11.0)	80 (80.0)	< 0.001*
BCLC stage (0-A/B/C)		-	35(35.0)/50(50.0)/15(15.0)	-

**Table 1.** Baseline characteristic of the validation cohort.



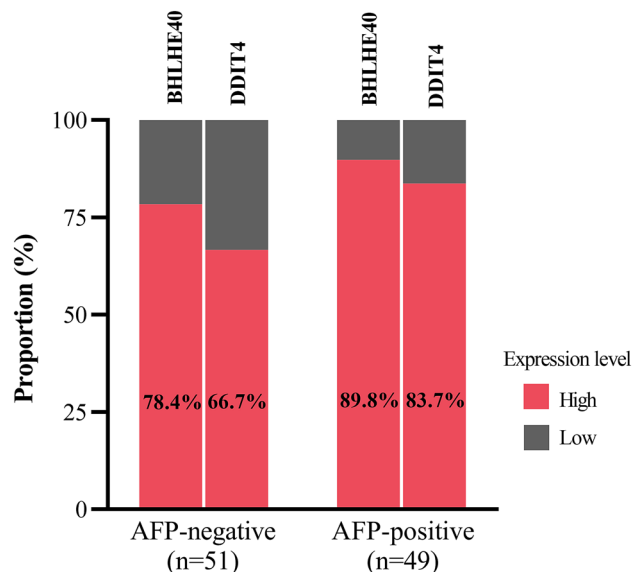
**Figure 3.** Relative expression of candidate cancer-induced genes in PBMCs of patients with HBV-HCC, CHB and healthy controls. (A) BHLHE40, (B) AREG, (C) SOCS1, (D) CCL5 and (E) DDIT4. Relative expression of genes represents as log<sub>2</sub> on y-axis.



**Figure 4.** Receiver operating characteristic (ROC) curves of the cancer-induced genes of PBMCs in differentiating patients with HCC and non-HCC.

vs healthy controls were also analyzed (Supplementary Fig. S3). The data showed that there was a similar trend between HCC vs. non-HCC and HCC vs CHB. Additionally, the expression of these genes was not useful in distinguishing patients with CHB from healthy controls.

According to the ROC analysis, BHLHE40 and DDIT4 were considered the best biomarkers among the studied candidate genes, their diagnostic role was further assessed. In this respect, the optimal cut-off value of



**Figure 5.** Proportion of BHLHE40 and DDIT4 expression in PBMCs of AFP- and AFP+ patients with HBV-HCC.

BHLHE40 in differentiating HCC from non-HCC was 1.80 with a sensitivity of 84.0% and specificity of 64.0%. Similarly, the cut-off value of DDIT4 was 2.10 with a sensitivity of 75.0% and specificity of 76.0%. The combination of BHLHE40 and DDIT4 without AFP slightly increased AUROC to 0.86 and improved accuracy to 78.24% which was the highest accuracy in this study (Supplementary Table S3). Our results showed that BHLHE40 had the best sensitivity of 84% and AFP had highest specificity of 88%. Furthermore, CU-DREAMX analysis demonstrated that BHLHE40 and DDIT4 of PBMCs were not increased in patients with head and neck cancers and pancreatic cancer (Supplementary Fig. S4). These results suggested that BHLHE40 and DDIT4 could be used as potential biomarkers for HCC.

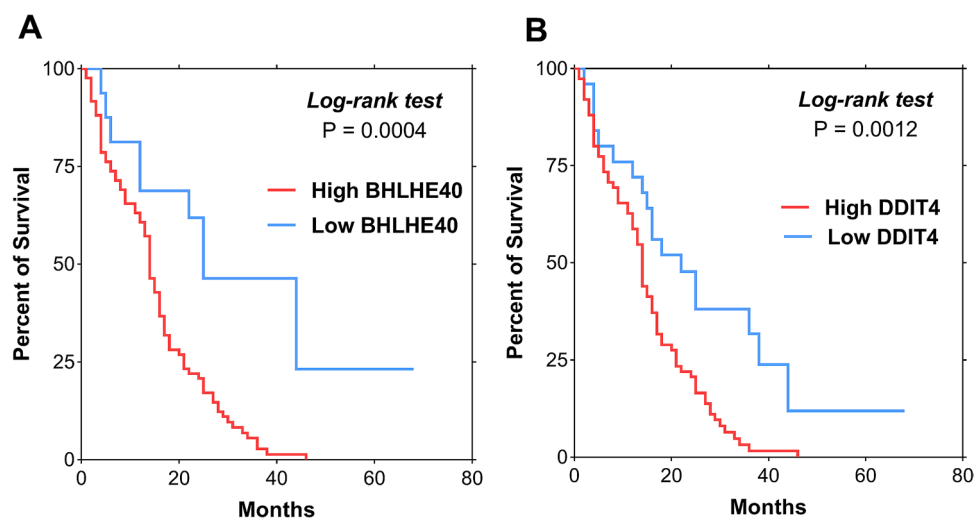
Based on their optimal cut-off values, the correlation of BHLHE40, AREG, SOCS1, CCL5 and DDIT4 expression with clinical parameters are shown in Supplementary Table S4.

**The diagnostic role of BHLHE40 and DDIT4 in AFP-negative HCC and small HCC.** Among patients with HCC, there was a strong correlation between BHLHE40 and DDIT4 ( $r = 0.826$ ;  $P < 0.001$ ). However, either BHLHE40 and DDIT4 was not correlated with AFP values ( $r = 0.197$ ;  $P = 0.050$  and  $r = 0.154$ ;  $P = 0.126$ , respectively). Using the normal upper limit of AFP (20 ng/mL) as a reference, there were 51 (51%) and 49 (49%) patients with AFP-negative HCC and AFP-positive HCC, respectively (Fig. 5). Among the AFP-negative group, 78.4% (40/51) of patients had elevated BHLHE40 level ( $\geq 1.8$ ) and 66.7% (34/51) of patients had high DDIT4 level ( $\geq 2.1$ ). For the AFP-positive group, high levels of BHLHE40 and DDIT4 were detected in 89.8% (44/49) and 83.7% (41/49), respectively. Furthermore, 82.6% (42/51) of patients with AFP-negative HCC had elevated BHLHE40 and/or DDIT4 (Supplementary Fig. S5). Among early HCC (stages 0 and A), we found that 27.8% (10/36) patients had elevated AFP concentration, while 86.1% (31/36) and 69.4% (25/35) patients had elevated levels of BHLHE40 and DDIT4, respectively. Together, these results suggested that BHLHE40 and DDIT4 could be promising biomarkers for detecting AFP-negative HCC and early HCC, as well as they could be complementary to AFP in diagnosis of HCC.

**Prognostic performance of BHLHE40 and DDIT4 in patients with HCC.** The potential prognostic values of BHLHE40 and DDIT4 in terms of overall survival were also analyzed. Based on Kaplan–Meier analysis, the median overall survival of patients with low BHLHE40 levels ( $< 1.8$ ) was significantly better than that of patients whose levels were  $\geq 1.8$  (33.8 vs. 15.2 months,  $P < 0.001$  by log rank test) (Fig. 6A). Similarly, the median overall survival of patients with low DDIT4 levels ( $< 2.1$ ) was significantly better than that of patients whose levels were elevated (26.7 vs. 15.0 months,  $P = 0.001$ ) (Fig. 6B).

BHLHE40 and DDIT4 were entered into the multivariate analysis together with other variables that might influence overall survival of the patients. These factors included age, gender, platelet counts, serum TB, AST, ALT, albumin, AFP level, presence of cirrhosis, tumor size and BCLC stage. The multivariate analysis using the Cox proportional hazards model revealed that high BHLHE40 and BCLC stage were independent predictive factors of overall survival. However, DDIT4 was not selected as an independent factor associated with overall survival (Table 2).





**Figure 6.** Kaplan–Meier survival curves for overall survival analysis of patients with HBV-HCC (A) BHLHE40 (B) DDIT4.

Variables	Category	Overall survival			
		Univariate analysis		Multivariate analysis	
		OR (95%CI)	P	OR (95%CI)	P
Age (years)	< 60 vs. $\geq$ 60	1.02 (0.67–1.55)	0.924		
Gender	Male vs. Female	1.16 (0.66–2.08)	0.608		
Total bilirubin (mg/dL)	< 1.2 vs. $\geq$ 1.2	0.77 (0.49–1.27)	0.303		
Serum albumin (g/dL)	< 3.5 vs. $\geq$ 3.5	1.02 (0.63–1.65)	0.946		
Aspartate aminotransferase (IU/L)	< 60 vs. $\geq$ 60	1.62 (1.06–2.49)	0.027*	1.34 (0.83–2.16)	0.230
Alanine aminotransferase (IU/L)	< 50 vs. $\geq$ 50	1.63 (1.04–2.53)	0.032*	1.18 (0.73–1.92)	0.493
Platelet count ( $10^9/L$ )	$\geq$ 100 vs. < 100	1.43 (0.85–2.40)	0.178		
Presence of cirrhosis	No vs. Yes	0.97 (0.58–1.64)	0.907		
Alpha fetoprotein (ng/mL)	< 100 vs. $\geq$ 100	1.93 (1.22–3.05)	0.005*	1.46 (0.88–2.42)	0.146
Tumor size (cm)	< 5.0 vs. $\geq$ 5.0	1.53 (1.01–2.32)	0.043*	1.09 (0.63–1.90)	0.757
BCLC stage	0-A vs. B vs. C	1.94 (1.33–2.83)	0.001*	2.00 (1.22–3.27)	0.006*
BHLHE40	< 1.8 vs. $\geq$ 1.8	3.21 (1.59–6.48)	0.001*	3.54 (1.52–8.34)	0.004*
DDIT4	< 2.1 vs. $\geq$ 2.1	2.29 (1.35–3.91)	0.002*	0.94 (0.48–1.82)	0.850

**Table 2.** Variables associated with overall survival in patients with HCC.

## Discussion

Recent advances have revealed that genetic and epigenetic alterations accumulated through repeated destruction and regeneration of the hepatocytes are responsible for the development of HCC<sup>14</sup>. In fact, hepatocarcinogenesis is linked to various etiological factors, which in turn result in aberrant activation of different signaling pathways and imbalance between oncogene activation and tumor suppressor gene inactivation. In Thailand, chronic HBV infection is the most important risk factor for HCC development accountable for at least 60% of all cases<sup>15</sup>. Detection of HCC at an early stage is crucial, which allows the possibility of receiving curative treatment and can improve overall survival. Although AFP is currently the most common biomarker for screening HCC in clinical setting, the overall sensitivity and specificity of this tumor marker are approximately 60% and 80%, respectively<sup>16</sup>. Additionally, its sensitivity declines significantly in detecting early HCC because elevated AFP level is typically correlated with large tumor size, poor tumor differentiation and presence of vascular invasion<sup>2</sup>. In our report, for example, approximately 25% of small HCC were AFP-positive (AFP level  $\geq$  20 ng/mL). Thus, new biomarkers that individually or in complementary with AFP could increase the diagnostic accuracy of early HCC are highly needed.

Given their accessibility, recent data have revealed the alteration of gene expression patterns in PBMCs as a novel source for clinical diagnosis and monitoring in several types of cancers including HCC. In this report, we initially aimed to characterize DEGs of RNA-Seq data derived from PBMCs of patients with HCC compared with healthy controls as a potential diagnostic biomarker of HCC. Additionally, DEGs from co-culture model was generated to mimic human body condition where normal PBMCs were incubated with HCC cell lines. After integrating these results with published microarray database, 24 DEGs with 18 up-regulated and 6 down-regulated

genes were identified. Previous data demonstrated that PBMCs from patients with HCC shared distinct and similar features of gene expression profiles with HCC-infiltrating mononuclear inflammatory cells in the liver<sup>9</sup>. Similarly, a recent report also demonstrated a concordance of transcriptomic analysis between PBMCs and tumor tissues in colorectal cancer<sup>17</sup>. In this context, our results indicate transcriptomic changes in PBMCs induced by HCC could represent a readily accessible biomarker for the tumor microenvironment.

Through functional analysis, DEGs identified from our data integration were shown to be associated with various biological processes, including cytokine activity, cell adhesion, inflammatory responses and immune regulation. Additionally, the enriched KEGG pathways of DEGs were mainly involved in TNF signaling, Toll-like receptor signaling, IL-17 signaling, NOD-like signaling pathways and cytokine-cytokine receptor interactions, most of which are closely related to pathogenesis and progression of HCC<sup>14</sup>. For instance, a recent report demonstrated that TNF- $\alpha$  promoted HCC carcinogenesis through the activation and proliferation of hepatic progenitor cells via TNFR2/STAT3 signaling pathway<sup>18</sup>. Regarding the role of IL-17, it was shown that this pro-inflammatory cytokine secreted by Th17 cells could facilitate tumor growth in vitro and in vivo through IL-6/STAT3 pathway in HBV-related HCC<sup>19</sup>. Additionally, previous data suggested that accumulation of intra-tumoral IL-17 accelerated tumor progression through promoting angiogenesis and its detection in cancerous tissues could serve as a potential prognostic marker of HCC<sup>20</sup>.

Among the identified DEGs, 5 candidate cancer-induced genes including BHLHE40, AREG, SOCS1, CCL5 and DDIT4 were selected for external validation on an independent set of PBMC samples of patients with various stages of HCC. The selection of these DEGs was made based on their homogeneous expression patterns in the integrated data set of PBMCs and is known to be involved in pathogenesis of various malignant tumors. According to ROC analysis, the overall results demonstrated that BHLHE40 and DDIT4 expression in PBMCs represented potential diagnostic biomarkers in distinguishing HCC from the non-HCC group. BHLHE and DDIT4 were also found to be specific markers for HCC compared with other cancers such as head and neck cancers and pancreatic cancer.

BHLHE40, also known as DEC1/Stra13/Sharp2, is a stress-responsive transcription factor directly targeted by hypoxia-inducible factor-1 $\alpha$  (HIF-1 $\alpha$ ) in modulating several cell physiological responses<sup>21</sup>. In Addition, BHLHE40 is emerging as a key regulator of immune response during autoimmunity and various inflammatory conditions<sup>22</sup>. Specifically, BHLHE40 has recently identified as a transcriptional regulator of T cell persistence and activity by coordinating metabolic and epigenetic programming, the mechanism of which is critically essential in immunological functions<sup>23</sup>. Current evidence has also indicated that BHLHE40 involve in regulating cell growth, differentiation, proliferation and apoptosis of several cancers<sup>24</sup>. Thus, dysregulation of BHLHE40 could cause alteration of intracellular homeostasis, leading to abnormal cell proliferation, differentiation and subsequent malignant transformation. In breast cancer, it was shown that BHLHE40 constituted an important signaling to promote tumor metastasis by modulating the secretion of epidermal growth factor (EGF), which is known to induce proliferation and invasion of tumor cells under hypoxia<sup>25</sup>. In gastric cancer, BHLHE40 was necessary for anti-apoptotic activity of tumor cells under hypoxic condition by promoting surviving expression<sup>26</sup>. Moreover, BHLHE40 was upregulated in cancerous tissue compared with normal gastric specimens and its expression enhanced during disease progression from well to poorly differentiated, indicating its associated with tumor differentiation status<sup>27</sup>.

With regard to HCC, BHLHE40 was shown to be activated by HIF-1 $\alpha$ , suggesting its role in adaptation to a hypoxic microenvironment associated with tumor progression<sup>28</sup>. Indeed, adaptation to hypoxia represents a crucial step in the expansion and transformation of rapidly proliferative cancerous cells, including HCC. A systematic review with meta-analysis has revealed that HIF-1 $\alpha$  overexpression is correlated with poor prognosis and aggressive clinicopathological features of HCC<sup>29</sup>. Moreover, it was previously showed in experimental models that overexpression of BHLHE40 promoted epithelial-mesenchymal transition (EMT) process that enhanced metastatic capability of the cancer<sup>30,31</sup>. A previous study also reported that higher BHLHE40 expression was detected in cancerous tissue compared with adjacent normal tissues and might be associated with histological differentiation of HCC<sup>32</sup>. Although these available data suggest that BHLHE40 might be participating in HCC development and progression, they were mostly conducted in cell line and tissue-based experiments. However, the role of circulating BHLHE40 expression as a diagnostic and prognostic biomarker in patients with HCC remains to be explored.

In this study, we demonstrate for the first time that BHLHE40 expression in PBMCs could be used as a promising biomarker for HCC. In particular, BHLHE40 was accurately discriminative of AFP-negative and early HCC and its diagnostic performance was more superior than AFP. These findings indicate the potential use of BHLHE40 as a sensitive biomarker for early HCC, as well as a complementary biomarker with AFP-negative HCC in patients with chronic HBV infection. Regarding its predictive role, Kaplan–Meier analysis showed that BHLHE40 overexpression was positively correlated with poor overall survival in patients with HCC. Additionally, multivariate analysis confirmed that an increased BHLHE40 levels in PBMCs was an independently unfavorable predictor of overall survival. Together, our data provide evidence supporting a novel role of circulating BHLHE40 expression in early detection and prognostic indicator of HCC. Given its strong link to adverse clinical outcome, our results might also suggest that targeting BHLHE40 and its related signaling pathways could be a potential therapeutic approach for HBV-related HCC. The DNA damage inducible transcript 4 (DDIT4, also known as REDD1 or RTP801), ubiquitously expressed at low levels in most human tissues, is induced by several transcription factors in response to various stress stimuli such as hypoxic conditions, metabolic alteration and chronic inflammation<sup>33,34</sup>. Dysfunction of DDIT4 has been shown to be associated with multiple disorders including various types of cancers<sup>35</sup>. It was shown that DDIT4 over-expression provided an advantage on cancer cell survival and metastasis in hypoxic conditions through decreased energy consumption, leading to cancer progression, angiogenesis and resistance to chemotherapy or radiotherapy<sup>36</sup>. A recent in silico analysis demonstrated that high levels of DDIT4 were significantly associated with poor prognosis of hematologic malignancies and



several solid tumors, such as breast, colon and lung cancers<sup>37</sup>. On the contrary, increased DDIT4 expression was associated with an improved prognosis in gastric cancer but was not related to clinical outcome of ovarian cancers<sup>37</sup>. These data apparently indicate that the role of DDIT4 might not be similar among different cancer types and their aggressiveness<sup>35,36</sup>.

Regarding HCC, the potential role of DDIT4 in this type of cancer, especially HBV-related HCC, remains to be determined. Previous data suggested that DDIT4 might be involved in the pathogenesis of HCC, notably through the expression of miR-802 and programmed cell death protein 1 (PD-1)<sup>38</sup>. In our report, we firstly demonstrated that DDIT4 expression was significantly increased in PBMCs of patients with HCC in comparison with the non-HCC group and healthy individuals, suggesting its potential role in HCC carcinogenesis. Of note, DDIT4 was superior to AFP in differentiating early HCC from the non-HCC group, which had a similar trend as observed in BHLHE40. Additionally, high DDIT4 expression was positively correlated with poor survival by univariate analysis. Although its significance was not reached in multivariate analysis. This result might reflect a strong relationship of circulating levels of BHLHE40 and DDIT4 expression demonstrated in our study.

This report had some limitations as being a retrospective study that enrolled relatively small number of patients with HCC. Moreover, this study focused on patients with chronic HBV infection, which might not be relevant to other chronic liver disease, such as chronic HCV infection and fatty liver disease. Additionally, the cut-off values for relative mRNA expressions identified based on the ROC curves could be variable from study to study depending on patient cohorts and laboratory techniques. Finally, the selection of candidate genes for validation using qRT-PCR was based entirely on their relevance to the pathobiology of various cancers, as well as their consistently up-regulated expression in our study. As a result, down-regulated genes identified from RNA-seq analysis, including CCR2 and ANKRD50, were not chosen for further investigation. Despite such limitations, our data demonstrated that circulating BHLHE40 and DDIT4 were differentially expressed in patients with HCC compared to individuals without cancer. Apart from its diagnostic role, circulating BHLHE40 also emerged as an independent prognosis factor of patients with HBV-related HCC. As current knowledge on the role of BHLHE40 and DDIT4 in HCC remains infancy, further studies are needed to confirm our observations and to elucidate the mechanisms by which these genes play important roles in the pathogenesis and aggressiveness of HCC.

## Materials and methods

**Sample collection.** Blood samples were obtained in EDTA tube before treatment procedure from patients with HBV-related HCC, who were diagnosed and treated at King Chulalongkorn Memorial Hospital, Bangkok, Thailand between 2018 and 2020. The diagnostic of HCC was based on the imaging studies results of dynamic computed tomography (CT) or magnetic resonance imaging (MRI) in concordance with the American Association for the Study of Liver Diseases (AASLD) guideline<sup>39</sup>. The demographic and clinical characteristics of patients were collected, which included sex, age, liver function tests, serum AFP level and HCC staging classified by the Barcelona Clinic Liver Cancer (BCLC) system<sup>40</sup>. The blood samples were also obtained from healthy controls and chronic HBV-infected patients without evidence of HCC as control groups.

This study was performed in concordance with the Declaration of Helsinki for the participation of human individuals. The written informed consents were obtained from all patients and all health controls involved in the study. The protocols in this study has been approved by the Institute Ethics Committee of Faculty of Medicine, Chulalongkorn University (IRB No. 313/62).

**PBMCs isolation.** PBMCs were isolated from fresh EDTA blood tube by Ficoll-Hypaque density gradient centrifugation using Percoll PLUS density gradient media (GE Healthcare) at 1500 rpm for 30 min at room temperature and were then washed 2 times with PBS. PBMCs were resuspended in 10% DMSO in fetal bovine serum and stored at  $-80^{\circ}\text{C}$ .

**Cell lines and Co-culture.** Liver cancer cells (Huh7, JCRB0403) were obtained from Nation Institutes of Biomedical Innovation, Health and Nutrition JCRB Cell Bank (Osaka, Japan). Cells were grown with DMEM medium (Gibco) supplemented with 10% FBS at  $37^{\circ}\text{C}$  5%  $\text{CO}_2$  in culture flask. The cells were harvested at 80% confluence by using 0.05% Trypsin with 0.5 mM EDTA and were then washed by using phosphate buffer saline (PBS).

Co-culture liver cancer cells with PBMCs from healthy individuals were performed in Transwell culture six well plates (Costar). The liver cancer cells ( $10^6$  cells) were seeded into lower with DMEM and were then incubated overnight at  $37^{\circ}\text{C}$  and 5%  $\text{CO}_2$ . PBMCs from 3 healthy individuals ( $2 \times 10^6$  cells) were individually plated on transwell membrane and incubated for 4 h at  $37^{\circ}\text{C}$  and 5%  $\text{CO}_2$ . Finally, PBMCs were collected for RNA extractions.

**RNA preparation and sequencing.** A total RNA was extracted from PBMCs using TRIzol reagent (Gibco) according to manufacturer's instruction. For RNA-sequencing, concentration of total RNA samples was quantified using Qubit RNA assay kit (Invitrogen) and RNA integrity was accessed by RNA Electrophoresis with the 2100 Bioanalyzer System (Agilent). After that, RNA samples were then performed library preparation and sequencing by Vishuo Biomedical (Vishuo Biomedical, Singapore). For briefly, library preparation was performed by NEBNext Ultra RNA Library Prep Kit (NEB) according to for Illumina RNA library preparation protocol. mRNA was captured and isolated by NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB). Double strand cDNA was synthesized using random primer, ProtoScript II Reverse Transcriptase and Second strand synthesis Enzyme mix. Then, double strand cDNA was repaired and added a A-tail at both ends of cDNA and was followed by a T-A ligation to add adaptors at both ends of cDNA. Each sample was amplified by PCR for 11 cycles using P5 and P7 primers. The PCR products were validated by 2100 Bioanalyzer and quantified by Qubit 2.0 Fluorometer. After that, libraries with different indices were multiplexed and loaded on an Illumina

HiSeq sequencer (Illumina). Sequencing data was obtained as 2 × 150 bp paired end. Ultimately, we obtained an average of 25 million read pairs per samples, which ranged from 20 to 37 million reads. The percentage of data above Q30 and mapping with reference genes were more than 90% of reads (Supplementary Table S5).

**Data processing.** The paired-end raw reads were performed by FasTQC version 0.11.2<sup>41</sup> to check overall sequencing quality and were then trimmed by Trimmomatic version 0.32<sup>42</sup> to remove the sequencing adaptor and low quality of sequences (lower than Q30). The trimmed reads were aligned to the Homosapiens reference genome (GRCh38) with HISAT2 version 2.1.0<sup>43</sup> software. Aligned reads were assembled into transcripts by StringTie version 1.3.3b<sup>44</sup>. Then DESeq2 version 1.16.1<sup>45</sup> was performed to identify DEGs between PBMCs from patients with HCC and healthy controls, as well as PBMCs from co-culture with HCC and without HCC. The up-regulated and down-regulated genes of DEGs were filtered by setting cut-off at 1.5-foldchange and  $P$ -value ≤ 0.05. The function of DEGs was classified in terms of Gene Ontology (GO) by gProfiler with default settings<sup>46</sup>. The DEGs were then used for cross comparing with microarray data by CU-DREAMX.

**Retrieving data from GenBank and CU-DREAMx analysis.** Gene expression profiles of PBMCs from patients with HCC (GSE 49515 and 58208)<sup>12</sup> were used to compare with RNA-seq data of our study by CU-DREAMX. Briefly, DEGs in PBMCs of these datasets derived from 38 patients with HCC cases and 14 healthy controls were identified ( $P < 0.05$ ) and then compared with DEGs in PBMCs from patients with HCC and co-culture model to identify the intersection of DEGs among three datasets. Additionally, the intersect DEGs in PBMCs of HCC from CU-DREAMX were compared with DEGs in PBMCs from 27 patients with head and neck cancers (GSE39400) and 3 patients with pancreatic cancer (GSE49515).

**Quantitative RT-PCR analysis.** The performance of identified DEGs was further validated in PBMCs of patients with HBV-related HCC and non-HCC individuals by qRT-PCR. The criteria for selecting candidate genes based on the consistency of up-regulated expression in PBMCs of patients with HCC from RNA-seq and known to be involved in pathogenesis of various malignant cancers. In this regard, BHLHE40, AREG, SOCS1, CCL5 and DDIT4 genes were selected. The total RNA was extracted from PBMCs of 100 patients with HBV-related HCC, 100 patients with chronic hepatitis B and 100 healthy controls using TRIzol reagent (Thermo Scientific). Then, cDNA was synthesized using RevertAid First Strand cDNA Synthesis (Thermo Scientific). The qRT-PCR reaction contained 6.25 μL of QPCR Green Master Mix HRox 2 × (Biotech rabbit), 0.25 μL of primers and 1 μL of cDNA and nuclease-free water in a total volume of 12.5 μL. The reactions carried out on a QuantStudio 5 Real-Time PCR System (Applied Biosystems). Primer sequences and thermal cycle condition are presented in Supplementary Table S6. All reactions were performed in duplicate. Positive controls for each target genes and negative controls were included to ensure correct interpretation. The expression of target genes was normalized by β-globin endogenous reference gene. The data are shown in log<sub>2</sub> fold change format.

**Statistical analysis.** Statistical analysis was performed using the Statistical Package for the Social Sciences (SPSS) version 23 (<https://www.ibm.com/analytics/spss-statistics-software>) and GraphPad prism version 8 for Window (<https://www.graphpad.com/scientific-software/prism>). Comparisons between groups were analyzed by Chi's square or Fisher's exact test for categorical variables and by Student's t-test or one-way ANOVA for quantitative variables. Spearman's rank test was used for correlations between parameters. Kaplan–Meier analysis and log-rank test were used for survival analysis. The Cox regression analysis was conducted to identify independent factors associated with overall survival of patients with HCC.  $P$ -value < 0.05 was considered statistically significant. Adjusted  $P$ -value for the multiple hypothesis testing was not performed in this study due to small number of samples. However, we further validate the results with highly sensitive method using qPCR in the independent cohort.

Received: 25 February 2021; Accepted: 12 May 2021

Published online: 27 May 2021

## References

- Kulik, L. & El-Serag, H. B. Epidemiology and management of hepatocellular carcinoma. *Gastroenterology* **156**, 477–491 (2019).
- Galle, P. R. *et al.* Biology and significance of alpha-fetoprotein in hepatocellular carcinoma. *Liver Int.* **39**, 2214–2229 (2019).
- Chen, S. *et al.* Identification of human peripheral blood monocyte gene markers for early screening of solid tumors. *PLoS ONE* **15**, e0230905 (2020).
- Puttipanyalears, C., Kitkumthorn, N., Buranapraditkun, S., Keelawat, S. & Mutirangura, A. Breast cancer upregulating genes in stromal cells by LINE-1 hypermethylation and micrometastatic detection. *Epigenomics* **8**, 475–486 (2016).
- Mosallaei, M. *et al.* PBMCs: A new source of diagnostic and prognostic biomarkers. *Arch Physiol. Biochem.* <https://doi.org/10.1080/13813455.2020.1752257> (2020).
- Kitkumthorn, N., Tuangsintanakul, T., Rattanananyong, P., Tiwawech, D. & Mutirangura, A. LINE-1 methylation in the peripheral blood mononuclear cells of cancer patients. *Clin. Chim. Acta* **413**, 869–874 (2012).
- Showe, M. K. *et al.* Gene expression profiles in peripheral blood mononuclear cells can distinguish patients with non-small cell lung cancer from patients with nonmalignant lung disease. *Cancer Res.* **69**, 9202–9210 (2009).
- Twine, N. C. *et al.* Disease-associated expression profiles in peripheral blood mononuclear cells from patients with advanced renal cell carcinoma. *Cancer Res.* **63**, 6069–6075 (2003).
- Sakai, Y. *et al.* Common transcriptional signature of tumor-infiltrating mononuclear inflammatory cells and peripheral blood mononuclear cells in hepatocellular carcinoma patients. *Can. Res.* **68**, 10267–10279 (2008).

10. Iwata, T. *et al.* PD-L1(+)-MDSCs are increased in HCC patients and induced by soluble factor in the tumor microenvironment. *Sci. Rep.* **6**, 39296 (2016).
11. Shen, Y. *et al.* Screening effective differential expression genes for hepatic carcinoma with metastasis in the peripheral blood mononuclear cells by RNA-seq. *Oncotarget* **8**, 27976–27989 (2017).
12. Shi, M. *et al.* A blood-based three-gene signature for the non-invasive detection of early human hepatocellular carcinoma. *Eur. J. Cancer* **50**, 928–936 (2014).
13. Apontewan, C. & Mutirangura, A. Connection up-and down-regulation expression analysis of microarrays (CU-DREAM): A physiogenomic discovery tool. *Asian Biomed.* **5**, 257–262 (2011).
14. Llovet, J. M. *et al.* Hepatocellular carcinoma. *Nat. Rev. Dis. Primers* **7**, 6 (2021).
15. Tangkijvanich, P., Hirsch, P., Theamboonlers, A., Nuchprayoon, I. & Poovorawan, Y. Association of hepatitis viruses with hepatocellular carcinoma in Thailand. *J. Gastroenterol.* **34**, 227–233 (1999).
16. Forner, A., Reig, M. & Bruix, J.  $\alpha$ -fetoprotein for hepatocellular carcinoma diagnosis: The demise of a brilliant star. *Gastroenterology* **137**, 26–29 (2009).
17. Shaath, H., Toor, S., Nair, V. S., Elkord, E. & Alajez, N. M. Transcriptomic analyses revealed systemic alterations in gene expression in circulation and tumor microenvironment of colorectal cancer patients. *Cancers* **11**, 1994 (2019).
18. Jing, Y. *et al.* Tumor necrosis factor- $\alpha$  promotes hepatocellular carcinogenesis through the activation of hepatic progenitor cells. *Cancer Lett.* **434**, 22–32 (2018).
19. Hu, Z. *et al.* IL-17 Activates the IL-6/STAT3 signal pathway in the proliferation of hepatitis B virus-related hepatocellular carcinoma. *Cell. Physiol. Biochem.* **43**, 2379–2390 (2017).
20. Zhang, J. P. *et al.* Increased intratumoral IL-17-producing cells correlate with poor survival in hepatocellular carcinoma patients. *J. Hepatol.* **50**, 980–989 (2009).
21. Azmi, S., Sun, H., Ozog, A. & Taneja, R. mSharp-1/DEC2, a basic helix-loop-helix protein functions as a transcriptional repressor of E box activity and Stra13 expression. *J. Biol. Chem.* **278**, 20098–20109 (2003).
22. Cook, M. E., Jarjour, N. N., Lin, C. C. & Edelson, B. T. Transcription factor Bhlhe40 in immunity and autoimmunity. *Trends Immunol.* **41**, 1023–1036 (2020).
23. Li, C. *et al.* The transcription factor Bhlhe40 programs mitochondrial regulation of resident CD8+ T cell fitness and functionality. *Immunity* **51**, 491–507 (2019).
24. Kiss, Z., Mudryj, M. & Ghosh, P. M. Non-circadian aspects of BHLHE40 cellular function in cancer. *Genes Cancer* **11**, 1–19 (2020).
25. Sethuraman, A. *et al.* BHLHE40 confers a pro-survival and pro-metastatic phenotype to breast cancer cells by modulating HBEGF secretion. *Breast Cancer Res.* **20**, 117 (2018).
26. Jia, Y. *et al.* DEC1 is required for anti-apoptotic activity of gastric cancer cells under hypoxia by promoting Survivin expression. *Gastric Cancer* **21**, 632–642 (2018).
27. Jia, Y. F. *et al.* Differentiated embryonic chondrocyte-expressed gene 1 is associated with hypoxia-inducible factor 1 $\alpha$  and Ki67 in human gastric cancer. *Diagn. Pathol.* **8**, 37 (2013).
28. Ma, W., Shi, X., Lu, S., Wu, L. & Wang, Y. Hypoxia-induced overexpression of DEC1 is regulated by HIF-1 $\alpha$  in hepatocellular carcinoma. *Oncol. Rep.* **30**, 2957–2962 (2013).
29. Ding, Z. N. *et al.* Effects of hypoxia-inducible factor-1 $\alpha$  and hypoxia-inducible factor-2 $\alpha$  overexpression on hepatocellular carcinoma survival: A systematic review with meta-analysis. *J. Gastroenterol. Hepatol.* <https://doi.org/10.1111/jgh.15395> (2021).
30. Xiong, J. *et al.* The anti-metastatic effect of 8-MOP on hepatocellular carcinoma is potentiated by the down-regulation of bHLH transcription factor DEC1. *Pharmacol. Res.* **105**, 121–133 (2016).
31. Murakami, K. *et al.* DEC1 promotes hypoxia-induced epithelial-mesenchymal transition (EMT) in human hepatocellular carcinoma cells. *Biomed. Res.* **38**, 221–227 (2017).
32. Shi, X.-H. *et al.* DEC1 nuclear expression: A marker of differentiation grade in hepatocellular carcinoma. *World J. Gastroenterol.* **17**, 2037 (2011).
33. Sofer, A., Lei, K., Johannessen, C. M. & Ellisen, L. W. Regulation of mTOR and cell growth in response to energy stress by REDD1. *Mol. Cell. Biol.* **25**, 5834–5845 (2005).
34. Yoshida, T. *et al.* Rtp801, a suppressor of mTOR signaling, is an essential mediator of cigarette smoke-induced pulmonary injury and emphysema. *Nat. Med.* **16**, 767–773 (2010).
35. Tirado-Hurtado, I., Fajardo, W. & Pinto, J. A. DNA damage inducible transcript 4 gene: The switch of the metabolism as potential target in cancer. *Front. Oncol.* **8**, 106 (2018).
36. Britto, F. A., Dumas, K., Giorgetti-Peraldi, S., Ollendorff, V. & Favier, F. B. Is REDD1 a metabolic double agent? Lessons from physiology and pathology. *Am. J. Physiol. Cell Physiol.* **319**, C807–C824 (2020).
37. Pinto, J. A. *et al.* In silico evaluation of DNA Damage Inducible Transcript 4 gene (DDIT4) as prognostic biomarker in several malignancies. *Sci. Rep.* **7**, 1526 (2017).
38. Jiang, C., Liu, X., Wang, M., Lv, G. & Wang, G. High blood miR-802 is associated with poor prognosis in HCC patients by regulating DNA damage Response 1 (REDD1)-mediated function of T cells. *Oncol. Res.* **27**, 1025–1034 (2019).
39. Bruix, J. & Sherman, M. Management of hepatocellular carcinoma: An update. *Hepatology* **53**, 1020 (2011).
40. Forner, A., Reig, M. & Bruix, J. Hepatocellular carcinoma. *The Lancet* **391**, 1301–1314 (2018).
41. Andrews, S. *FastQC: A Quality Control Tool for High Throughput Sequence Data.* <https://www.bioinformatics.babraham.ac.uk/projects/fastqc> (2010).
42. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
43. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
44. Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
45. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
46. Raudvere, U. *et al.* g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).

## Acknowledgements

We would like to thank all of members in Center of Excellence in Hepatitis and Liver Cancer, Faculty of Medicine, Chulalongkorn University for their efforts in conducting the experiments.

## Author contributions

P.K. and N.C. contributed to designing the study and concept of the experiment. S.S., N. P and P.S. contributed to clinical data and specimen collection. P.K. and N.C. performed an experiment. P.K., C.A. and N.C. contributed

to data analysis and interpretation of the data. P.K. and N.C. wrote the manuscript. I.N. and P.T. contributed to editing the manuscript. All authors read and approved the final manuscript.

### Funding

This research was supported by the Thailand Research Fund (RTA6280004), the Royal Golden Jubilee Ph.D. Program (Grant No. PHD/0151/2558) and Center of Excellence in Hepatitis and Liver Cancer, Faculty of Medicine, Chulalongkorn University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-90515-w>.

**Correspondence** and requests for materials should be addressed to N.C. or P.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021